

Теория экспериментов с конечными автоматами

2-е издание, исправленное

Сперанский Д.В.

Национальный Открытый Университет "ИНТУИТ"
2016

УДК 519.713(076.6)

ББК 8

С71

Лекции по теории экспериментов с конечными автоматами / Сперанский Д.В. - М.: Национальный Открытый Университет "ИНТУИТ", 2016 (Основы информационных технологий)

ISBN 978-5-9963-0268-0

Конечные автоматы представляют собой удобные и адекватные математические модели, широко применяющиеся для описания структур и процессов функционирования цифровой аппаратуры, при разработке программных систем и трансляторов и во многих других предметных областях.

В данном курсе лекций излагаются результаты теории экспериментов с автоматами, востребованные при решении задач технической диагностики дискретных устройств, кодирования и декодирования информации, задач распознавания, расшифровки и идентификации и т.п.

(с) ООО "ИНТУИТ.РУ", 2010-2016

(с) Сперанский Д.В., 2010-2016

Эксперименты с автоматами, имеющими взвешенный входной алфавит

В лекции приводятся определения основных понятий теории экспериментов: синхронизирующих, установочных и диагностических последовательностей, конструкции дерева преемников. Описывается конструкция синхронизирующего (установочного, диагностического) дерева, используемая для построения минимальной по весу синхронизирующей (установочной, диагностической) последовательности, а также процедура ее построения.

Основные определения

Введем некоторые основные понятия теории автоматов, которые используются в дальнейшем изложении.

Под автоматом Мили будем понимать пятерку $A = (S, X, Y, \delta, \lambda)$, где S, X, Y - конечные множества состояний, входной и выходной алфавиты соответственно, δ и λ - функции переходов и выходов, задающие отображения $\delta : S \times X \rightarrow S, \lambda : S \times X \rightarrow Y$. Через S_0 обозначим подмножество множества состояний S , которое назовем множеством допустимых начальных состояний автомата A .

Предполагается, что множества X и Y автомата A конечны (такие автоматы называются конечными).

Через X^* обозначим множество всех последовательностей (слов) конечной длины в алфавите X . Через e обозначим пустое слово (нулевой длины) в любом алфавите. Полагаем, что $e \in X^*$. Распространим функции переходов и выходов автомата на множество X^* следующим образом:

$$\delta(s, e) = s, \delta(s, px) = \delta(\delta(s, p), x), \lambda(s, e) = e, \lambda(s, px) = \lambda(s, p)\lambda(\delta(s, p), x)$$

где $s \in S, x \in X, p \in X^*$. Последовательность выходов автомата $\lambda(s, p)\lambda(\delta(s, p), x)$ получена с помощью операции конкатенации

(приписывания) двух слов: $\lambda(s, p)$ и $\lambda(\delta(s, p), x)$.

Автомат Мили A может быть задан в виде таблиц переходов и выходов (иногда они совмещаются в одну таблицу), с помощью матриц и, наконец, с помощью графа. Форма представления автомата в виде графа является наиболее наглядной. В дальнейшем изложении мы будем пользоваться в основном графовой и табличной формами представления автомата, которые подробно описаны в [18]. Напомним, что множеством вершин графа переходов автомата A являются состояния из S , а множество дуг есть совокупность четверок (s, x, y, t) , где $t = \delta(s, x)$, $y = \lambda(s, x)$. Пара (x, y) называется отметкой дуги, s - ее началом, t - концом.

Пусть $T \subseteq S$, тогда полагаем, что $\delta(T, x) = Y\delta(s, x)$ и $\lambda(T, x) = Y\lambda(s, x)$, где объединение выполняется по всем $s \in T$.

Состояние t назовем достижимым из s , если $t \in \delta(s, p)$ для некоторого $p \in X^*$.

Автомат A назовем сильно связным, если его граф переходов сильно связан, т. е. из любого состояния автомата достижимы все его состояния. Состояние s называется переходящим в автомате A , если s не является концом ни одной дуги автомата, т. е. s не достижимо ни из одного состояния, отличного от s . Состояние s называется тупиковым в автомате A , если из s не достижимо ни одно состояние, отличное от s .

Каждому входному символу x_i из алфавита X автомата A поставим в соответствие некоторое положительное число $w(x_i)$ и назовем его весом символа x_i . Далее такой алфавит будем называть взвешенным. Весом входной последовательности $p = x_{i_1}, x_{i_2}, \dots, x_{i_n} \in X^*$ назовем число

$$w(p) = \sum_{j=1}^n w(x_{i_j})$$

Следуя традиции классической теории экспериментов с автоматами,

будем различать безусловные и условные эксперименты [18]. Безусловный эксперимент предполагает подачу на вход автомата такой последовательности, которая определена заранее, т. е. до начала эксперимента. При проведении условного эксперимента на вход автомата подается две или более подпоследовательности, причем каждая последующая подпоследовательность, кроме первой, формируется на основании реакций, вызываемых предыдущими подпоследовательностями.

Преимущества условных экспериментов состоят в том, что они, во-первых, иногда могут оказаться короче безусловных и, во-вторых, в некоторых случаях их строить легче, чем безусловные.

Один автомат будем называть копией другого, если они имеют одинаковые графы переходов и перед началом эксперимента оба они находятся в одном и том же состоянии. По числу требуемых для эксперимента копий автомата различают простые, когда требуется лишь один экземпляр автомата, и кратные эксперименты, когда требуется более чем один экземпляр автомата.

В дальнейшем изложении мы в основном будем иметь дело с простыми безусловными экспериментами.

В классической теории экспериментов с автоматами были исследованы две задачи, получившие наименование диагностической и установочной. В обеих задачах предполагается известным автомат, т. е., например, его граф переходов, и множество S_0 его допустимых начальных состояний, т. е. истинное начальное (стартовое) состояние. Это состояние есть одно из состояний множества S_0 . Диагностическая задача состоит в отыскании начального (стартового) состояния заданного автомата, а установочная - в установке автомата в известное состояние. Эксперименты, решающие перечисленные задачи, принято называть диагностическими и установочными соответственно. Интерес к этим задачам объясняется возможностью приложений результатов их решения к контролю и диагностике реальных цифровых устройств.

Проведение установочных и диагностических экспериментов с автоматами требует наличия специальных входных последовательностей, которые мы также будем именовать

установочными (УП) и диагностическими (ДП) последовательностями соответственно.

Дадим теперь формальные определения этих последовательностей. Входная последовательность $p = x_{i_1}, x_{i_2}, \dots, x_{i_n}$ называется установочной последовательностью для автомата A и множества S_0 его допустимых начальных состояний, если

$$\forall s_1, s_2 \in S_0 \lambda(s_q, p) \rightarrow \delta(s_1, p) = \delta(s_2, p) \quad (1.1)$$

Содержательно это определение означает, что по наблюдаемой реакции на УП автомата A , находящегося в неизвестном для наблюдателя начальном состоянии, всегда однозначно можно определить его конечное состояние.

Входная последовательность $p = x_{i_1}, x_{i_2}, \dots, x_{i_n}$ называется диагностической последовательностью для автомата A и множества S_0 его допустимых начальных состояний, если

$$\forall s_1, s_2 \in S_0 \lambda(s_1, p) = \lambda(s_2, p) \rightarrow s_1 = s_2 \quad (1.2)$$

Содержательно это определение означает, что по наблюдаемой реакции на ДП автомата A , находящегося в неизвестном для наблюдателя начальном состоянии, всегда однозначно можно определить это начальное (стартовое) состояние.

Определим еще один тип последовательности, которую назовем синхронизирующей (СП). Входная последовательность $p = x_{i_1}, x_{i_2}, \dots, x_{i_n}$ называется синхронизирующей последовательностью для автомата A и множества S_0 его допустимых начальных состояний, если

$$\forall s_1, s_2 \in S_0 \delta(s_1, p) = \delta(s_2, p) \quad (1.3)$$

Содержательно это определение означает, что после подачи на автомат A СП этот автомат оказывается в одном и том же известном наблюдателю конечном состоянии независимо от того, из какого начального состояния он стартовал.

Очевидно, что СП по существу является УП, поскольку после ее подачи автомат оказывается в известном состоянии, хотя никакого наблюдения реакций производить не нужно. Другими словами, подачу на вход автомата СП можно трактовать как синхронизирующий эксперимент, но эксперимент "вырожденный", поскольку он не требует наблюдения реакции автомата.

Ниже мы займемся разработкой методов построения СП, УП и ДП в предположении, что автомат A имеет взвешенный входной алфавит X . При этом нас будут интересовать все три типа последовательностей с минимальным весом. Как уже было упомянуто выше, известные ранее методы построения минимальных по длине последовательностей, описанные в [18], в рассматриваемой ситуации оказываются непригодными. Вместе с тем базисная конструкция дерева преемников, использованная в [18] для синтеза УП и ДП, минимальных по длине, оказывается вполне работоспособной и в нашем случае, однако она, что вполне естественно, требует соответствующих изменений.

Понятно, что, располагая методами построения минимальных по весу СП, УП и ДП для автоматов со взвешенным входным алфавитом, всегда можно построить соответствующие последовательности, минимальные по длине, положив вес каждого символа алфавита X равным единице. Таким образом, описываемые ниже методы построения упомянутых последовательностей являются более общими, чем известные ранее.

Заметим, что в этой лекции мы не касаемся вопросов существования упомянутых выше трех типов последовательностей, поскольку для автоматов Мили они подробно отражены в [18]. Легко сообразить, что "взвешенность" алфавита никак не влияет на условия существования соответствующих последовательностей.

Синтез синхронизирующих последовательностей, минимальных по весу

В основе построения СП, УП и ДП, минимальных по весу, лежит так называемое дерево преемников, формируемое для автомата A и множества S_0 его допустимых начальных состояний. Оно незначительно отличается от аналогичной структуры, определенной в

[18]. Остановимся на способе его построения.

Под σ -множеством автомата A понимается любая конечная совокупность состояний A , не все из которых обязательно различны. Если σ -множество содержит один элемент, то оно называется простым. Если σ -множество содержит два или более одинаковых элементов, оно именуется кратным. σ -множество однородно, если все его элементы совпадают друг с другом.

Назовем A -группой множество, состоящее из σ -множеств, общее число элементов которых равно $m = |S_0|$. A -группа называется простой (однородной), если все σ -множества в ней просты (однородны).

Пусть G есть A -группа, состоящая из σ -множеств $g_1, g_2, \dots, g_r, x_{i_1}, x_{i_2}, \dots, x_{i_j}$ -преемник G есть другая A -группа, построенная по следующим правилам:

1. σ -множество $g_i = (i = \overline{1, r})$ разбиваем на такие подмножества, каждое из которых содержит все те состояния из g_i , которые вырабатывают одинаковую реакцию на входную последовательность $x_{i_1}, x_{i_2}, \dots, x_{i_j}$. Каждое полученное подмножество далее интерпретируется как σ -множество, а совокупность всех таких подмножеств интерпретируется как A -группа G' .
2. В σ -множествах из G' заменяем каждое состояние его преемником относительно входной последовательности $x_{i_1}, x_{i_2}, \dots, x_{i_j}$.

Получаемая в результате такого построения A -группа является $x_{i_1}, x_{i_2}, \dots, x_{i_j}$ -преемником G .

Определим теперь структуру, называемую деревом преемников для заданного автомата A и множества допустимых начальных состояний S_0 . Она состоит из ветвей, расположенных в последовательных уровнях, нумерация которых начинается с нуля. Нулевой уровень дерева содержит единственную, называемую начальной, ветвь. Условимся

считать, что она связана с A -группой S_0 . Если входной алфавит $X = \{x_1, \dots, x_p\}$ автомата A содержит p символов, то каждая ветвь в k -м уровне ($k \geq 0$) расщепляется на p ветвей, представляющих символы x_1, x_2, \dots, x_p соответственно и являющихся ветвями $(k+1)$ -го уровня. Ветвь x_i $(k+1)$ -го уровня считается связанной с A -группой, являющейся x_i -приемником соответствующей A -группы уровня k . Будем называть вершины простыми, кратными или однородными, если они связаны с соответствующими A -группами.

Из описанного метода построения вытекает, что дерево преемников является бесконечной структурой. Для синтеза СП, УП и ДП нами будут из упомянутой бесконечной структуры выделяться конечные деревья за счет введения специальных правил обрыва ветвей. Каждая такая оборванная ветвь связана с вершиной дерева, которую мы будем называть листом, следуя терминологии, принятой в теории графов.

Для построения минимальной по весу СП используем граф, который назовем синхронизирующим деревом (СД). Основой для его построения является описанное выше дерево преемников.

Вершинами СД являются A -группы дерева преемников. Корень СД представляет собой A -группу S_0 нулевого уровня. Вершиной 1-го уровня СД, связанной с x_i -ветвью, является A -группа S , представляющая собой x_i -приемник A -группы S_0 , при этом из вершины S_0 в S проводится дуга, помеченная символом x_i . Аналогично строятся вершины СД всех последующих уровней и соединяющие их дуги. Каждой вершине S в СД поставим в соответствие флажок (на рисунках он будет обозначаться прямоугольником), в него будет вписано число, равное сумме весов входных символов последовательности, которая соответствует пути по СД, ведущему из корня в вершину S .

Условимся теперь об обозначениях вершин СД. Если вершине S СД соответствует A -группа g_1, g_2, \dots, g_r , где $g_i (i = \overline{1, r})$ - σ -множества, то для ее обозначения будем использовать множество состояний автомата A , равное

$$\bigcup_{1 \leq i \leq r} g_i.$$

Для того чтобы выделить из дерева преемников конечное СД, введем правила обрыва ветвей. Правила эти должны определять момент, когда некоторая вершина дерева преемников становится листом, то есть из нее в процессе дальнейшего построения не исходит никаких дуг.

При построении СД автомата A с множеством S_0 допустимых начальных состояний вершина S k -го уровня становится листом, если выполняется хотя бы одно из следующих условий:

1. $|S| = 1$;
2. на уровнях, предшествующих k -му, имеется такая вершина S^* , что $S = S^*$, но значение флага вершины S больше значения флага вершины S^* ;
3. на уровнях, предшествующих k -му, имеется такая вершина S^* , что $|S^*| = 1$, а значение флага вершины S больше значения флага вершины S^* .

Проиллюстрируем процесс построения СД для автомата, заданного табл. 1.1, считая, что $S_0 = \{1, 2, 3\}$, $w(\alpha) = 1$, $w(\beta) = 2$, $w(\gamma) = 20$. Этот граф приведен на рис. 1.1. В овале, изображающем вершину, перечислены состояния множества, соответствующего этой вершине.

Таблица 1.1.

Переходы-
выходы
автомата

s\х	α	β	γ
1	1/0	1/1	5/0
2	3/1	2/0	4/0
3	2/1	4/1	4/0
4	5/1	1/1	5/1

5 3/0 2/0 4/1

Так, вершина $\{1, 2, 3\}$ первого уровня, две вершины $\{4, 5\}$ второго уровня, вершины $\{4, 5\}$ и $\{1, 2\}$ третьего уровня и т. д. являются листьями в силу пункта 2 правил обрыва. Вершины $\{1\}$ и $\{5\}$ пятого уровня, вершина $\{4\}$ третьего уровня являются листьями в силу пункта 1 правил обрыва. Две вершины $\{4, 5\}$ четвертого уровня, имеющие значение флага 43, являются листьями в силу пункта 3 тех же правил.

Из способа построения СД вытекает, что последовательность входных символов, соответствующая пути по СД из корня в одну из вершин S , где $|S| = 1$, является СП для заданного автомата.

Понятно, что построение минимальной по весу СП сводится к следующей простой процедуре. Среди всех вершин S СД, у которых $|S| = 1$, отыскивается вершина с минимальным значением флага $w(s)$ и в СД находится путь в нее из корня дерева. Очевидно, что соответствующая этому пути последовательность символов и является СП, имеющей минимальный вес, равный упомянутому значению флага $w(s)$.

Так, в СД на [рис.1.1](#) имеется 4 таких вершины, из которых вершина $\{1\}$ пятого уровня имеет минимальное значение флага, равное 27.

Поскольку пути из корня в эту вершину соответствует входная последовательность $\gamma\beta\alpha\beta\beta$, то она и является минимальной по весу СП. Заметим, кстати, что минимальной по длине СП для этого автомата и того же множества допустимых начальных состояний является последовательность $\gamma\alpha\gamma$, соответствующая пути по СД из корня в вершину $\{4\}$ третьего уровня.

Приведенный пример показывает, что в общем случае СП для автомата может быть минимальной по весу, но не являться минимальной по длине.

простая вершина S^* , значение флажка которой меньше значения флажка вершины S .

Заметим, что для превращения вершины S в лист достаточно выполнения одного из четырех перечисленных пунктов.

Прокомментируем эти правила. Пункт 3 правил означает, что в УД уже имеется более "легкий" по весу отрезок пути, приводящий к некоторой A -группе, чем путь, приведший к вершине S с точно такой же A -группой. Понятно, что при продолжении построения УД из вершины S принципиально возможно построить некоторую УП, но вес ее заведомо не может быть минимальным. Именно это является причиной того, что дальнейшее продолжение пути из S_0 в S не производится, т. е. этот путь обрывается.

Пункт 4 правил означает, что если в УД на некотором уровне меньше k -го уже имеется путь, приведший к синтезу УП определенного веса, то продолжение поиска УП на пути от S_0 к S и далее если и даст положительный результат, то заведомо приведет к синтезу новой УП большего веса, чем ранее найденная. Поскольку такая УП не представляет интереса, то путь из S_0 в S продолжать не имеет смысла.

Легко сообразить, что если некоторая вершина S дерева преемников становится листом вследствие выполнения пунктов 1 и/или 2 правил, то последовательность входных символов, соответствующая пути в УД из корня в вершину S , является УП. Отсюда вытекает следующая процедура построения УП с минимальным весом. Среди всех вершин УД, являющихся листьями, которым соответствуют однородные и/или простые A -группы состояний рассматриваемого автомата, отыскивается вершина S с минимальным значением $w(s)$ флажка. Входная последовательность, соответствующая пути из корня УД в вершину S , и является УП с минимальным весом, равным $w(s)$.

Проиллюстрируем процесс построения УП для автомата, заданного табл. 1.1, полагая, что $S_0 = \{1, 2, 3, 4, 5\}$, $w(\alpha) = 10$, $w(\beta) = 20$, $w(\gamma) = 5$. УД для этого

автомата изображено на рис.1.2. Вершины УД, являющиеся однородными или простыми, отмечены знаком "+". Всего таких вершин в УД имеется 16, из которых вершина $\{5, 5; 3; 5; 3\}$ второго уровня имеет минимальное значение флагажка, равное 15. Поскольку пути в УД из корня в эту вершину соответствует входная последовательность $\gamma\alpha$, то она и является минимальной по весу УП. Заметим, что в данном примере эта УП является одновременно и минимальной по длине, однако в общем случае этот факт места не имеет.

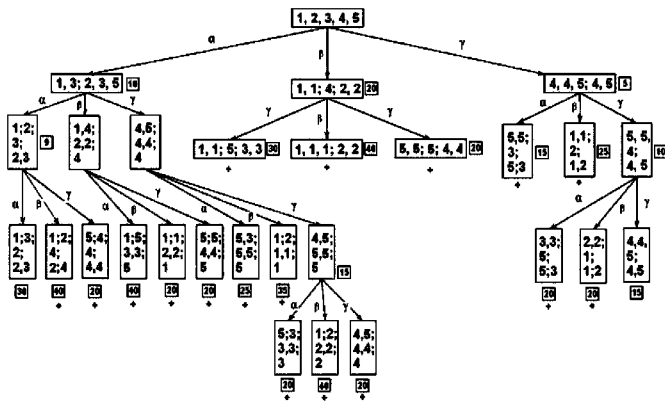


Рис. 1.2. Установочное дерево

Синтез диагностических последовательностей, минимальных по весу

Для построения минимальной по весу ДП используем граф, который назовем диагностическим деревом (ДД). Его построение осуществляется точно так же, как и построение УД, но при этом используются несколько иные правила обрыва ветвей дерева приемников.

При построении ДД автомата A с множеством S_0 допустимых начальных состояний вершина S_k -го уровня становится листом, если

выполняется хотя бы одно из следующих условий:

1. она является простой;
2. она является кратной;
3. на уровнях, предшествующих k -му, имеется такая вершина S^* , что $S = S_0$, но значение флажка вершины S больше значения флажка вершины S^* ;
4. на уровнях, предшествующих k -му, имеется простая вершина S^* , значение флажка которой меньше значения флажка вершины S .

Прокомментируем эти правила. Выполнение пункта 1 этих правил означает, что рассматриваемый автомат имеет ДП, которая соответствует входной последовательности на пути в ДД, ведущем из его корня в простую вершину S . Выполнение пункта 2 правил означает, что существует такая последовательность, подача которой на вход автомата, находящегося в различных состояниях из множества S_0 , вызывает одинаковые реакции автомата и переводит его в одно и то же состояние. Понятно, что в этом случае любое продолжение упомянутой входной последовательности заведомо не может привести к построению ДП и поэтому такое продолжение не имеет смысла. Что касается пунктов 3 и 4, то комментарии, данные к таким же пунктам в разделе 1.3, в полной мере относятся и к нашему случаю.

Из сказанного выше вытекает следующая процедура построения минимальной по весу ДП. Среди всех вершин ДД, являющихся листьями, которым соответствуют простые A -группы состояний рассматриваемого автомата, отыскивается вершина S с минимальным значением $w(s)$ флажка. Входная последовательность, соответствующая пути из корня ДД в вершину S , и является ДП с минимальным весом, равным $w(s)$.

Проиллюстрируем процесс построения ДП для автомата, заданного табл. 1.1, полагая множество состояний S_0 и веса символов входного алфавита теми же, что и в примере предыдущего раздела 1.3. ДД для этого автомата изображено на рис.1.3.

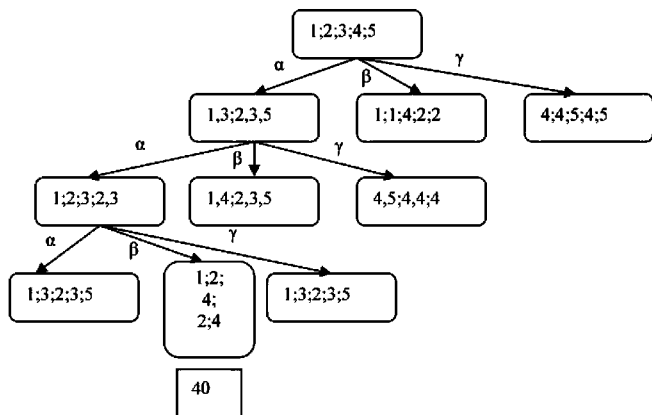


Рис. 1.3. Диагностическое дерево

Из этого рисунка видно, что все вершины ДД, кроме вершины третьего уровня $\{1, 2, 4, 2, 4\}$ со значением флажка, равным 40, являются листьями в силу пункта 2 правил, приведенных выше, поскольку они содержат кратные σ -множества. Единственная вершина - лист $\{1, 2, 4, 2, 4\}$ - является простой, и потому входная последовательность $\alpha\alpha\beta$, соответствующая пути из корня ДД в эту вершину, является минимальной ДП, имеющей вес, равный 40.

Заметим, что та же входная последовательность является одновременно минимальной и по длине.

Вопросы и упражнения

1. Дайте определение автомата Мили.
2. Что означает "взвешенность" входного алфавита автомата?
3. Что такое эксперимент с автоматом?
4. Приведите классификацию экспериментов с автоматами.
5. Дайте определения синхронизирующих, установочных и

диагностических последовательностей и поясните их содержательный смысл.

6. Опишите структуру дерева преемников для автомата со взвешенным входным алфавитом.
7. Каковы правила обрыва ветвей в дереве преемников при построении синхронизирующего (установочного, диагностического) дерева?
8. Опишите процедуру построения минимальной по весу синхронизирующей (установочной, диагностической) последовательности с использованием конструкции синхронизирующего (установочного, диагностического) дерева.
9. Для автомата, заданного приведенной ниже таблицей переходов-выходов, постройте минимальную по весу синхронизирующую последовательность, если $w(\alpha) = 10$, $w(\beta) = 1$ и множество допустимых начальных состояний есть $\{1, 2, 3, 4\}$.

x \ s	α	β
1	4,1	3,0
2	1,0	4,0
3	2,0	3,1
4	3,0	1,1

10. Для автомата из предыдущего упражнения постройте УП с минимальным весом, если $w(\alpha) = 10$, $w(\beta) = 20$ и множество допустимых начальных состояний есть $S_0 = \{1, 2, 3\}$.
11. Для автомата из упражнения 9 постройте ДП с минимальным весом, если $w(\alpha) = 10$, $w(\beta) = 20$ и множество $S_0 = \{1, 2, 3\}$.

Синтез экспериментов методами динамического программирования

Описываются рекурсивные методы построения графов синхронизации, установки и диагностики автомата. Показано, что задачи построения минимальных по весу СП и УП сводятся к задачам выбора наискорейшего пути по сети дорог, эффективно решаемым методами динамического программирования из теории оптимального управления. Задача синтеза минимальной по весу ДП по графу диагностики автомата сводится к вырожденной в данном случае задаче поиска минимального пути между двумя вершинами.

Синтез синхронизирующей последовательности

Решение задачи синтеза СП, минимальной по весу, будет состоять из двух этапов, первый из которых заключается в построении по таблице переходов-выходов автомата Мили графа, который будет именоваться графом синхронизации автомата (ГСА), а второй - в поиске в нем минимальной по весу СП методом динамического программирования.

Построение ГСА осуществляется рекурсивно. Вначале в число его вершин включается вершина S_0 , соответствующая множеству допустимых начальных состояний автомата. Вершину S_0 поместим на вертикальную прямую, называемую далее линией уровня, и присвоим ей номер 1. Определим далее подмножества

$S_i^{(2)} = \sigma(S_j, x_i), i = \overline{S_0, x_i}$, и все полученные попарно различные подмножества, отличные от S_0 , также включим в число вершин графа.

Ясно, что $|S_i^{(2)}| \leq |S_0|$. Упорядочим множества $S_i^{(2)}$ по мощности и будем располагать соответствующие им вершины графа на второй, третьей и т. д. линиях уровня по следующему принципу: на 2-ю линию

уровня поместим все те вершины $|S_i^{(2)}| = |S_0| - 1$, у которых $S_i^{(2)}$, на 3-ю линию - все те вершины, у которых $|S_i^{(2)}| = |S_0| - 2$ и т. д.

Затем вершина S_0 соединяется с каждой вершиной $S_{i_0}^{(2)}$ дугой с символом $x \in X$, если $S_{i_0}^{(2)} = \sigma(S_0, x)$. В частности, если

существует $x \in X$, такой, что $\sigma(S_0, x) = S_0$, то у вершины S_0 проводится петля с символом x . Пусть $S_1^{(2)}, \dots, S_a^{(2)}$ - все попарно различные вершины, появившиеся в графе на втором шаге. Определим теперь подмножества $S_i^{(3)} = \sigma(S_i^{(2)}, x_j)$ для всех $x_j \in X$ и для всех $S_i^{(2)}, i = \overline{1, a}$. Все полученные попарно различные подмножества, отличные от появившихся на втором шаге, также включим в число вершин графа. Затем упорядочиваем их по мощности, размещаем их на соответствующих линиях уровня, как это описано выше, и соединяем все вершины $S_i^{(2)}$ с вершинами $S_i^{(3)}$ соответствующими дугами по ранее определенному правилу. В силу того что $|S_i^{(3)}| \leq |S_i^{(2)}|$, все упомянутые дуги либо связывают между собой вершины, находящиеся на одной и той же линии уровня (в случае, когда $|S_{i_0}^{(2)}| = |S_{j_0}^{(3)}|$), либо идут от вершин $S_i^{(2)}$ к вершинам $S_i^{(3)}$, находящимся на линиях уровня с большими номерами (в случае, когда $|S_{i_0}^{(2)}| > |S_{j_0}^{(3)}|$).

Описанный процесс продолжается далее аналогичным образом до тех пор, пока на очередном шаге построения ГСА будут отсутствовать вершины, отличные от тех, что уже появились в графе на предшествующих шагах. Понятно, что построение ГСА обязательно завершится, поскольку число всевозможных подмножеств множества S состояний автомата конечно, причем число линий уровня ГСА не превосходит величины $n = |s|$.

В полученном графе осуществим параллельный перенос всех линий уровня вправо так, чтобы 1-я из них оказалась между вершинами 1-го и 2-го уровней, 2-я - между вершинами 2-го и 3-го уровней и т. д. В результате такого переноса смещенные линии уровня будут пересекаться дугами, соединяющими между собой вершины i -го и $(i+1)$ -го уровней ($i = 1, 2, \dots$). Все упомянутые точки пересечения будем также считать вершинами графа промежуточными и каким-либо образом их перенумеруем. Именем каждой из появившихся промежуточных вершин будем считать присвоенный ей при

упомянутой нумерации номер.

Удалим из этого графа все петли и поставим в соответствие каждой дуге графа число, трактуемое как длина дуги, равное весу помеченного ею входного символа. С каждой парой дуг графа $(s_1, l_1)(l_1, s_1)$ (здесь l_1 - номер некоторой промежуточной вершины), полученной из дуги (s_1, s_2) с пометкой x_l за счет переноса линий уровня, поступим следующим образом. Дугу (s_1, l_1) пометим пустым (\emptyset) символом и поставим ей в соответствие число 0, а дуге (l_1, s_2) оставим пометку x_i ; дуги (s_1, s_2) и поставим ей в соответствие число, равное весу символа x_l . Полученный в результате граф будем называть ГСА.

Из самого способа построения ГСА вытекает, что заданный автомат будет иметь СП тогда и только тогда, когда в соответствующем ему ГСА имеются вершины, помеченные символами одноэлементных подмножеств множества S состояний автомата. При этом последовательность входных символов, соответствующая пути по ГСА из вершины S_0 в одну из одноэлементных вершин, очевидно, будет являться СП.

Для иллюстрации на рис.2.1 изображен граф синхронизации автомата, заданного табл. 1.1. Граф построен в предположении, что $S_0 = \{1, 2, 3, 4, 5\}$.

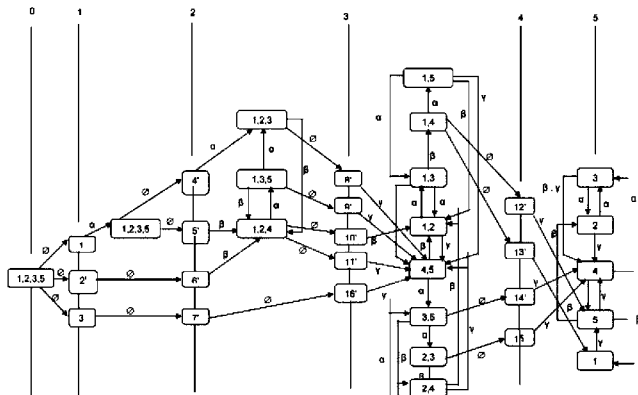


Рис. 2.1.

Далее покажем, что рассматриваемая нами задача синтеза СП может быть сведена к задаче выбора наискорейшего пути, решаемой методом динамического программирования [15] из теории оптимального управления. Последняя задача решается поэтапно с помощью процесса перемещения с одной опорной прямой на соседнюю. В нашем случае роль таких опорных прямых играют линии уровня, введенные нами при построении ГСА. Напомним, что из содержательного смысла задачи о выборе наискорейшего пути сетка дорог, ведущая из начального в конечный пункт, имеет некоторые особенности. Так, во-первых, в ней не должно быть петель, поскольку движение по ним не приближает нас к конечному пункту, а, во-вторых, все дуги, соответствующие участкам дорог от пунктов, расположенных на i -й опорной прямой, должны либо вести к пунктам на $(i + 1)$ -й опорной прямой, либо соединять между собой пункты на i -й опорной прямой, но не должно быть "обратных" дуг.

Понятно, что если веса дуг в ГСА интерпретировать как время движения по ним, а сам ГСА как сеть дорог, то задача о синтезе СП минимального веса эквивалентна задаче о выборе наискорейшего пути

из вершины S_0 ГСА в одну из одноэлементных вершин того же графа. Действительно, ГСА удовлетворяет по построению всем требованиям, предъявляемым к сети дорог в упомянутой задаче. В качестве примера рассмотрим применение метода динамического программирования к построению минимальной по весу СП, переводящей автомат, заданный табл. 1.1, в конечное состояние 4, предполагая что $w(\alpha) = 1$, $w(\beta) = 2$, $w(\gamma) = 3$, а множество допустимых начальных состояний есть $S_0 = \{1, 2, 3, 4, 5\}$.

Вначале введем две дополнительные линии уровня: нулевую, на которой будет располагаться вершина S_0 ГСА, и μ -ю, на которой будут располагаться вершины ГСА, лежащие правее линии уровня с номером $\mu - 1$, где $\mu - 1$ есть максимальный номер уровня, появившийся в процессе построения ГСА, описанном выше. В рассматриваемом примере максимальный номер уровня, появившийся при построении ГСА на рис.2.1, равен 4, поэтому на 5-й линии уровня будут расположены все одноэлементные вершины $\{1\}, \{2\}, \{3\}, \{4\}, \{5\}$, а на нулевом - вершина $\{1, 2, 3, 4, 5\}$.

Напомним, что в методе динамического программирования каждый шаг представляет собой перемещение с одной из опорных прямых (линии уровня) на соседнюю, более близкую к конечному пункту (в нашем примере к вершине $\{1, 2, 3, 4, 5\}$). При этом, как обычно, предполагается, что движение от шага к шагу допускается только в положительном направлении (от вершины $\{4\}$ к вершине $\{1, 2, 3, 4, 5\}$). Иными словами, после того как некоторый шаг пройден, возвращение обратно, в ту же полосу между двумя линиями уровня, не допускается.

В соответствии с числом линий уровня на рис.2.1 процесс перемещения из вершины $\{4\}$ в вершину $\{1, 2, 3, 4, 5\}$ разделим на пять шагов и начнем построение оптимального пути с последнего пятого шага.

Наметим на четвертой линии уровня все возможные пункты (вершины) нашего движения в момент окончания предпоследнего четвертого шага.

В нашем примере такими вершинами будут вершины $\{12'\}, \{3'\}, \{14'\}, \{15'\}$. Далее находим оптимальные по времени пути из этих вершин в вершину $\{4\}$. Заметим, что для поиска таких путей можно применить, например, известные алгоритмы Флойда и Данцига поиска всех кратчайших путей в ориентированном графе [34].

После завершения этого шага переходим к планированию четвертого шага. Для каждой из вершин $\{12'\}, \{3'\}, \{14'\}, \{15'\}$ теперь необходимо найти оптимальное управление, т. е. такой путь с 3-й линии уровня на 4-ю, который совместно с уже оптимизированным последним шагом дает возможность достигнуть вершины $\{4\}$ за минимальное время.

Чтобы найти это условное оптимальное управление, для каждой вершины на 3-й линии уровня необходимо перебрать всевозможные способы перехода на 4-ю линию уровня и время, которое требуется на этот переход, сложить с минимальным временем последнего шага. Из всех возможных путей выбирается тот, для которого это суммарное время минимально, и соответствующий путь отмечается. Понятно, что и на этом шаге для поиска кратчайших путей между соответствующими вершинами ГСА 3-й и 4-й линий уровня можно применить алгоритмы Флойда и Данцига.

В результате цепочки таких построений, перемещая шаг за шагом с одной линии уровня на другую, дойдем до исходной вершины S_0 . Для нее определим оптимальный путь на первую линию уровня. Таким образом, теперь мы располагаем всеми данными для построения оптимального пути, так как для каждой из намеченных вершин на линиях уровня известно оптимальное продолжение пути. Для нашего примера в результате завершения процесса будет найден оптимальный путь

$$\{1, 2, 3, 4, 5\} \xrightarrow{0(\emptyset)} \{3'\} \xrightarrow{0(\emptyset)} \{7'\} \xrightarrow{0(\emptyset)} \{16'\} \xrightarrow{3(\gamma)} \{4, 5\} \xrightarrow{I(\alpha)} \{3, 5\} \xrightarrow{0(\emptyset)} \{14'\} \xrightarrow{3(\gamma)} \{4\}$$

Ему соответствует последовательность входных сигналов $\emptyset\emptyset\emptyset\gamma\alpha\emptyset\gamma$. Если исключить из нее пустые сигналы, то получится искомая

минимальная по весу СП $\gamma\alpha\gamma$, переводящая заданный автомат из любого состояния в одно и то же конечное состояние 4.

Выбирая далее в качестве конечной вершины пути все остальные одноэлементные вершины ГСА, находящиеся на 5-й линии уровня, найдем описанным выше образом минимальные по весу СП, переводящие автомат в эти конечные состояния. Заметим, что в общем случае не все такие пути могут существовать, что говорит о невозможности перевода автомата в соответствующее конечное состояние, если он стартует из любого состояния множества S_0 . Наконец, выбрав из всех построенных условно оптимальных по весу СП минимальную, получим искомую СП. Легко убедиться, что в нашем примере такой минимальной по весу СП будет $\gamma\alpha\gamma$, построенная выше.

Синтез установочной последовательности

Решение задачи синтеза УП, минимальной по весу, будет состоять из двух этапов, первый из которых заключается в построении по таблице переходов-выходов автомата графа, называемого графом установки автомата (ГУА), а второй - в поиске на нем минимальной УП.

ГУА - это ориентированный граф, каждая вершина которого отмечается A - группой σ -множеств. Отметка вершины используется в качестве ее имени. Если A - группа, которой отмечена некоторая вершина S_1 ГУА, под воздействием входного сигнала $x \in X$ переходит в A - группу, которой отмечена вершина S_2 , то в ГУА из S_1 в S_2 проводится дуга, на которой ставится входной сигнал x . Правила преобразования одной A - группы в другую под воздействием входного сигнала x остаются такими же, как они определены выше.

Построение ГУА осуществляется рекурсивно и по существу совпадает с процессом построения установочного дерева из [18].

Вначале в число вершин графа включается вершина S_0 , где S_0 - множество допустимых начальных состояний. Вершину S_0 поместим на вертикальную прямую и назовем ее линией 1-го уровня. Далее строим все A - группы, в которые переходит A - группа S_0 под

воздействием всех символов входного алфавита X , включаем их в число вершин ГУА и помещаем на линию уровня с номером 2. Из вершины S_0 в вершины второго уровня проводим по упомянутому правилу дуги, помеченные соответствующими входными символами. Процесс построения ГУА продолжается далее аналогичным образом (как и в случае установочного дерева). При этом если на очередном шаге построения появилась вершина, соответствующая однородной A -группе, то такая вершина считается листом, т. е. при продолжении процесса построения ГУА из нее не будет больше исходить никаких дуг.

У каждой вершины S_i ГУА установим флажок (на рисунке он, как и ранее, изображается в виде квадрата); в него записывается число, равное сумме весов входных символов последовательности, которая ведет из S_0 в S_i . Теперь введем некоторые дополнительные правила, используемые при построении ГУА:

- если на очередной линии уровня появилась вершина S^* , а на одной из линий уровня с меньшим номером имеется вершина \tilde{S}^* , помеченная такой же A -группой, что и S^* , причем число во флажке у вершины S^* больше, чем во флажке у вершины \tilde{S}^* , то вершина S^* становится листом, а если она является непосредственным преемником вершины с такой же A -группой, то она удаляется из ГУА;
- если значение флажка у некоторой вершины S , появившейся на очередном шаге построения ГУА, больше, чем у какой-либо вершины, помеченной однородной A -группой, находящейся на предыдущих уровнях, то вершина S удаляется из ГУА.

Понятно, что описанный процесс построения ГУА всегда завершается за конечное число шагов. Проиллюстрируем построение ГУА для автомата, заданного табл. 1.1, считая, что

$S_0 = \{1, 2, 3, 4, 5\}$, $w(\alpha) = 1$, $w(\beta) = 2$, $w(\gamma) = 3$. Обратимся к рис.2.2, на котором изображен соответствующий ГУА. На третьем шаге построения ГУА, как это видно из рисунка, возникло 9 вершин, среди которых

$\{1, 2, 3, 2, 3\}$, $\{1, 4, 2, 2, 4\}$, $\{4, 5, 4, 4, 4\}$, $\{5, 5, 4, 2, 4, 5\}$ не являются однородными, а остальные - однородные. Последние именно

по этой причине становятся листьями. Заметим, что вершина $\{1, 1, 3, 3, 5\}$ среди однородных имеет минимальное значение флага, равное 3. Теперь перейдем к четвертому шагу. Так, самая верхняя вершина 3-го уровня $\{1, 2, 3, 2, 3\}$ под воздействием сигнала α переходит в вершину, помеченную такой же A -группой. Поскольку она является непосредственным преемником вершины с такой же A -группой, по правилу 1 она удаляется из ГУА. Аналогичная ситуация имеет место для вершины $\{5, 5, 4, 4, 5\}$ (самая нижняя на 3-м уровне) при построении вершины 4-го уровня, в которую она переходит под воздействием входного символа γ .

Легко убедиться, что при построении остальных вершин 4-го уровня в их флажках будут стоять числа, большие 3, но тогда по правилу 2 они все должны быть удалены из ГУА. Таким образом, граф ГУА содержит только три уровня.

Из самого способа построения ГУА вытекает, что последовательность входных символов, соответствующая пути по ГУА из вершины S_0 в одну из однородных вершин, будет являться УП для заданного автомата.

Покажем теперь, что задача синтеза минимальной по весу УП всегда может быть сведена к задаче выбора наискорейшего пути, как это было и в случае СП.

Роль опорных прямых здесь выполняют линии уровня, а сам ГУА можно интерпретировать как сеть дорог, удовлетворяющую необходимым требованиям для упомянутой задачи, поскольку в нем, во-первых, отсутствуют петли, а, во-вторых, все дуги направлены от вершин i -го уровня к вершинам $(i + 1)$ -го уровня, т. е. отсутствуют "обратные" дуги.

Понятно, что если положить время движения по дуге равным весу входного символа, которым она помечена, то задача синтеза УП минимального веса сводится к задаче о выборе наискорейшего пути из вершины S_0 в однородные вершины ГУА. Отсюда вытекает, что для решения рассматриваемой нами задачи можно применить метод динамического программирования.

Вместе с тем заметим, что ГУА представляет собой дерево, а в дереве, как известно, если между некоторой парой его вершин путь существует, то он единственен. Следовательно, применительно к ГУА задача поиска пути из вершины S_0 в некоторую однородную вершину превращается в тривиальную задачу. Более того, для каждой однородной вершины в ее флажке записано число, равное времени, затрачиваемому на путь в нее из S_0 .

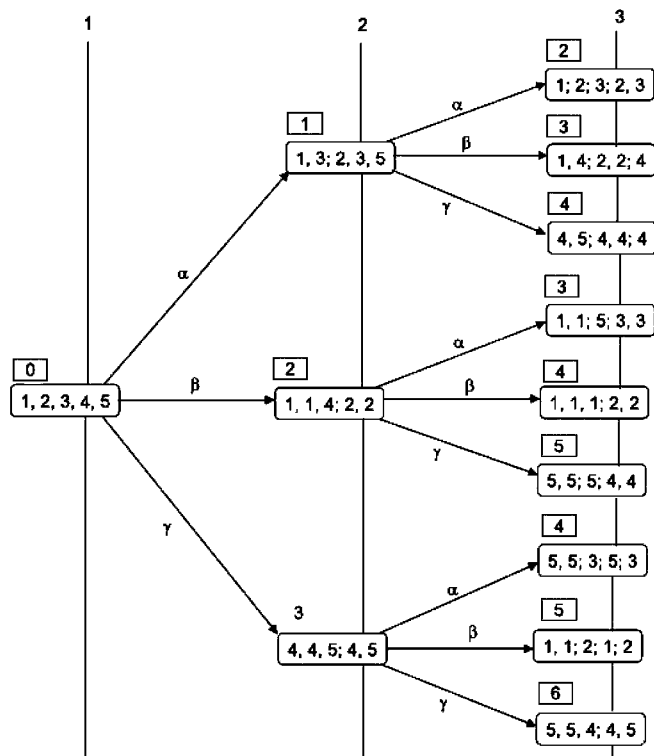


Рис. 2.2.

Таким образом, располагая этими данными для нахождения оптимального пути, достаточно найти ту из однородных вершин, у которой число во флажке минимально. Очевидно, что искомой УП минимального веса будет входная последовательность, соответствующая пути по ГУА из вершины S_0 в однородную вершину с минимальным числом во флажке.

Так, в ГУА на рис.2.2 имеется 5 однородных вершин, среди которых минимальное число во флажке, равное 3, имеет вершина $\{1, 1, 5, 3, 3\}$. Поскольку пути по ГУА из вершины S_0 в нее соответствует входная последовательность β, α , то она и является минимальной по весу УП.

Синтез диагностической последовательности

Решение задачи синтеза ДП, минимальной по весу, по аналогии с изложенным выше, будет состоять из двух этапов. На первом из них по таблице переходов-выходов заданного автомата строится граф, называемый графом диагностики автомата (ГДА), а на втором - производится поиск на построенном графе минимальной ДП.

ГДА представляет собой ориентированный граф, множество вершин которого и способ проведения дуг между ними полностью совпадает с тем, как это было определено для ГУА. По аналогии с предыдущим разделом у каждой вершины ГДА устанавливается флажок, значение которого определяется точно так же, как и в случае ГУА.

Вершины ГДА, которые отмечены A -группами, содержащими кратные (только простые) σ -множества, будем называть кратными (простыми).

Построение ГДА аналогично построению диагностического дерева автомата, описанному в [18], но при этом вводятся несколько иные, чем в [18], правила, по которым вершина S (по терминологии [18] - ветвь) k -го уровня становится листом.

Вершина S k -го уровня ГДА становится листом, если:

- S является простой вершиной;
- S является кратной вершиной;
- в уровнях ГДА, предшествующих k -му, имеется вершина S^* , помеченная такой же A -группой, что и вершина S , но у вершины S значение флажка больше, чем у S^* ;
- в уровнях, предшествующих k -му, имеется простая вершина, значение флажка которой меньше, чем значение флажка вершины S .

Легко показать, что введенные правила обрыва ветвей гарантируют построение ГДА за конечное число шагов.

Проиллюстрируем предложенный способ построения ГДА на примере автомата, таблица переходов-выходов которого отличается от табл. 1.1 только содержимым левой верхней клетки: под воздействием сигнала α автомат из состояния 1 переходит опять в состояние 1, но при этом его выходной сигнал равен 1, а не 0, как в табл. 1.1. Предполагается, что множество допустимых начальных состояний есть $S_0 = \{1, 2, 3, 4, 5\}$, а $w(\alpha) = 1$, $w(\beta) = 20$, $w(\gamma) = 30$.

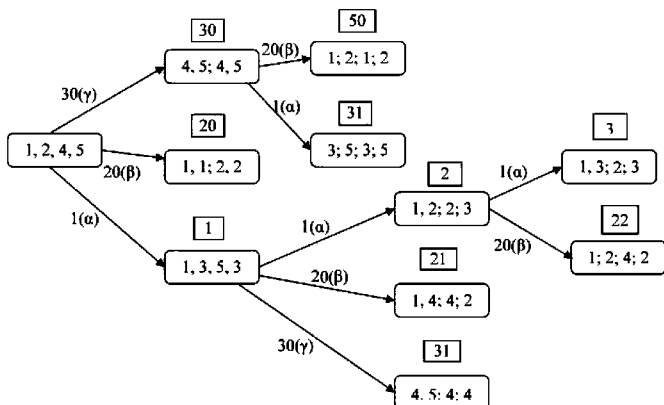


Рис. 2.3.

На рис.2.3 изображен ГДА для этого автомата. Проиллюстрируем теперь, как работают введенные правила завершения. Так, на третьем шаге построения ГДА появляется вершина $\{4, 5, 4, 5\}$ со значением флажка, равным 60, в которую ведет дуга, помеченная входным символом γ , из вершины $\{4, 5, 4, 5\}$ второго уровня со значением флажка, равным 30. Тогда, в силу пункта 3 правил завершения, вершина $\{4, 5, 4, 5\}$ 3-го уровня удаляется. Далее: среди вершин 4-го уровня, появившихся в процессе построения ГДА, имеется простая вершина $\{1, 2, 4, 2\}$ со значением флажка, равным 22. Легко убедиться, что на 5-м шаге построения ГДА у всех появляющихся вершин значения флажков будут больше 22, но тогда в силу пункта 4 правил завершения все эти вершины должны быть удалены из ГДА. Таким образом, ГДА для нашего примера будет содержать только четыре уровня. Отметим также, что вершина $\{1, 1, 2, 2\}$ второго уровня является окончательной в силу пункта 2 правил завершения, вершины $\{1, 1, 2, 2\}$, $\{3, 5, 3, 5\}$ второго уровня и вершина $\{1, 2, 4, 2\}$ третьего уровня становятся окончательными в силу пункта 1 правил завершения.

Из способа построения ГДА вытекает, что последовательность входных символов, соответствующая пути по ГДА из вершины S_0 в одну из простых вершин, будет являться ДП для заданного автомата.

Отметим, что поиск упомянутых путей представляет собой тривиальную задачу, поскольку ГДА является деревом, и потому каждый такой путь единственен. Понятно, что построение ДП с минимальным весом сводится к следующему. Среди всех простых вершин ГДА отыскивается вершина S с минимальным значением флажка. Затем в этом графе находится единственный путь из S_0 в S , которому соответствует определенная последовательность входных символов. Построенная таким образом входная последовательность и является искомой ДП, минимальной по весу.

Так, в ГДА на рис.2.3 имеется 3 простых вершины, из которых у вершины $\{1, 2, 4, 2\}$ четвертого уровня значение флажка минимально и равно 22. Поскольку пути по ГДА из вершины S_0 в эту вершину соответствуют входная последовательность α, α, β , то она и является

минимальной по весу ДП. Заметим, кстати, что минимальной ДП в классическом смысле для этого примера будет являться входная последовательность γ, β , соответствующая пути по ГДА из вершины S_0 в вершину $\{1, 2, 1, 2\}$ третьего уровня.

Вопросы и упражнения

1. Сформулируйте метод построения графа синхронизации (установки, диагностики) автомата.
2. Сформулируйте процедуру построения минимальной по весу СП с использованием ГСА.
3. Какие линии в ГСА выступают в роли опорных прямых при построении искомой СП методом динамического программирования?
4. Для автомата из упражнения 9 к лекции 1 построить ГСА и по этому графу методом динамического программирования найти минимальную по весу СП в предположении, что $w(\alpha) = 10$, $w(\beta) = 1$ и $S_0 = \{1, 2, 3, 4\}$.
5. Для того же автомата построить ГУА и по этому графу методом динамического программирования найти минимальную по весу УП в предположении, что $w(\alpha) = 10$, $w(\beta) = 20$ и $S_0 = \{1, 2, 3\}$.
6. Для автомата из предыдущего упражнения при тех же предположениях о весах входных символов и множестве S_0 начальных состояний построить ГДА и по этому графу найти минимальную по весу ДП.

Обобщенные автоматы без потери информации

Определен новый класс автоматов без потери информации, названных обобщенными автоматами БПИ. Приводится критерий принадлежности автомата этому классу в терминах состояний с потерей информации. Описывается алгоритм распознавания проекции неизвестного входного слова по заданному множеству каналов. Предложена конструкция так называемого проверочного графа автомата, и в его терминах формулируется еще один критерий принадлежности автомата классу ОБПИ.

Автоматы, рассматриваемые как преобразователи информации, используются в системах передачи сообщений и выполняют различные функции, в том числе кодирования этих сообщений. В последней ситуации принятое получателем закодированное сообщение должно быть однозначно декодировано. В этой главе рассматриваются автоматы, удовлетворяющие этому требованию.

Основополагающие работы в области исследования автоматов, являющихся моделями таких устройств, принадлежат Д. Хаффмену [76] и С. Ивену [74]. В этих работах введены и изучены два класса автоматов, названных соответственно автоматами без потери информации (БПИ-автоматами) и автоматами без потери информации конечного порядка (БПИК-автоматами), которые позволяют однозначно декодировать принятое сообщение. В упомянутых работах исследуемые автоматы предполагались инициальными, т. е. стартовыми из известного начального состояния.

В этой лекции предложены обобщения БПИ- и БПИК-автоматов в двух направлениях. Одно из них связано с отказом от предположения об их инициальности, а второе - с исследованием такого рода автоматов со структурированными входными и выходными алфавитами. В последней ситуации восстановление неизвестных входных последовательностей осуществляется не в полном объеме, а на заданном подмножестве компонент структурированного входного символа.

В этом разделе в качестве математической модели ДУ используется слабоинициальный автомат Мили $A = (S, X, Y, \delta, \lambda, s_0)$, где s_0 - множество его допустимых начальных состояний, причем $|S_0| \geq 1$.

Далее предполагается, что входной и выходной алфавиты автомата являются структурированными, т. е. $X = X_1 \times X_2 \times \dots \times X_n$ и $Y = Y_1 \times Y_2 \times \dots \times Y_m$.

Пусть $\bar{p} = \bar{x}_1 \dots \bar{x}_k$ - последовательность входных символов (входное слово) автомата A . Число $d(\bar{p}) = k$ назовем длиной слова \bar{p} .

Пусть $\bar{x} = (x_1, \dots, x_n) \in X, 1 \leq i_1 \leq i_2 \leq \dots \leq i_\nu$, - натуральные числа, $1 \leq i_j \leq n$, тогда проекцией \bar{x} по каналам с номерами i_1, i_2, \dots, i_ν назовем вектор $(x_{i_1}, \dots, x_{i_\nu})$ и будем обозначать ее $pr_{i_1, \dots, i_\nu} \bar{x}$, а проекцией слова $bar{p} = \bar{x}_1 \dots \bar{x}_k$ по тем же каналам назовем упорядоченную последовательность $(pr_{i_1, \dots, i_\nu} \bar{x}_1, \dots, pr_{i_1, \dots, i_\nu} \bar{x}_k)$ и будем обозначать ее $pr_{i_1, \dots, i_\nu} \bar{p}$.

Рассмотрим следующую задачу. На автомат A , находящийся в одном из состояний множества S_0 , но неизвестно, в каком именно, подается неизвестное входное слово \bar{p} , и наблюдается реакция автомата на это слово по выходным каналам с номерами $j_1 < j_2 < \dots < j_\mu$, где $1 \leq j_i \leq m$. Требуется построить эксперимент, проводимый с автоматом A после приложения слова \bar{p} , позволяющий распознать проекцию этого слова по каналам с номерами i_1, i_2, \dots, i_ν , где $i_1 < i_2 < \dots < i_\nu$.

Автоматы, для которых сформулированная задача может быть решена независимо от входного слова \bar{p} и действительного начального состояния из множества S_0 , назовем обобщенными автоматами без потери информации (ОБПИ-автоматами).

Далее, не теряя общности, положим, что $i_1 = 1, i_2 = 2, \dots, i_\nu = \nu$ и соответственно $j_1 = 1, j_2 = 2, \dots, j_\mu = \mu$, т. е. реакция автомата A на слово \bar{p} наблюдается по первым μ выходным каналам, а в каждом символе $\bar{x}_i (i = \overline{1, k})$ слова \bar{p} восстановлению подлежат сигналы, поступающие на автомат по первым ν входным каналам.

Из формулировки задачи следует, что ОБПИ-автоматы должны обладать следующим свойством:

$$\forall s, t \in S_0 \forall \bar{p}, \bar{q} \in X^* pr_{1, \dots, \mu} \lambda(s, \bar{p}) = pr_{1, \dots, \mu} \lambda(t, \bar{p}) \rightarrow pr_{1, \dots, \nu} \bar{p} = pr_{1, \dots, \nu} \bar{q}$$

Отметим, что определенные нами ОБПИ-автоматы в качестве частного случая включают в себя ранее известные БПИ-автоматы, введенные Д. Хаффменом [76], и автоматы существенно без потери информации (СБПИ-автоматы) [9] [43], введенные автором предлагаемой монографии. БПИ-автоматы получаются при $|S_0| = 1, \nu = n, \mu = m$, а СБПИ-автоматы получаются при $S_0 = S, \nu = n, \mu = m$.

Определение 3.1. Пару состояний s и t автомата A назовем состояниями с потерей информации (СПИ-состояниями), если

$$\exists \bar{p}, \bar{q} \in X^* \delta(s, \bar{p}) = \delta(t, \bar{q}) pr_{1, \dots, \mu} \lambda(s, \bar{p}) = pr_{1, \dots, \mu} \lambda(t, \bar{q}) \rightarrow pr_{1, \dots, \nu} \bar{p} \neq pr_{1, \dots, \nu} \bar{q}$$

Заметим, что если $s = t$, то приведенное определение совпадает с определением из [76], и в этом случае s называется СПИ-состоянием.

Теорема 3.1. Для того чтобы автомат A был ОБПИ-автоматом, необходимо и достаточно, чтобы он не имел СПИ-состояний.

Доказательство. Необходимость условий теоремы очевидна, покажем их достаточность.

Пусть условия теоремы выполняются. Предположим, что на вход автомата A было подано неизвестное слово \bar{v} , а на первых μ выходных каналах наблюдалось слово \bar{w} . По таблице переходов-выходов автомата A определим пары "состояние - входное слово" $(s_{i_1}; \bar{v}_1), \dots, (s_{i_r}; \bar{v}_r)$, такие, что $s_{i_j} \in S_0$ и $pr_{1, \dots, \mu} \lambda(s_{i_j}, \bar{v}_j) = \bar{w}$, где $j = 1, \dots, r$. Понятно, что среди слов $\bar{v}_1, \dots, \bar{v}_r$ находится и искомое слово.

Пусть $\delta(s_{i_j}, \bar{v}_j) = s_{i_j}^*; j = 1, \dots, r$. Покажем, что между парами (s_{i_j}, \bar{v}_j) и состояниями $s_{i_j}^*$ существует взаимно однозначное соответствие. Последнее означает, что такое соответствие существует и

между $pr_{1,\dots,\nu}\bar{v}_j$ и состояниями $s_{i_j}^*$, $j = 1, \dots, r$.

Если все состояния $s_{i_1}^*, \dots, s_{i_r}^*$ попарно различны, справедливость сформулированного утверждения очевидна. Предположим, что среди состояний $s_{i_1}^*, \dots, s_{i_r}^*$ имеется по крайней мере два совпадающих. Без ограничения общности можно считать, что $s_{i_1}^* = s_{i_2}^*$. Из последнего равенства следует, что $\delta(s_{i_1}, n\bar{u}_1) = \delta(s_{i_2}, \bar{v}_2)$. По построению пары (s_{i_1}, \bar{v}_1) и (s_{i_2}, \bar{v}_2) различны, что возможно только в следующих трех случаях:

1. $s_{i_1} = s_{i_2}, pr_{1,\dots,\nu}\bar{v}_1 \neq pr_{1,\dots,\nu}\bar{v}_2$
2. $s_{i_1} \neq s_{i_2}, pr_{1,\dots,\nu}\bar{v}_1 \neq pr_{1,\dots,\nu}\bar{v}_2$
3. $s_{i_1} \neq s_{i_2}, pr_{1,\dots,\nu}\bar{v}_1 = pr_{1,\dots,\nu}\bar{v}_2$.

В первом случае состояние S_{i_1} является по определению СПИ-состоянием, что противоречит нашему предположению. Во втором случае пара состояний S_{i_1} и s_{i_2} является по определению СПИ-состояниями, что также противоречит нашему предположению. Таким образом, в действительности может иметь место только третий случай, когда у двух входных слов \bar{v}_1 и \bar{v}_2 их проекции по каналам $1, \dots, \nu$ совпадают. Это и доказывает справедливость утверждения о существовании взаимно однозначного соответствия между $pr_{1,\dots,\nu}\bar{v}_j$ и состояниями $s_{i_j}^*$, $j = 1, \dots, r$.

Продолжим доказательство теоремы. Предположим, что среди состояний $s_{i_1}^*, \dots, s_{i_r}^*$ все попарно различные состояния образуют множество $\hat{S} = \{s_{k_1}^*, \dots, s_{k_f}^*\}$, где $f \leq r$. Рассматривая \hat{S} в качестве множества допустимых начальных состояний автомата A , построим для него установочную последовательность L . Обозначим через s состояние, в которое перейдет автомат A после подачи слова L , а выходную последовательность, которую он выдает при этом, обозначим через M . Если пара (L, M) порождается только из одного состояния множества \hat{S} (этот факт легко установить по таблице

переходов-выходов автомата A), тогда, очевидно, существует взаимно однозначное соответствие между заключительным состоянием s и состоянием \hat{s}_{k_i} , между \hat{s}_{k_i} и $s_{i_j}^*$ и, наконец, между $s_{i_j}^*$ и $pr_{1, \dots, \nu, \bar{\nu}_j}$. В силу сказанного, знание заключительного состояния s позволяет однозначно восстановить проекцию неизвестного входного слова, поданного на автомат A , по каналам с номерами $1, \dots, \nu$.

Если пара (L, M) порождается из двух различных состояний множества \hat{S} , например \hat{s}_{k_1} и \hat{s}_{k_2} , то им взаимно однозначно соответствуют состояния $s_{i_a}^*$ и $s_{i_b}^*$, последним - состояния S_{i_a} и S_{i_b} из множества S_0 , а состояниям S_{i_a} и S_{i_b} , в свою очередь, взаимно однозначно соответствуют входные слова $\bar{\nu}_a$ и $\bar{\nu}_b$. В силу выполненных выше построений имеем следующие равенства:

$$\delta(s_{i_a}, \bar{\nu}_a L) = \delta(s_{i_b}, \bar{\nu}_b L), pr_{1, \dots, \mu} \lambda(s_{i_a}, \bar{\nu}_a L) = pr_{1, \dots, \mu} \lambda(s_{i_b}, \bar{\nu}_b L) = \bar{w} pr_{1, \dots, \mu} M$$

Если входные слова $\bar{\nu}_a$ и $\bar{\nu}_b$ таковы, что $pr_{1, \dots, \mu} \bar{\nu}_a = pr_{1, \dots, \mu} \bar{\nu}_b$, то искомая проекция восстанавливается однозначно. Если же выписанные только что проекции различны, то это означает, что пара состояний S_{i_a} и S_{i_b} является СПИ-состояниями. Этот факт противоречит условию теоремы.

Заметим, что аналогичные рассуждения справедливы и для случая, когда пара (L, M) порождается более чем из двух состояний.

Теорема доказана.

Доказательство теоремы фактически содержит алгоритм распознавания проекции неизвестного входного слова по заданному множеству каналов, который сформулируем в явном виде.

1. Наблюдаем проекцию \bar{w} выходного слова автомата A по каналам $1, \dots, \mu$, являющуюся реакцией на подачу неизвестного входного слова $\bar{\nu}$. По таблице переходов-выходов автомата A определяем все пары "состояние - входное слово" $(s_{i_1}; \bar{\nu}_1), \dots, (s_{i_r}; \bar{\nu}_r)$, такие, что $s_{i_j} \in S_0$ и

$pr_{1,\dots,\mu}\lambda(s_{i_j}, \nu_j) = \bar{w}$, где $j = 1, \dots, r$.

2. Строим установочную последовательность L , считая множеством \hat{S} допустимых начальных состояний автомата множество всех попарно различных состояний из совокупности состояний $\delta(s_{i_1}, \bar{\nu}_1), \dots, \delta(s_{i_r}, \bar{\nu}_r)$.
3. Подаем на вход автомата A слово L и наблюдаем выходное слово M . По паре (L, M) определяем заключительное состояние автомата A и все те состояния $\hat{S}_{k_a}, \dots, \hat{S}_{k_b}$, которые этой парой порождаются.

Из множества пар $(s_{i_1}, \bar{\nu}_1), \dots, (s_{i_r}, \bar{\nu}_r)$ выделяем все те его элементы, для которых $\delta(s_{i_j}, \bar{\nu}_j) = s_{k_f}^*$, $f = 1, \dots, b$. Если этот элемент один, то проекция второго члена выделенной пары по каналам $1, \dots, \nu$ и есть искомое входное слово. Если число таких элементов больше двух, то у них упомянутые проекции вторых членов пар совпадают и являются искомым входным словом.

Рассмотрим пример. Пусть ОБПИ-автомат задан графом на рис.3.1. Условимся, что для наблюдения реакции выделен 1-й выходной канал автомата (по нему выдается левый символ выходной пары) и проекция неизвестного входного слова восстанавливается по 1-му входному каналу, т. е. $\nu = \mu = 1$. Пусть $S_0 = \{1, 2\}$, на автомат подано неизвестное входное слово длиной 3, а по 1-му выходному каналу при этом наблюдалась реакция 0,1,1. Строим последовательность пар "состояние - входное слово", упомянутую в пункте 1 алгоритма:

(1;10,00,10), (1;10,00,11), (1;10,00,00), (1;10,00,11), (1;10,01,10),
 (1;10,01,11), (1;10,01,00), (1;10,01,10), (1;11,00,10), (1;11,00,11),
 (1;11,00,00), (1;11,00,10), (1;11,01,10), (1;11,01,11), (1;11,01,00),
 (1;11,01,10), (2;10,10,10), (2;10,10,11), (2;10,11,10), (2;10,11,11),
 (2;11,10,10), (2;11,10,11), (2;11,11,10), (2;11,11,11).

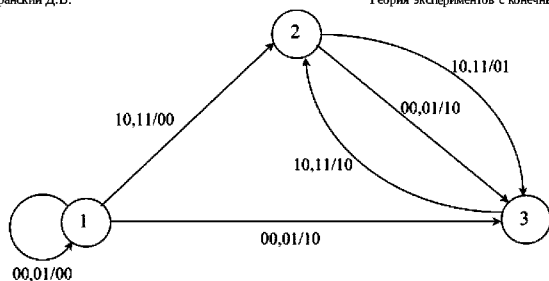


Рис. 3.1.

Нетрудно убедиться, что упомянутое в пункте 2 алгоритма множество есть $\hat{S} = \{1, 2, 3\}$. Установочной последовательностью для рассматриваемого автомата с множеством допустимых начальных состояний \hat{S} является, например, слово длиной два $L = 01, 01$. Пользуясь [таблицей 3.1](#), по действительной реакции автомата на слово L находим состояния $s_{\hat{k}_i}$, упомянутые в пункте 3 алгоритма.

Таблица 3.1.

Состояние автомата	ВХОДНАЯ ПОСЛЕДОВАТЕЛЬНОСТЬ	Реакция автомата	Заключительное состояние
1	01,01	00,00	1
2	01,01	10,10	1
3	01,01	10,00	1

Если $s_{\hat{k}_i} = 1$, то ему соответствуют пары $(1; 10, 00, 00)$, $(1; 10, 00, 01)$, $(1; 10, 01, 00)$, $(1; 10, 01, 01)$, ..., $(1; 11, 01, 00)$ и проекцией по 1-му входному каналу неизвестного входного слова является $1, 0, 0$. При $s_{\hat{k}_i} = 2$ искомая проекция есть $1, 0, 1$ и, например, при $s_{\hat{k}_i} = 3$ проекция равна $1, 1, 1$.

Хотя теорема 3.1 позволила обосновать метод восстановления искомой

проекции неизвестного входного слова, однако ответить с ее помощью на вопрос, является ли конкретный автомат ОБПИ-автоматом, сложно из-за отсутствия эффективного способа проверки условий теоремы. Ввиду этого ниже сформулируем еще одно необходимое и достаточное условие принадлежности автомата классу ОБПИ, но в других по сравнению с теоремой 3.1 терминах, допускающее простую проверку.

С этой целью по аналогии с [74] введем понятие проверочного графа автомата. Условимся считать, что автомат задан с помощью конечного ориентированного графа. Будем говорить, что дуга $s_i \rightarrow s_j$ графа автомата A , где $s_i, s_j \in S$, имеет пометку $\bar{y} \in Y$, если $\exists x \in X \delta(s_i, \bar{x}) = s_j \lambda(s_j, \bar{x}) = \bar{y}$.

Аналогичное определение сформулируем и для случая, когда выходной символ \bar{y} наблюдается не по всем, а только по μ выделенным каналам с номерами $j_1 < j_2 < \dots < j_m$: дуга $s_i \rightarrow s_j$ имеет пометку $pr_{j_1 \cdot j_2 \cdot \dots \cdot j_m} \bar{y}$, если

$$\exists x \in X \delta(s_i, \bar{x}) = s_j pr_{j_1 \cdot j_2 \cdot \dots \cdot j_m} \lambda(s_j, \bar{x}) = pr_{j_1 \cdot j_2 \cdot \dots \cdot j_m} \bar{y}$$

Понятно, что две вершины графа могут связываться несколькими дугами с разными пометками.

Пусть $\Gamma\{S_0\}$ - множество всех состояний автомата A , достижимых из множества допустимых начальных состояний S_0 путями любой длины.

По графу автомата A построим другой ориентированный конечный граф $G(A)$ следующим образом.

1. Вершинами $G(A)$ являются все возможные неупорядоченные пары $\{s_i, s_l\}$, где $s_i, s_j \in \Gamma\{S_0\}$.
2. Из вершины $\{s_k, s_l\}$ в вершину $\{s_i, s_j\}$ проводится дуга с пометкой $pr_{j_1 \cdot j_2 \cdot \dots \cdot j_m} \bar{y}$, если

$$\exists x_1, x_2 \in X \delta(s_k, \bar{x}_1) = s_i \delta(s_l, \bar{x}_2) = s_j p r_{j_1, j_2, \dots, j_\mu} \lambda(s_k, \bar{x}_1) = p r_{j_1, j_2, \dots, j_\mu} \lambda(s_l, \bar{x}_2) = p t_{j_1, j_2, \dots, j_\mu} \bar{y}$$

В графе $G(A)$ две вершины могут быть связаны несколькими дугами с разными пометками.

Пусть неизвестное входное слово необходимо восстановить по каналам с номерами i_1, \dots, i_ν , где $i_1 < i_2 < \dots < i_\nu$.

Дугу $\{s_k, s_l\} \rightarrow \{s_i, s_j\}$ с пометкой $p r_{j_1, j_2, \dots, j_\mu} u$ назовем выделенной, если

$$\exists x_1, x_2 \in X \delta(s_k, \bar{x}_1) = s_i \delta(s_l, \bar{x}_2) = s_j p r_{j_1, \dots, j_\mu} \lambda(s_k, \bar{x}_1) = p r_{j_1, j_2, \dots, j_\mu} \lambda(s_l, \bar{x}_2) = p r_{j_1, j_2, \dots, j_\mu} y \rightarrow p r_{i_1, \dots, i_\nu} \bar{x}_1 \neq p r_{i_1, \dots, i_\nu} \bar{x}_2$$

Из графа $G(A)$ удалим все вершины вида $\{s, s\}$ вместе с инцидентными им дугами, если последние, в свою очередь, инцидентны только вершинам такого же вида, а также изолированные вершины. Полученный в результате такого удаления ориентированный конечный граф назовем проверочным графом автомата A и обозначим его $T(A)$.

Теорема 3.2. Автомат $T(A)$ является ОБПИ-автоматом тогда и только тогда, когда все пути в проверочном графе $T(A)$, ведущие в вершины вида $\{s, s\}$, не содержат выделенных дуг.

Необходимость. Пусть A есть ОБПИ-автомат, но в $T(A)$ существует путь с выделенной дугой $\{s_k, s_l\} \rightarrow \{s_i, s_j\}$, ведущий в вершину вида $\{s, s\}$. Предположим, что \bar{p} и \bar{q} - два таких входных слова, что $\delta(s_k, \bar{p}) = \delta(s_l, \bar{q}) = s$, и они соответствуют пути по графу в $T(A)$, начинающемуся с упомянутой выделенной дуги. Последнее означает, что

$$\exists s_k, s_l \in T|S_0 \exists p, q \in X^* p r_{j_1, j_2, \dots, j_\mu} \lambda(s_k, \bar{p}) = p r_{j_1, j_2, \dots, j_\mu} \lambda(s_l, \bar{q}) \rightarrow p r_{i_1, \dots, i_\nu} \bar{p} \neq p r_{i_1, \dots, i_\nu} \bar{q}$$

Из сравнения этого высказывания с (3.1) вытекает, что A не может быть ОБПИ-автоматом, а это противоречит исходной посылке.

Достаточность. Докажем ее методом от противного. Пусть A не является ОБПИ-автоматом. Покажем тогда, что в графе $T(A)$ существует путь, ведущий в вершину вида $\{s, s\}$, содержащий выделенную дугу. Поскольку A не есть ОБПИ-автомат, то

$$\exists s_k, s_l \in S_0 \exists p, q \in X^* \delta(s_k, \bar{p}) = \delta(s_l, \bar{q}) = spr_{j_1, j_2, \dots, j_\mu} \lambda(s_k, \bar{p}) = pr_{j_1, j_2, \dots, j_\mu} \lambda(s_l, \bar{q}) \rightarrow pr_{i_1, \dots, i_\nu} \bar{p} \neq pr_{i_1, \dots, i_\nu} \bar{q}$$

Пусть \bar{x}_p и \bar{x}_q - первые слева символы в словах \bar{p} и \bar{q} , проекции которых по каналам i_1, \dots, i_ν различны. Пусть при подаче этих символов автомат A осуществляет переходы из состояний s_1, s_2 в состояния \hat{s}_1, \hat{s}_2 соответственно. Отсюда следует, что в графе $T(A)$ существует выделенная дуга, соединяющая вершины $\{s_1, s_2\}$ и $\{\hat{s}_1, \hat{s}_2\}$. Это означает существование в графе $T(A)$ пути из $\{s_1, s_2\}$ в $\{s, s\}$, первая дуга которого выделенная. Этим завершается доказательство достаточности условий теоремы.

В качестве примера на рис.3.2 изображен проверочный граф для автомата, представленного на рис.3.1, в предположении, что $\nu = \mu = 1$. На рисунке двойные стрелки соответствуют выделенным дугам, над каждой дугой стоит соответствующая ей пометка. В этом проверочном графе все возможные пути, заканчивающиеся в вершинах вида $\{s, s\}$, могут начинаться только в вершинах $\{1, 1\}$ либо $\{3/3\}$, при этом конечной вершиной в любом из этих путей является вершина $\{1, 1\}$. Из рис.3.2 видно, что ни один из этих путей не содержит выделенных дуг, следовательно, рассматриваемый автомат относится к классу ОБПИ-автоматов.

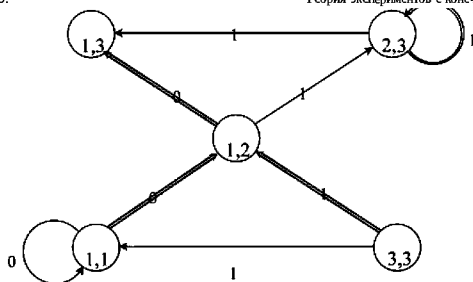


Рис. 3.2.

Если задан автомат A с n входными и m выходными каналами, то возникает вопрос: для каких подмножеств входных и выходных каналов он является ОБПИ-автоматом? Очевидно, что A может не быть БПИ-автоматом в классическом смысле, но в то же время быть ОБПИ-автоматом, если восстановление входных сигналов производится по некоторому подмножеству входных каналов.

При отыскании всех возможных подмножеств входных и выходных каналов, для которых заданный автомат A является ОБПИ-автоматом, в первую очередь представляет интерес случай, когда подмножество входных каналов - максимальное, а подмножество выходных каналов - минимальное по мощности. Далее такой автомат будем называть оптимальным. Это соответствует ситуации, когда при минимальных затратах, определяемых объемом наблюдаемой информации на выходах, удается однозначно восстановить максимальный объем информации, поступившей на вход автомата.

Опишем способ, позволяющий получить ответ на сформулированный выше вопрос. Условимся далее через A_M^R обозначать автомат, у которого выходные символы наблюдаются по каналам с номерами из множества M , а восстановление сигналов ведется по каналам с номерами из множества R .

По графу автомата A построим его проверочный граф $T(A)$ при

$\nu = n$ и $\mu = m$, т. е. предполагая, что реакция A наблюдается по всем его выходным каналам, а восстановление неизвестного входного слова также ведется по всем каналам. Используя $T(A)$, найдем в нем все пути, ведущие в вершины вида $\{s, s\}$, и образуем на этой основе множество $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_t\}$ всех входящих в эти пути выделенных дуг. Если $\Gamma = \emptyset$, то по теореме 3.2 автомат A есть ОБПИ-автомат, и при $\mu = m$ он будет, очевидно, таковым для любого ν , где $1 \leq \nu \leq n$. Если $\Gamma \neq \emptyset$, то для каждого $i = \overline{1, t}$ построим множество R_i^* целых чисел, представляющих собой номера входных каналов, на которых на дуге γ_i происходит потеря информации. Последнее означает, что если реакция автомата A соответствует отметке дуги γ_i , то в каналах с номерами из множества R_i^* сигналы могут иметь значения как 0, так и 1. Используя множество $R^* = R_1 \cup \dots \cup R_t^*$, построим $R = \{1, \dots, n\} \setminus R^*$. Из способа построения следует, что, во-первых, R содержит все те номера входных каналов, по каждому из которых и по их совокупности рассматриваемый автомат A является ОБПИ-автоматом, и, во-вторых, R - максимальный по мощности, сохраняющий свойство ОБПИ.

Далее через $T_Q^R(A)$ обозначим проверочный граф автомата A , предполагая, что наблюдение реакций ведется по каналам с номерами из множества Q , а восстановление неизвестного входного слова - по каналам с номерами из множества R .

Пусть $M = \{1, \dots, m\}$. Для каждого $j, 1 \leq j \leq m$, построим проверочный граф $T_{Mj}^R(A)$. Нетрудно сообразить, что дуги графа $T_M^R(A)$ одновременно являются и дугами графа $T_{Mj}^R(A)$, но в последнем, кроме того, могут появиться и новые. Если последний упомянутый граф удовлетворяет условиям теоремы 3.2, то, очевидно, j -й выходной канал избыточен в том смысле, что отказ от наблюдения реакции по нему не повлияет на возможности восстановления неизвестного входного слова по каналам из множества R . Пусть $M^* = \{j_1, j_2, \dots, j_s\}$ - множество номеров всех избыточных

выходных каналов автомата A . Если $M^* = \emptyset$, то построенный на 1-м этапе автомат A_M^R - искомый. При $M^* \neq \emptyset$ рассмотрим все возможные сочетания пар каналов из множества M^* , число которых равно $C_{|M^*}^2$. Для каждой из упомянутых пар каналов $J + \{j_k, j_l\}$ построим проверочный граф $T_M^R(A)$ и проверим для него выполнение условий теоремы 3.2. Если ни один из таких графов не удовлетворяет условиям этой теоремы, то автоматы $A_{M \setminus J_i}^R, i = \overline{1, s}$, - искомые оптимальные автоматы. В противном случае формируем все возможные сочетания J троек номеров из множества M^* и ищем среди проверочных графов $T_M^R(A)$ такие, которые удовлетворяют условиям теоремы 3.2. Указанный процесс продолжается до того момента, когда на очередном этапе все графы вида $T_M^R(A)$, где $1 \leq |J| \leq s$, не удовлетворяют условиям теоремы 3.2. Тогда автоматы A_M^R , построенные на предыдущем этапе и являющиеся ОБПИ-автоматами, - искомые оптимальные.

Отметим, что для поиска всех оптимальных ОБПИ-автоматов в худшем случае может потребоваться построение $\sum_{i=1}^s C_s^i$ различных проверочных графов, где s - число избыточных выходных каналов автомата A .

Вопросы и упражнения

1. Дайте определение пары состояний с потерей информации.
2. Сформулируйте определение обобщенного автомата без потери информации.
3. Приведите критерий принадлежности автомата классу автоматов ОБПИ в терминах СПИ-состояний.
4. Опишите процедуру распознавания проекции неизвестного входного слова, поданного на ОБПИ-автомат.
5. Опишите процедуру построения проверочного графа автомата.
6. Приведите критерий принадлежности автомата классу ОБПИ-автоматов в терминах проверочного графа.
7. Дайте определение оптимального ОБПИ-автомата.

8. ОБПИ-автомат задан графом на рис.3.1. Пусть для наблюдения реакции выделен 1-й выходной канал автомата (он соответствует левому символу входной пары), а проекция неизвестного входного слова должна быть восстановлена по первому входному каналу. Пусть множество допустимых начальных состояний автомата есть $\{1, 2\}$. На этот автомат подано неизвестное входное слово длины 2, а по первому выходному каналу наблюдалась реакция 0,1. Требуется восстановить проекцию по первому каналу неизвестного входного слова.

Обобщенные автоматы без потери информации конечного порядка

Определен новый класс автоматов без потери информации конечного порядка, названных обобщенными автоматами БПИК. Приводится критерий принадлежности автомата классу БПИК в терминах проверочного графа. Описан способ определения порядка БПИК-автомата. Описан метод построения комбинационной схемы, восстанавливающей проекцию неизвестного входного слова.

Автоматы без потери информации конечного порядка (ОБПИК-автоматы) являются в некотором смысле аналогами БПИК-автоматов, введенных С. Ивеном [74].

В этом разделе нами используется в качестве математической модели дискретного устройства слабоинициальный автомат Мили со структурированными входными и выходными алфавитами. Условимся также применять далее те же обозначения, понятия и конструкции, что были введены в предшествующем разделе.

Для однозначности определения числа $N(A)$ будем считать, что оно является минимальным из всех возможных для заданного автомата A чисел такого рода, и назовем его порядком автомата A .

Отметим, что определенные нами ОБПИК-автоматы в качестве частного случая включают в себя ранее известные БПИК-автоматы, введенные С. Ивеном [74], и автоматы существенно без потери информации конечного порядка (СБПИК-автоматы), введенные автором представленного курса лекций в [9], [46]. БПИК-автоматы получаются при $|S_0| = 0$, $\nu = n, \mu = m$, а СБПИК-автоматы получаются при $S_0 = S, \nu = n, \mu = m$, где n - число входных каналов автомата.

Пусть $\bar{p} = p_1 p_2 \dots p_\alpha$ и $\bar{q} = q_1 q_2 \dots q_\beta$ - слова в структурированном алфавите X . Из определения ОБПИК-автомата порядка $N(A)$ следует, что он должен обладать следующим свойством:

$$\forall s, t \in S, \forall \bar{p}, \bar{q} \in X^* d(\bar{p}) = d(\bar{q}) = N(A) p r_{1, \dots, \mu} \lambda(s, \bar{p}) = p r_{1, \dots, \mu} \lambda(s, \bar{q}) \rightarrow p r_{1, \dots, \nu} \bar{p} = p r_{1, \dots, \nu} \bar{q}$$

Здесь, не теряя общности, предполагается, что наблюдение реакции ведется по первым μ выходным каналам с номерами $1, 2, \dots, \mu$, а неизвестное слово будет восстанавливаться по первым ν входным каналам с номерами $1, 2, \dots, \nu$.

Теорема 4.1. Автомат A является ОБПИК-автоматом тогда и только тогда, когда в его проверочном графе $T(A)$ все контуры и все пути, ведущие в эти контуры, не содержат выделенных дуг.

Необходимость. Пусть A есть ОБПИК-автомат порядка $N(A)$. Покажем, что он одновременно является и ОБПИ-автоматом. Пусть на вход A подано неизвестное входное слово $\bar{\nu}$, такое, что $p r_{1, \dots, \nu} \bar{\nu} = \nu_1 \nu_2 \dots \nu_{d(\nu)}$, а реакция \bar{w} автомата A на это слово такова, что $p r_{1, \dots, \mu} \bar{w} = w_1 w_2 \dots w_{d(\nu)}$. Подадим на вход A произвольное слово $\bar{\nu}^*$ длины $N(A)$ (пусть $p r_{1, \dots, \nu} \bar{\nu}^* = \nu_1 * \dots \nu_{N(A)}^*$), реакция на которое есть слово \bar{w}^* (пусть $p r_{1, \dots, \mu} \bar{w}^* = w_1 * \dots w_{N(A)}^*$). Поскольку A есть ОБПИК-автомат порядка $N(A)$, то по начальному отрезку проекции выходного слова $w_1 w_2 \dots w_{d(\nu)} w_1 * \dots w_{N(A)}^*$ длины $N(A)$ можно определить 1-й символ ν_1 неизвестного входного слова, аналогично по начальному отрезку длины $N(A)$ слова $w_2 w_3 \dots w_{d(\nu)} w_1 * \dots w_{N(A)}^*$ - символ ν_2 и, наконец, по начальному отрезку длины $N(A)$ слова $w_{d(\nu)} w_1 * \dots w_{N(A)-1}^*$ - символ $\nu_{d(\nu)}$. Таким образом, подача на вход автомата A произвольного слова длины $N(A)$ является тем экспериментом, который всегда позволяет определить искомую проекцию неизвестного входного слова. Следовательно, A является ОБПИ-автоматом.

Предположим, что в $T(A)$ имеется контур $\{s_1, s_j\} \rightarrow \{s_{i_1}, s_{j_1}\} \rightarrow \dots \rightarrow \{s_{i_{t-1}}, s_{j_{t-1}}\} \rightarrow \{s_i, s_j\}$ длины t

, первая дуга которого - выделенная. В силу теоремы 3.2 предшествующей лекции все вершины $\{s_a, s_b\}$ этого контура таковы, что $s_a \neq s_b$. Обозначим через \bar{p} и \bar{q} входные слова, соответствующие проходу автомата A по состояниям контуров $s_i, s_{i_1}, \dots, s_{i_{t-1}}$ и $s_j, s_{j_1}, \dots, s_{j_{t-1}}$. Поскольку 1-я дуга контура в $T(A)$ является выделенной, то $\bar{p} = p_1 \bar{v}$ и $\bar{q} = q_1 \bar{v}$, где $p_1 \neq q_1$. Если к автомату A , находящемуся в состоянии s_j или s_i , приложить входные слова вида $p_1 \bar{v}^r \bar{v}^*$ и $q_1 \bar{v}^r \bar{v}^*$, где r - целое положительное число, \bar{v}^* - некоторый начальный отрезок слова \bar{v} , \bar{v}^r - r -кратное повторение слова \bar{v} , то понятно, что наблюдаемые проекции автомата в обоих случаях будут одинаковыми. Следовательно, по наблюдаемой реакции восстановить однозначно 1-й символ неизвестного входного слова невозможно. Поскольку сказанное справедливо для любого r и начального отрезка \bar{v}^* любой длины, то упомянутая потеря информации будет иметь место для слов произвольной длины. Последний факт противоречит исходному предположению о том, что A является ОБПИК-автоматом.

Нетрудно убедиться, что аналогичная ситуация имеет место и в случае, когда в графе $T(A)$ существует путь, ведущий в контур и начинающийся с выделенной дуги.

На этом завершается доказательство необходимости условий теоремы.

Достаточность. Пусть $T(A)$ не содержит контуров и, следовательно, в нем имеются только пути конечной длины. Рассмотрим вначале случай, когда эти пути не содержат выделенных дуг. Покажем, что в этих предположениях однозначное восстановление проекции первого символа неизвестного входного слова длины $N(A)$ возможно.

Предположим противное: пусть в $T(A)$ существует два пути

$$\{s_{i_1}, s_{j_1}\} \xrightarrow{x_1/y_1} \{s_{i_2}, s_{j_2}\} \rightarrow \dots \xrightarrow{x_{N(A)}/y_{N(A)}} \{s_{i_{N(A)}}, s_{j_{N(A)}}\}$$

и

$$\{s_{k_1}, s_{l_1}\} \xrightarrow{\tilde{x}_1/y_1} \{s_{k_2}, s_{l_2}\} \rightarrow \dots \xrightarrow{x_{N(A)}/y_{N(A)}} \{s_{i_{N(A)+1}}, s_{j_{N(Q)+1}}\}$$

где $x_1 \neq \tilde{x}_1$. Очевидно, что если проекция реакции автомата A есть $y_1 y_2 \dots y_{N(A)}$, то проекцию 1-го искомого символа восстановить однозначно невозможно. Однако в этом случае в $T(A)$ должны существовать выделенные дуги

$$\{s_{i_1}, s_{k_1}\} \xrightarrow{y_1} \{s_{i_2}, s_{k_2}\}, \{s_{i_1}, s_{l_1}\} \xrightarrow{y_1} \{s_{i_2}, s_{l_2}\} \text{ и т. д.,}$$

что противоречит исходному предположению.

Теперь рассмотрим случай, когда в графе $T(A)$ содержится путь $\{s_{i_1}, s_{j_1}\} \rightarrow \{s_{i_2}, s_{j_2}\} \rightarrow \dots \rightarrow \{s_{i_a}, s_{j_a}\}$, у которого 1-я дуга является выделенной. Отсутствие в вершине $\{s_{i_a}, s_{j_a}\}$ исходящей дуги означает, что при любом входном сигнале реакции автомата A из состояний s_{i_a} и s_{j_a} различны. Но тогда по этой реакции идентифицируется истинное состояние автомата (s_{i_a} или s_{j_a}). Поскольку по теореме 3.2 предшествующей лекции для вершины $\{s_{i_k}, s_{j_k}\}$ упомянутого выше пути $s_{i_k} \neq s_{j_k}$ при $k = 2, \dots, a$, то обратным ходом можно идентифицировать истинную последовательность сменяющихся состояний автомата A из двух возможных: s_{i_1}, \dots, s_{i_a} либо s_{j_1}, \dots, s_{j_a} . Отсюда вытекает, что проекция первого символа неизвестного входного слова восстанавливается однозначно.

Перейдем к рассмотрению случая, когда в графе $T(A)$ имеются контуры. Доказательство для него проведем от противного. Пусть A не является ОБПИК-автоматом. В частности, в нем могут существовать два контура:

$$s_1 \xrightarrow{x_1/y_1} s_2 \xrightarrow{x_2/y_2} \dots \rightarrow s_k \xrightarrow{x_b/y_b} s_1$$

и

$$s_{1*} \xrightarrow{\tilde{x}_1/y_1} s_{2*} \xrightarrow{x_2/y_2} \dots \rightarrow s_{b*} \xrightarrow{x_b/y_b} s_{1*}$$

где происходит потеря информации о первом входном символе, если

$x_1 \neq \hat{x}_1$. Но тогда в графе $T(A)$ существует контур $\{s_1, s_1^*\} \rightarrow \{s_2, s_2^*\} \rightarrow \dots \rightarrow \{s_b, s_b^*\} \rightarrow \{s_1, s_1^*\}$, у которого 1-я дуга является выделенной. Отсюда вытекает справедливость достаточности условий теоремы для контуров с выделенными дугами. По аналогии устанавливается справедливость условий теоремы и для случая путей с выделенными дугами, ведущими в контур. На этом завершается доказательство достаточности условий теоремы.

Следующее утверждение позволяет определить порядок ОБПИК-автомата.

Теорема 4.2. Если в проверочном графе $T(A)$ ОБПИК-автомата A длина максимального пути, начальная дуга которого является выделенной, равна t , то порядок ОБПИК-автомата $N(A) = t + 1$.

Доказательство. Пусть в $T(A)$ существует путь, начинающийся в вершине $\{s_i, s_j\}$ и заканчивающийся в $\{s_k, s_l\}$, 1-я дуга которого является выделенной. Наличие его означает существование у автомата A двух входных слов \bar{v}_1 и \bar{v}_2 , различающихся проекциями первых символов и обладающих следующим свойством:

$$\delta(s_i, \bar{v}_1) = s_k \delta(s_j, \bar{v}_2) = s_l \rightarrow pr_{1, \dots, \mu} \lambda(s_i, \bar{v}_1) = pr_{1, \dots, \mu} \lambda(s_i, \bar{v}_1) = pr_{1, \dots, \mu} \lambda(s_j, \bar{v}_2)$$

При распознавании проекции неизвестного входного слова автомата A мы не располагаем информацией об истинном его начальном состоянии. Если на вход автомата A будет подано одно из слов \bar{v}_1 или \bar{v}_2 , то мы не сможем определить и его истинное конечное состояние. В этой ситуации определение проекции 1-го символа входного слова по наблюдаемой реакции длины не более t символов невозможно. Это означает, что порядок автомата A заведомо больше t . Поскольку путь рассматриваемого вида максимален по длине, то из вершины $\{s_k, s_l\}$ графа $T(A)$ не может исходить ни одной дуги. Для автомата A этот факт интерпретируется следующим образом: если автомат A находится в одном из состояний s_k или s_l , то подача любого входного символа и наблюдение соответствующей реакции автомата позволяет всегда однозначно определить истинное начальное состояние из двух

допустимых. Отсюда можно сделать вывод, что знание проекции реакции автомата на неизвестное входное слово длины $t + 1$ всегда позволяет однозначно определить проекцию 1-го символа этого входного символа.

Докажем это утверждение от противного. Пусть существуют два различных входных слова $\bar{\nu}_1 \bar{x}$ и $\bar{\nu}_2 \bar{x}$, где $\bar{\nu}_1$ и $\bar{\nu}_2$ - слова, упомянутые ранее, а \bar{x} - произвольный символ из X . Предположим, что автомат находится в состоянии s_i или s_j . Если на вход автомата A будет подано одно из слов $\bar{\nu}_1 \bar{x}$ или $\bar{\nu}_2 \bar{x}$, то по проекции последнего выходного символа можно определить состояние, в котором автомат находился перед подачей символа \bar{x} . Неоднозначность определения проекции 1-го символа неизвестного входного слова в нашей ситуации означало бы, что

$$\exists \nu_1, \nu_2 \in X * \delta(s_i, \bar{\nu}_1) = \delta(s_j, \bar{\nu}_2) \text{pr}_{1, \dots, \mu} \lambda(s, \bar{\nu}_1) = \text{pr}_{1, \dots, \mu} \lambda(s, \bar{\nu}_2) \rightarrow \text{pr}_{1, \dots, \nu} \bar{\nu}_1 \neq \text{pr}_{1, \dots, \nu} \bar{\nu}_2$$

Тогда в соответствии с определением 3.1 состояния s_i и s_j образуют пару СПИ-состояний. Отсюда следует, что автомат A не может принадлежать классу ОБПИ-автоматов, а в силу теоремы 4.1 он не может принадлежать и классу СБПИК-автоматов. Последний факт противоречит нашему предположению.

Рассмотрим автомат, заданный графом на рис.4.1. Пусть его реакция наблюдается по 1-му и 2-му выходным каналам ($\mu = 2$), а проекция неизвестного входного слова должна восстанавливаться по 1-му каналу ($\nu = 1$). Пусть $S_0 = \{1, 2, 3\}$, тогда проверочный граф для этого автомата изображен на рис.4.2. Легко проверить, что условия теоремы 4.1 для графа на рис.4.2 выполняются, следовательно, рассматриваемый автомат есть ОБПИК. Поскольку длины всех путей в графе $T(A)$, начальные дуги которых являются выделенными, равны $1(t = 1)$, то порядок автомата $N(A) = 2$.

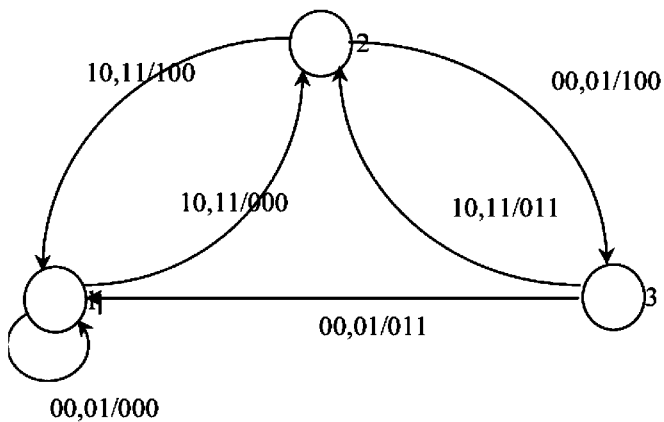


Рис. 4.1.

Рассмотрим вопрос о восстановлении проекции 1-го символа неизвестного входного слова. Из (4.1) следует, что каждой комбинации из $N(A)$ символов, являющихся проекциями реакций автомата A по выходным каналам с номерами $1, \dots, \mu$, однозначно соответствует искомая проекция 1-го символа неизвестного входного слова. В частности, такое соответствие для автомата на [рис.4.1](#) задано [табл. 4.1](#). Подобную таблицу будем интерпретировать далее как совокупность ν таблиц истинности частичных булевых функций, где ν - число каналов, по которым производится восстановление неизвестного входного слова. Строки первого столбца таблицы типа [табл. 4.1](#) содержат двоичные наборы длины $\mu * N(A)$, на которых упомянутые булевы функции считаются заданными, а значение на соответствующем наборе i -й булевой функции полагается равным величине, стоящей в i -й позиции второго столбца таблицы, где $i = \overline{1, \nu}$.

Исходя из предложенной интерпретации, общая схема восстановления некоторой проекции неизвестного входного слова для ОБПИК-автомата порядка $N(A)$ с использованием комбинационного устройства, реализующего совокупность ν частичных булевых функций, изображена на [рис.4.3](#). На этом рисунке D означает элемент единой задержки. На [рис.4.4](#) изображена схема восстановления, которая построена на основе [табл. 4.1](#) для автомата, представленного на [рис.4.1](#).

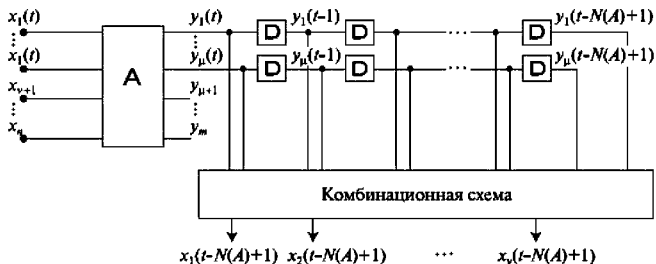


Рис. 4.3.

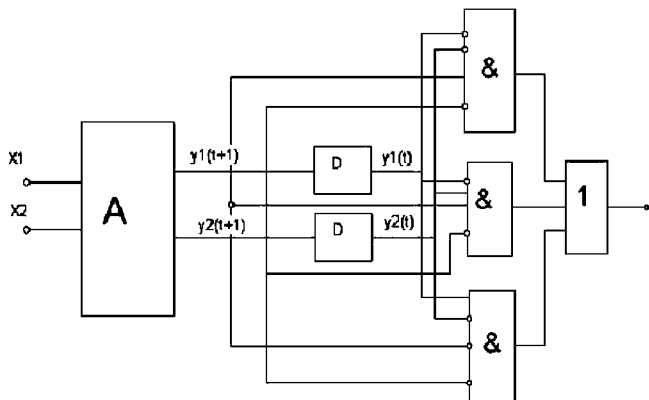


Рис. 4.4.

По аналогии с оптимальными ОБПИ-автоматами, введенными в предыдущей лекции, рассмотрим ОБПИК-автоматы. Для их отыскания по графу автомата A построим проверочный граф $T(A)$, предполагая, что наблюдение реакции ведется по всем m выходным, а восстановление - по всем n входным каналам. Используя $T(A)$, найдем в нем все контуры и все пути, ведущие в эти контуры, образуем на этой основе множество $\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_t\}$ всех входящих в них выделенных дуг. Если $\Gamma = \emptyset$, то по теореме 4.1 автомат A есть ОБПИК и при $\mu = m$ он будет таковым для любого $\nu = \overline{1, n}$. Если $\Gamma \neq \emptyset$, то для каждого $i = \overline{1, t}$ построим множество R_i^* целых чисел, являющихся номерами входных каналов, в которых на дуге γ_i происходит потеря информации в смысле, определенном в предыдущем разделе. По аналогии с ОБПИ-автоматами можно утверждать, что A является ОБПИК для любого подмножества каналов из множества $R = \{1, 2, \dots, n\} \{R_1^* \cup \dots \cup R_t^*\}$ и R есть максимальное по мощности множество каналов, сохраняющее свойство ОБПИК.

Пусть $M = \{1, 2, \dots, \mu\}$. Для $j = \overline{1, \mu}$ строим проверочный граф $T_{Mj}^R(A)$. Если он удовлетворяет условиям теоремы 4.1, то j -й канал является избыточным в смысле, определенном в предыдущей лекции. Пусть $M^* = \{j_1, j_2, \dots, j_s\}$ - множество номеров всех избыточных выходных каналов автомата A . Если $M^* = \emptyset$, то A_M^R является искомым оптимальным автоматом. При $M^* \neq \emptyset$ рассмотрим всевозможные сочетания пар, троек и т. д. номеров каналов из M^* . Обозначим S_λ множество, содержащее λ номеров из M^* , где $\lambda = \overline{2, |M^*|}$. Для каждого из множеств S_λ начиная с $\lambda = 2$ строим проверочный граф $T_{S_\lambda}^R(A)$. Если для всевозможных множеств S_λ при $\lambda = 2$ соответствующие им проверочные графы удовлетворяют условиям теоремы 4.1, то увеличиваем λ на 1 и продолжаем этот процесс. Если на очередном этапе каждый из построенных для всевозможных множеств S_λ проверочных графов не удовлетворяет условиям теоремы 4.1, то процесс прекращается. При этом очевидно, что оптимальными будут все те автоматы $A_{S_\lambda}^R$, рассмотренные на предпоследнем этапе (для них λ имеет значение на 1 меньше, чем на последнем этапе), для которых соответствующие графы удовлетворяют условиям теоремы 4.1.

Вопросы и упражнения

1. Сформулируйте определение обобщенного автомата без потери информации конечного порядка.
2. Приведите критерий принадлежности автомата классу автоматов ОБПИК в терминах проверочного графа.
3. Опишите процедуру определения порядка ОБПИК-автомата.
4. Опишите процедуру построения комбинационной схемы, позволяющей восстанавливать проекцию неизвестного входного слова ОБПИК-автомата.
5. Автомат задан графом на [рис.4.1](#). Пусть множество допустимых начальных состояний автомата есть $\{1, 2, 3\}$. Пусть для наблюдения реакции доступны оба выходных канала, а проекция неизвестного входного слова должна быть восстановлена по 2-му

каналу. Требуется:

- построить проверочный граф этого автомата;
- проверить критерий принадлежности этого автомата классу ОБПИК-автоматов;
- если этот критерий выполняется, то определить порядок автомата.

Преобразования автоматов в автоматы без потери информации

Исследуется задача преобразования произвольного автомата в автомат без потери информации на заданном регулярном множестве входных слов. Предложен метод решения, основанный на расширении выходного алфавита автомата. Описывается процедура построения проверочного графа для заданного автомата и заданного регулярного события. В терминах этого графа дается критерий того, что автомат является автоматом БПИ на словах этого события. Приводится процедура построения по исходному автомату автомата БПИ на заданном регулярном множестве слов за счет выведения дополнительных контрольных точек.

Отсутствие потери информации в автомате является полезным свойством при решении ряда задач. В качестве примера рассмотрим задачу контроля правильности функционирования сети автоматов, изображенной на [рис.5.1](#). В этой сети $A_i (i = 1, \dots, 7)$ - автоматы Мили. Задачу контроля сети автоматов можно свести к задаче контроля одного автомата A , являющегося математической моделью этой сети, и решить ее известными методами. Однако при этом исследуемый автомат A может иметь столь значительную размерность, что известные методы окажутся практически неприемлемыми из-за большого объема вычислений. Другой возможный способ состоит в сведении контроля сети к последовательному контролю каждого из автоматов-компонентов.

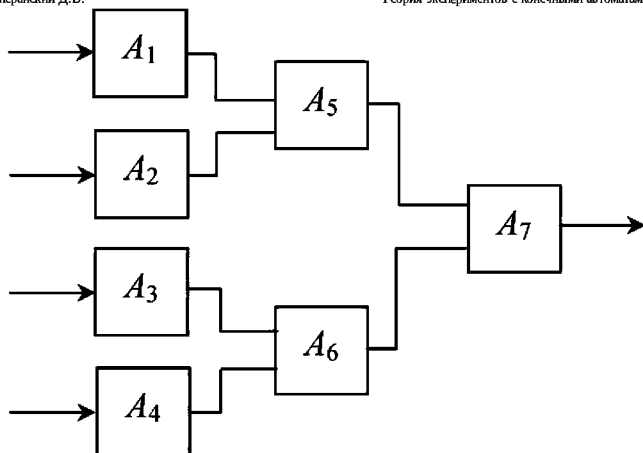


Рис. 5.1.

Второй подход более перспективен с точки зрения практической реализации, но применим он не ко всякой сети. При контроле сети, изображенной на [рис.5.1](#), доступным для наблюдения выходом является только выход автомата A_7 . Поскольку выходы автоматов A_5 и A_6 являются входами автомата A_7 , для контроля каждого из них в отдельности необходимо распознавать входные слова A_7 по наблюдаемой реакции сети. Аналогичным образом обстоит дело и с входными словами автоматов A_5 , A_6 . Таким образом, одно из условий, упрощающих решение задачи контроля рассматриваемой сети, состоит в том, чтобы ее компоненты A_5 , A_6 , A_7 были БПИ-автоматами. Поскольку в конечных автоматах представлены регулярные события и только они, на входы указанных компонентов сети могут поступать только слова, принадлежащие регулярным событиям. С этой точки зрения нам достаточно потребовать, чтобы компоненты A_5 , A_6 , A_7 были БПИ-автоматами лишь на регулярных множествах входных слов (ниже мы уточним это понятие).

В связи с изложенным возникает задача преобразования заданного автомата в БПИ-автомат на регулярном множестве входных слов, которая и будет исследована ниже.

Заметим, что эта задача представляет и самостоятельный интерес, один из возможных ее вариантов рассмотрен в [80].

Обратим внимание, что сформулированная задача будет рассматриваться в предположении, что исходный произвольный автомат будет преобразовываться в БПИ-автомат, восстановление неизвестного входного слова по наблюдаемой реакции которого требует также знания начального состояния. Указанное преобразование будет осуществляться путем выведения из автомата дополнительных контрольных точек, что не влияет на его функционирование, т. е. не изменяет его таблицу переходов, но расширяет выходной алфавит и тем самым модифицирует его таблицу выходов.

Перейдем к подробному рассмотрению упомянутой задачи. Пусть P - регулярное множество непустых слов в алфавите X автомата $A = (S, X, Y, \delta, \lambda)$. Обозначим через Q следующее множество непустых выходных слов автомата A :

$$Q = \bigcup_{s \in S} \lambda_s(P)$$

где

$$\lambda_s(P) = \bigcup_{p \in X^*} \lambda(s, p)$$

Автомат A назовем P -БПИ-автоматом, если и только если

$$\forall s \in S \forall p, q \in P \lambda(s, p) = \lambda(s, q) \delta(s, p) \rightarrow p = q \quad (5.1)$$

Рассмотрим следующую задачу. Пусть задан автомат A и некоторое регулярное множество P непустых слов в алфавите X . Требуется вывести контрольные точки в автомате A таким образом, чтобы полученный в результате автомат был P -БПИ-автоматом.

По аналогии с [23] задачу нахождения минимального числа контрольных точек можно сформулировать следующим образом. Для

автомата A и заданного P найти такое разбиение $\Pi = \{K_1, \dots, K_r\}$ множества $Z = S \times X$, чтобы выполнялись следующие условия:

- для любого состояния s автомата $B = (S, X, Y, \times \Pi, \delta, \Lambda)$ и любой пары слов p и q из P имеет место соотношение

$$\Lambda(s, p) = \Lambda(s, p)\delta(s, p) = \delta(s, p) \rightarrow p = q$$

где $\Lambda(s, x) = (\lambda(s, x), K(s, x)), K(s, x)$, - код класса разбиения Π , содержащего пару (s, x) , в некотором стандартном алфавите;

- у всех разбиений Π' , удовлетворяющих пункту 1 и не равных Π , число классов k таково, что верно неравенство

$$E(\log_t k) \geq E(\log_t r) \quad (5.2)$$

где $E(x) = n$ при $n - 1 < x \leq n$; n - целое положительное число; t - мощность стандартного алфавита; $E(\log_t r)$ - число дополнительных выходных каналов (контрольных точек).

Сформулированная задача относится к числу NP -полных проблем, т. е. методы ее решения содержат перебор. Ниже описан способ вывода контрольных точек в автомате, не содержащий перебора, но в общем случае не минимизирующий их количество.

Пусть R - регулярное выражение события Q . Используя известную технику, описанную, например, в работе [77], преобразуем R таким образом, чтобы оно представляло собой объединение событий, не содержащих итерации. При этом воспользуемся следующими известными соотношениями:

$$\begin{aligned} R_1 * (R_2 Y R_3) &= R_1 * R_2 Y R_1 * R_3, \\ (R_1 Y R_2) * R_3 &= R_1 * R_3 Y R_2 * R_3, \end{aligned}$$

операции объединения и произведения событий, а вхождение итерации события R , обозначаемой через R^* , заменяется выражением $eYRYR^2Y \dots YR^m$. Здесь m - целое положительное число, $R^i = R * R * \dots * R - i$ раз, e - пустое слово.

Вопрос о выборе m должен решаться особо в каждом конкретном случае. На этом мы остановимся ниже. Далее будет использоваться преобразованное указанным способом регулярное выражение R , обозначаемое R' .

Предположим, что $R' = r_1 Y r_2 Y \dots Y r_i$, где $r_i = y_1^{(i)} * y_2^{(i)} * \dots * y_{m_i}^{(i)}$ ($1 \leq i \leq t$) и $y_j^{(i)} \in Y$. Для каждого r_i и каждого $s \in S$ построим ориентированный граф $G_s^A(r_i)$.

Множество вершин этого графа обозначим через T , а построение $G_s^A(r_i)$ будем выполнять следующим образом:

1. Полагаем множество T равным паре $\{s, s\}$, а значение индекса j равным 1. Переходим к пункту 2.
2. По выходному сигналу $y_j^{(i)}$ для каждого элемента $\{s_1, s_2\}$ множества T строим множество $M_{\{s_1, s_2\}}$ всех различных неупорядоченных пар $\{s'_1, s'_2\}$, где $s'_1, s'_2 \in S$, удовлетворяющих следующему условию: в алфавите X существует два таких символа x и x' (необязательно различных), что

$$\delta(s_1, x) = s'_1, \delta(s_2, x') = s'_2, \lambda(s_1.x) = \lambda(s_2.x') = y_j^{(i)} \quad (5.3)$$

Если (5.3) не выполняется ни для каких пар символов из X , то построение заканчивается. В этом случае полагаем, что граф $G_s^A(r_i)$ имеет пустое множество дуг и вершин.

3. К множеству Γ добавляем все множества $M_{\{s_1, s_2\}}$, построенные на предыдущем этапе. Соединим вершины $\{s_1, s_2\}$ и $\{s'_1, s'_2\}$, удовлетворяющие условию (5.3), дугой, началом которой служит вершина $\{s_1, s_2\}$, а концом $\{s'_1, s'_2\}$.
Этой дуге присваивается отметка $(x, x', y_j^{(i)})$. Переходим к пункту 4.
4. Если $j < i$, то $i := i + 1$ и переходим к пункту 2, в противном случае построение графа $G_s^A(r_i)$ заканчивается.

Если в графе $G_s^A(r_i)$ есть дуга с отметкой $(x, x', y_j^{(i)})$, где $x \neq x'$, то такую дугу назовем выделенной.

Используя графы $G_s^A(r_i)$, $1 \leq i \leq r$, построим теперь граф $G_s^A(R')$ по следующим правилам:

1. Множество вершин графа $G_s^A(R')$ является объединением множеств вершин графов $G_s^A(r_i)$ для всех $i = 1, \dots, l$.
2. Из вершины g_1 в вершину g_2 графа $G_s^A(R')$ ведет дуга, если существует такое $i_0, 1 \leq i_0 \leq l$, что в графе $G_s^A(r_{i_0})$ вершины g_1 и g_2 также соединены дугой. Полагаем, что отметка названной дуги в графе $G_s^A(R')$ совпадает с отметкой той же дуги в графе $G_s^A(r_{i_0})$.

Проверочный граф $G^A(R')$, который является конечной целью построений, определим следующим образом:

1. Множество вершин $G^A(R')$ является объединением множеств вершин графов $G_s^A(R')$ для всех $s \in S$.
2. Из вершины g_1 в вершину g_2 графа $G^A(R')$ ведет дуга, если существует такое s_0 , принадлежащее S , что в графе $G_{s_0}^A(R')$

вершины g_1 и g_2 также соединены дугой. Полагаем, что отметка названной дуги в графе $G^A(R')$ совпадает с отметкой той же дуги в графе $G_{s_0}^A(R')$.

3. Начальными вершинами графа $G^A(R')$ считаются все его вершины типа $\{s, s\}$.

Дугу графа $G^A(R')$, исходящую из вершины $\{s, t\}$ с отметкой (x, x', y) , обозначим $(s, t, (x, x', y))$.

Преобразование исходного регулярного выражения R в выражение R' описанным выше способом не является эквивалентным. Покажем, что, несмотря на это, проверочные графы $G^A(R)$ и $G^A(R')$ совпадают.

Рассмотрим событие Q^* и предположим, что Q не содержит итераций. Пусть $Q_j = eYQYQ^2Y \dots YQ^j$. Покажем при этих предположениях, что если событие Q^* в алфавите Y представимо в автомате $A = (S, X, Y, \delta, \lambda)$, $|S| = n$, то проверочные графы $G^A(Q^*)$ и $G^A(Q_m)$ совпадают, где $m = n(n+1)/2$.

Пусть $S_1 = \{s_1, \dots, s_k\}$, $k \leq n$, есть множество всех состояний автомата A , из которых представимо событие Q^* . Выберем произвольным образом состояние s из S_1 и обозначим \bar{F}_s^i множество неупорядоченных пар состояний из S , появляющихся на последнем этапе построения $C_s^A(Q^i)$. Элементы \bar{F}_s^i будем называть финальными.

Графу $G_s^A(Q_i)$, $i = 1, 2, \dots, m$, естественным образом поставим в соответствие автомат Рабина - Скотта $A'_i = (\sum_i, Y, \delta_i, (s, s), f_s^i)$, где $\sum_i \subseteq S \times S$, (s, s) - начальное состояние, $F_s^i = Y_{\nu=1}^i \bar{F}_s^\nu$ - множество финальных состояний, а функция переходов определена следующим образом: $\delta'_i(\{s, t\}, y) = \{\bar{s}, \bar{t}\}$, если в графе $G_s^A(Q^i)$ из

вершины $\{s, t\}$ в вершину $\{\bar{s}, \bar{t}\}$ ведет дуга с отметкой (x, x', y) . Легко видеть, что автомат A'_i , вообще говоря, представляет событие P , такое, что $Y_{j=1}^I Q^j \subseteq P$ и имеет место следующее соотношение:

$$F_s^1 \subseteq F_s^2 \subseteq \dots \subseteq F_s^I \subseteq \dots \quad (5.4)$$

Поскольку автомат A имеет n состояний, то число различных неупорядоченных пар $\{s, t\}, s, t, \in S$, являющихся состояниями автомата A'_i , не превышает величины $n(n+1)/2$. Тогда из (5.4) вытекает

$$|F_s^1| \leq |F_s^2| \leq \dots \leq |F_s^i| \leq \dots \leq n(n-1)/2$$

и, следовательно, существует такое $i_0 \leq n(n+1)/2$, что $|F_s^{i_0}| = |F_s^{i_0+k}|$ при любом целом положительном k . Но тогда на основании (5.4) можно сделать вывод, что $F_s^{i_0} = F_s^{i_0+1} = F_s^{i_0+k} = \dots$. Последнее соотношение означает, что автоматы $A'_{i_0}, A'_{i_0+1}, \dots$ представляют одно и то же событие, следовательно, графы $G_s^A(Q^m), m = n(n+1)/2$, и $G_s^A(Q^*)$ совпадают. Вследствие произвольности выбора s из S и способа построения проверочного графа это же утверждение справедливо для графов $G^A(Q^m)$ и $G^A(Q^*)$.

Рассуждения, аналогичные приведенным, показывают, что если Q содержит итерации некоторых событий, то эти итерации также можно заменить объединением конечного числа степеней событий и при этом такая замена не приводит к изменению проверочного графа автомата.

Таким образом, нами доказана справедливость следующего утверждения.

Теорема 5.1. Если событие Q в алфавите Y , регулярное выражение которого есть R , представимо в автомате $A = (S, X, Y, \delta, \lambda), |S| = n$, то проверочные графы $G^A(R)$ и

$G^A(R')$, где R' получено из R путем замены каждого вхождения итерации события T на $Y_{i=1}^{n(n+1)/2} T^i$, совпадают.

Из доказательства этой теоремы следует, что если R' является объединением всех возможных слов, принадлежащих R , длина которых не превосходит величины $n(n+1)/2$, то графы $G^A(R')$ и $G^A(R)$ для автомата A совпадают. Это замечание позволяет упростить практическое построение графа $G^A(R)$.

Сформулированная ниже теорема будет служить критерием, используемым при преобразовании произвольного автомата в P -БПИ-автомат. Ее доказательство аналогично доказательству теоремы 2 из [74].

Пусть R - регулярное выражение события $Q = Y_{s \in S} \lambda_s(P)$.

Теорема 5.2. Для того чтобы автомат A был P -БПИ-автоматом, необходимо и достаточно, чтобы в его проверочном графе $G^A(R)$ для любого слова $p \in P$ каждый соответствующий ему путь не содержал выделенной дуги, исходящей из вершины $\{s, s\}$, а также дуги, заходящей в вершину такого же типа.

Рассматриваемую задачу отыскания множества выходных контрольных точек можно разбить на две следующие.

Задача 1. В проверочном графе $G^A(R)$ найти такое множество L его дуг, чтобы полученный в результате их удаления новый граф удовлетворял условиям теоремы 5.2.

Задача 2. По заданному множеству L дуг построить такое максимальное разбиение Π множества $Z = S \times Y$, чтобы выполнялись условия:

- если дуга $(s, t, (x_i, x_j, y))$ графа $G^A(R)$ принадлежит множеству L , то (s, x_i) и (t, x_j) находятся в разных классах

разбиения;

- для всех Π' , удовлетворяющих условию (а), выполняется условие (5.2).

Обе сформулированные задачи принадлежат к числу сложных комбинаторных задач, методы решения которых, отличные от перебора, пока не известны. В работе [23] предложены методы решения задач, подобных задачам 1 и 2, не гарантирующие минимальности, но не содержащие перебора.

Решение задачи 2 можно получить по алгоритму $A1$ из [23], если положить в нем $Z = N$, а симметричное бинарное отношение $\varphi \subset Z \times Z$ определить следующим образом: $((s, x_i), (t, x_j)) \in \varphi$, если и только если существует такое $y \in Y$, что $\lambda(s, x_i) = \lambda(t, x_j) = y, (s, t, (x_i, x_j, y)) \neq L$.

Перейдем к рассмотрению первой задачи. По графу $G^A(R')$ построим множество C всех простых путей длины большей 1, удовлетворяющих следующим условиям:

- путь начинается выделенной дугой, исходящей из вершины типа $\{s, s\}$, и заканчивается дугой, заходящей в вершину такого же типа;
- ни один путь из C не должен являться отрезком никакого другого пути из C ;
- каждый путь в C должен соответствовать некоторому отрезку выходного слова, принадлежащего событию Q ;
- каждый путь в C не содержит выделенной дуги, соединяющей вершины типа $\{s, s\}$.

Каждой дуге графа $G^A(R')$, входящей в состав путей из C , присвоим вес, равный числу путей в C , содержащих эту дугу.

Пусть F' - множество выделенных дуг графа $G^A(R')$, каждая из которых соединяет вершины типа $\{s, s\}$.

Рассмотрим проверочный граф $G^A(R')$ автомата $A = (S, X, Y, \delta, \lambda)$, где R - регулярное выражение события $Q = \bigcup_{s \in S} \lambda_s(P)$. Выделим все слова P из P , такие, что соответствующие им пути в графе $G^A(R')$ содержат выделенные дуги, исходящие из вершины типа $\{s, s\}$, а также дуги, заходящие в вершину такого же типа. Множество всех таких путей обозначим через M . Легко видеть, что множество всех путей C и множество выделенных дуг F принадлежат M : все пути из M , не совпадающие с путями из C и F , содержат их в качестве отрезков.

Отсюда следует, что если в графе $G^A(R')$ отсутствуют пути из множеств C и F , то и пути из множества M в нем также отсутствуют. Таким образом, из теоремы 5.2 вытекает, что для преобразования автомата A в P -БПИ-автомат достаточно из графа $G^A(R')$ удалить все пути множества C и все выделенные дуги множества F .

Перейдем к рассмотрению алгоритма, позволяющего по множествам C и F строить такое множество L дуг графа $G^A(R')$, чтобы после их удаления из этого графа в нем отсутствовали все пути из C и все дуги из F .

Алгоритм выбора множества дуг L можно сформулировать следующим образом.

Полагаем $L = F$, где F - множество выделенных дуг графа $G^A(R')$, каждая из которых соединяет вершины типа $\{s, s\}$. Если $F = \emptyset$, то перейти к пункту 2, иначе к пункту 4.

1. Из множества C выделяем множество F_1 всех дуг, имеющих максимальный вес.
2. Добавляем в L одну дугу из F_1 , выбранную произвольно.
3. Из графа $G^A(R')$ удаляем дуги, входящие в множество $F = \emptyset$.
Вновь полученный граф обозначаем также $G^A(R')$.

4. Из множества C удаляем пути, содержащие дуги из множества C , построенного на предыдущем этапе. Если $C = \emptyset$, то построение L окончено, иначе переходим к пункту 6.
5. Произво

Пусть $L = \{(c_1, d_1, (x_{k_1}, x'_{k_1}, y_{k_1})), \dots, (c_\nu, d_\nu, (x_{k_\nu}, x'_{k_\nu}, y_{k_\nu}))\}$ получено по описанному алгоритму, причем дуги в L занумерованы в том порядке, в котором они присоединялись к L в соответствии с алгоритмом. Пусть $C - i$ - подмножество всех путей из множества C , проходящих через дугу $(c_i, d_i, (x_{k_i}, x'_{k_i}, y_{k_i}))$.

Лемма 5.1. Для любых $a \neq b, 1 \leq a, b \leq \nu$ справедливо соотношение $C_a \not\subset C_b$.

Доказательство. Для определенности положим, что $a < b$. На основании пункта 4 алгоритма из $G^A(R')$ удаляется дуга $(c_a, d_a, (x_{k_a}, x'_{k_a}, y_{k_a}))$, что влечет за собой удаление из этого же графа и всех путей из множества C_a . Это же множество путей удаляется из C на основании пункта 5 алгоритма. Присоединение дуги $(c_b, d_b, (x_{k_b}, x'_{k_b}, y_{k_b}))$ к множеству L на основании пункта 2 алгоритма возможно лишь в том случае, если в множестве C , полученном на предыдущем этапе, есть по крайней мере один путь, проходящий через эту дугу. Поскольку C после проведенной корректировки заведомо не содержит C_a , отсюда следует, что $C_a \not\subset C_b$. Обратное соотношение $C_b \not\subset C_a$ тоже справедливо, так как в противном случае после удаления дуги $(c_a, d_a, (x_{k_a}, x'_{k_a}, y_{k_a}))$ из графа $G^A(R)$ все пути множества C_b были также удалены, что противоречит предположению о наличии дуги $(c_b, d_b, (x_{k_b}, x'_{k_b}, y_{k_b}))$ во множестве L .

В качестве следствия леммы получаем, что для любого $I, 1 \leq I \leq \nu$, справедливы соотношения $\bigcup_{j=1}^{I-1} C_j \not\subset C_I$ и $C_I \not\subset \bigcup_{j=1}^{I-1} C_j$.

Отметим следующий очевидный факт: множество выделенных дуг F заведомо принадлежит решению задачи 1 независимо от того, получено ли оно по описанному алгоритму или иным способом.

Из изложенного выше следует, что по предложенному алгоритму строится множество L , не являющееся избыточным в том смысле, что на каждом его шаге из графа $G^A(R)$ удаляется непустое множество путей, принадлежащих C . Если $L = F$, то мощность L минимальна и L строится за один шаг.

Рассмотренная нами задача построения множества L дуг эквивалентна задаче о покрытии, или задаче построения минимальной формы булевой функции. Если проводить аналогию с последней задачей, то построенное с помощью предложенного алгоритма множество L соответствует в силу леммы 5.1 приведенной системе простых импликант. Таким образом, вопрос о близости минимального по мощности решения задачи 1 и решения, построенного с помощью предложенного алгоритма, эквивалентен вопросу о близости приведенной системы простых импликант к минимальной приведенной системе импликант (минимальной форме булевой функции). Продолжая аналогию, заметим, что из построенного по предложенному алгоритму множества L некоторые дуги можно удалять способом, подобным способу удаления импликант из приведенной системы [21].

Проиллюстрируем предложенный способ преобразования автомата в P -БПИ-автомат на примере.

Рассмотрим автомат A , заданный табл. 5.1. Пусть $R = a(ba^*) * a$.

Таблица 5.1.

s\х	0	1	2	3
1	2, a	3, a	2, a	1, c
2	4, a	1, b	2, c	3, c
3	2, b	4, a	3, d	2, f
4	3, a	2, b	4, f	3, d

Вследствие замечания к теореме 5.1 в рассматриваемом примере нам достаточно ограничиться при замене итерации по изложенному выше способу значением m , равным 8, так как в выражении R до и после итерации $(ba^*)^*$ стоит символ a и с учетом этого длина слов в R' будет доведена до 10 (значение $n(n+1)/2$ при $n=4$) и при этом значении m . В результате преобразования R получим следующее выражение:

$$\begin{aligned}
 R' = & a^2 \cup aba^2 \cup aba^3 \cup aba^4 \cup aba^5 \cup aba^6 \cup aba^7 \cup aba^8 \cup \\
 & \cup a(ab)^2a \cup a(ba)^2a^2 \cup a(ba)^2a^3 \cup a(ba)^2a^4 \cup a(ba)^2a^5 \cup \\
 & \cup a(ba^2)^2 \cup aba^2ba^3 \cup aba^2ba^4 \cup aba^2ba^5 \cup aba^3ba^2 \cup aba^3ba^3 \cup \\
 & \cup aba^3ba^4 \cup aba^4ba^2 \cup aba^4ba^3 \cup aba^5ba^2 \cup a(ba)^3a \cup a(ba)^3a^2 \cup \\
 & \cup a(ba)^3a^3 \cup a(ba)^2aba^2 \cup (ab)^2a^2ba^3 \cup (ab)^2a^3ba^2 \cup aba^2(ba)^2a \cup \\
 & \cup aba^2(ba)^2a^2 \cup aba^2(ba^2)^2 \cup aba^3(ba)^2a
 \end{aligned}$$

Граф $G^A(R')$, или, что то же самое, $G^A(R)$, имеет вид, изображенный на рис. 5.2, где двойные стрелки соответствуют выделенным дугам.

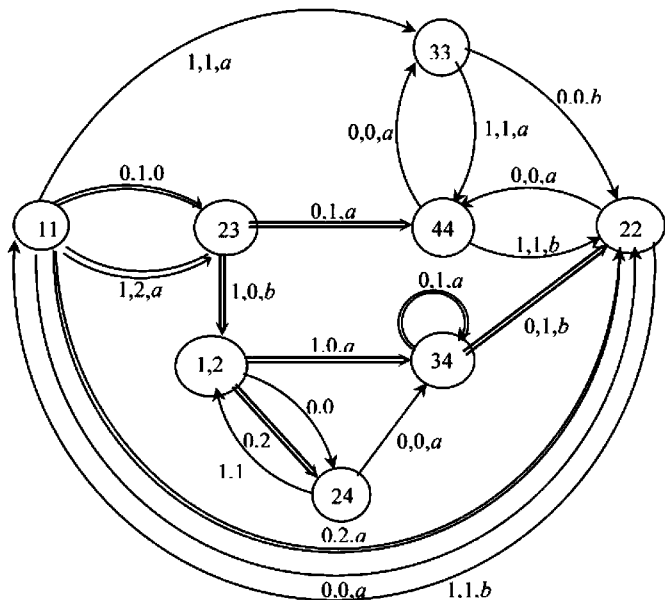


Рис. 5.2.

Множество путей C в нашем случае таково:

$$\begin{aligned}
 &1, 1 \xrightarrow{0,1,a} 2, 3 \xrightarrow{0,1,a} 4, 4; \\
 &1, 1 \xrightarrow{1,2,a} 2, 3 \xrightarrow{0,1,a} 4, 4; \\
 &1, 1 \xrightarrow{0,1,a} 2, 3 \xrightarrow{0,1,b} 1, 2 \xrightarrow{1,0,a} 3, 4 \xrightarrow{0,1,b} 2, 2; \\
 &1, 1 \xrightarrow{1,2,a} 2, 3 \xrightarrow{1,0,b} 1, 2 \xrightarrow{0,1,a} 3, 4 \xrightarrow{0,1,b} 2, 2 \\
 &1, 1 \xrightarrow{0,1,a} 2, 3 \xrightarrow{0,1,b} 1, 2 \xrightarrow{0,0,a} 2, 4 \xrightarrow{0,0,a} 3, 4 \xrightarrow{0,1,b} 2, 2 \\
 &1, 1 \xrightarrow{1,2,a} 2, 3 \xrightarrow{1,0,b} 1, 2 \xrightarrow{0,0,a} 2, 4 \xrightarrow{0,0,a} 3, 4 \xrightarrow{0,1,b} 2, 2;
 \end{aligned}$$

$$\begin{aligned}
1, 1 &\xrightarrow{0,1,a} 2, 3 \xrightarrow{1,0,b} 1, 2 \xrightarrow{1,0,a} 3, 4 \xrightarrow{0,1,a} 3, 4 \xrightarrow{0,1,b} 2, 2; \\
1, 1 &\xrightarrow{1,2,a} 2, 3 \xrightarrow{1,0,b} 1, 2 \xrightarrow{1,0,a} 3, 4 \xrightarrow{0,1,a} 3, 4 \xrightarrow{0,1,b} 2, 2; \\
1, 1 &\xrightarrow{0,1,a} 2, 3 \xrightarrow{0,1,b} 1, 2 \xrightarrow{0,2,a} 2, 4 \xrightarrow{0,0,a} 3, 4 \xrightarrow{0,1,b} 2, 2; \\
1, 1 &\xrightarrow{1,2,a} 2, 3 \xrightarrow{1,0,b} 1, 2 \xrightarrow{0,2,a} 2, 4 \xrightarrow{0,0,a} 3, 4 \xrightarrow{0,1,b} 2, 2.
\end{aligned}$$

В соответствии с предложенным выше алгоритмом выбора L на первом шаге получаем $L = \{(1, 1, (0, 2, a))\}$.

Далее, максимальный вес, равный 8, в множестве C имеют дуги $(3, 4, (0, 1, b))$ и $(2, 3, (1, 0, b))$; тогда полагаем $L = \{(1, 1, (0, 2, a)): (3, 4, (0, 1, b))\}$, выбрав первую из названных дуг. После корректировки в множестве C остается два первых пути из списка, приведенного выше. Поскольку вес дуги $(2, 3, (0, 1, a))$ максимален (он равен 2), то добавляем ее в L и на этом процесс построения L заканчивается, так как множество C после повторной корректировки пустое. Таким образом, $L = \{(3, 4, (0, 2, a)): (2, 3, (0, 1, a))\}$

Используя алгоритм $A1$ из [23], получаем следующее разбиение множества $Z = S \times X$:

$$\overline{\Pi = (1, 0), (1, 1), (1, 3), (2, 0), (2, 1), (2, 2), (2, 3), (3, 0), (3, 2), (3, 3), (4, 0), (4, 2), (4, 3); (1, 2), (3, 1), (4, 1)}$$

Это разбиение имеет два класса, следовательно, для кодирования символов классов достаточно, например, одной двоичной переменной. Пусть $K_1 = 0$ и $K_2 = 1$, тогда автомат с контрольными точками будет представлен в табл. 5.2.

Таблица 5.2.

s \ x	0	1	2	3
1	2, a 0 3, a 0 2, a 1 1, c 0			
2	4, a 0 1, b 0 2, c 0 3, c 0			
3	2, b 0 4, a 1 3, d 0 2, f 0			

4 3, a 0 2, b 1 4, f 0 3, d 0

Отметим, что идея изложенного метода преобразования автоматов очевидным образом может быть использована и для решения следующей задачи: для заданного регулярного множества P непустых входных слов вывести выходные контрольные точки в автомате таким образом, чтобы каждое слово из P стало диагностическим для автомата с контрольными точками.

Предложенный метод решения задачи 1 в качестве одного из этапов содержит построение путей множества C , выполняемого визуально по проверочному графу. Такое построение становится тем более затруднительным, чем большее число состояний имеет данный автомат. В связи с этим рассмотрим формализованный метод построения такого множества L' дуг автомата A , что удаление их превращает автомат A в P -БПИ-автомат. В основу метода положим представление автомата в виде матрицы переходов и выполнение операций над этой матрицей, описанных, например, в работе [18].

Напомним, что для автомата A , имеющего n состояний, матрица переходов состоит из n строк и n столбцов и обозначается $[A]$. Пусть $S = \{s_1, \dots, s_n\}$ - множество состояний автомата A и пусть $s_i \xrightarrow{x,y} s_j$ обозначает дугу графа автомата A , направленную от s_i к s_j под воздействием входного сигнала x с выдачей выходного сигнала y . Элемент (i, j) матрицы $[A]$, т. е. содержимое клетки, расположенной на пересечении i -й строки и j -го столбца, содержит все дуги, ведущие из вершины s_i в вершину s_j . Если такие дуги отсутствуют, то элемент (i, j) полагается равным нулю.

Условимся считать, что обозначение k -го состояния s_k приписано k -й строке и k -му столбцу матрицы переходов.

В работе [18] введена операция возведения в степень матрицы переходов $[A]$ и доказано, что элемент (i, j) матрицы $[A]^k, k = 1, 2, \dots$, является множеством всех путей длины k , ведущих из состояния s_i в состояние s_j автомата A . Этот факт и

является основой для формализованного нахождения тех путей в графе автомата A , которые являются причиной потери информации.

Построение множества L' дуг, упомянутого выше, будет проводиться в виде пошагового процесса. Одновременно с этим будет проводиться построение "каркаса" разбиения Π множества Z .

Вначале произведем следующие преобразования над матрицей переходов автомата A :

- удаляем из $[A]$ все дуги, имеющие в качестве выходных символы, не содержащиеся в выражении Q ;
- если элемент (i, j) матрицы $[A]$ содержит две дуги с различными входными, но одинаковыми выходными символами, то одну из этих дуг, выбираемую произвольно, включаем в множество L' и удаляем ее из матрицы $[A]$.

Отметим при этом, что удаленная и оставленная в матрице $[A]$ дуги должны находиться в разных классах разбиения Π (это и есть информация о "каркасе" разбиения Π).

Описанный процесс продолжается далее до тех пор, пока элемент (i, j) не будет содержать ни одной пары дуг с указанным выше свойством.

Легко видеть, что удаленные в соответствии с пунктом (а) дуги заведомо таковы, что путь в графе автомата A , соответствующий любому слову из заданного регулярного множества P , не может содержать их.

Удаление из матрицы $[A]$ дуг согласно пункту (б) соответствует удалению выделенных дуг множества F' из проверочного графа $G^A(R)$. Очевидно, что пара дуг графа автомата A с различными входными, но одинаковыми выходными символами является выделенной дугой проверочного графа $G^A(R)$.

Преобразованную описанным способом матрицу переходов будем, как и ранее, обозначать $[A]$.

Описываемые ниже преобразования будут производиться над матрицами $[A]^k$ при $k = 1, 2, \dots$, получаемыми из $[A]^{k-1}$ путем умножения ее слева на преобразованную матрицу переходов $[A]$.

Итак, рассмотрим матрицу переходов $[A]^k$. Произведем над ней следующие преобразования:

- Если элемент (i, j) матрицы $[A]^k$ содержит два пути, которым соответствуют различные входные, но одинаковые выходные слова, то каждой дуге, входящей в состав этих путей, присваиваем вес, равный числу вхождений этой дуги во все пути матрицы переходов $[A]^k$. Дугу с максимальным весом включаем в множество L' и удаляем из $[A]^k$ все пути, содержащие эту дугу. Из матрицы переходов $[A]$ дуга с максимальным весом также удаляется.

Этот процесс продолжается далее до тех пор, пока элемент (i, j) не будет содержать ни одной пары путей с описанным выше свойством.

- Среди всех путей матрицы $[A]^k$ находим пути, соответствующие входным (выходным) словам, которые не могут являться отрезками никаких слов из заданного множества $P(Q)$. Указанные пути удаляются из матрицы переходов $[A]^k$.

Как уже было отмечено, удаление дуг множества L' из графа автомата A соответствует удалению дуг некоторого множества \tilde{L} из проверочного графа $G^A(R)$.

Пусть выполнены описанные выше преобразования над матрицей

переходов $[A]$ и при этом получено множество L' дуг графа автомата A . Удаляем из $G^A(R)$ множество \tilde{L} дуг, соответствующее множеству L' , и проверяем выполнение условий теоремы 5.2. Если условия выполняются, то преобразования матриц переходов на этом завершаются. Для построения искомого разбиения Π можно далее воспользоваться уже описанным методом. Если же условия теоремы 5.2 не выполняются, то вычисляем следующую степень матрицы переходов $[A]$ и выполняем над ней описанные преобразования. Затем вновь следует проверка условий теоремы 5.2 и процесс продолжается аналогичным образом.

Из описанных преобразований следует, что выполнение условий теоремы 5.2 будет достигнуто за конечное число шагов.

Продемонстрируем предложенный метод на примере, который уже был рассмотрен выше. Итак, пусть автомат A задан табл. 5.1, а множество Q задается выражением $R = a(ba^*) * a$.

Исходная матрица переходов имеет следующий вид:

$$[A] = \begin{bmatrix} 1 \xrightarrow{1,c} 1 & 1 \xrightarrow{2,c} 2 & 1 \xrightarrow{1,a} 3 & 0 \\ & 1 \xrightarrow{0,a} 2 & & \\ 2 \xrightarrow{1,b} 1 & 2 \xrightarrow{2,c} 2 & 2 \xrightarrow{3,c} 3 & 2 \xrightarrow{0,a} 4 \\ 0 & 3 \xrightarrow{0,b} 2 & 3 \xrightarrow{2,d} 3 & 3 \xrightarrow{1,a} 4 \\ & 3 \xrightarrow{0,f} 2 & & \\ 0 & 4 \xrightarrow{1,b} 2 & 4 \xrightarrow{0,a} 3 & 4 \xrightarrow{2,f} 4 \\ & & 4 \xrightarrow{3,d} 3 & \end{bmatrix}$$

Преобразованная матрица после удаления дуг из элементов (1,1), (2,2), (2,3), (3,2), (3,3), (4,3), (4,4) на основании пункта (а) и дуги $1 \xrightarrow{2,a} 2$ из элемента (1,2) на основании пункта (б) примет вид

$$[A] = \begin{bmatrix} 0 & 1 \xrightarrow{0,a} 2 & 1 \xrightarrow{1,a} 3 & 0 \\ 2 \xrightarrow{1,b} & 0 & 0 & 2 \xrightarrow{0,a} \\ 0 & 3 \xrightarrow{0,b} 3 & 0 & 3 \xrightarrow{1,a} 4 \\ 0 & 4 \xrightarrow{1,b} 2 & 4 \xrightarrow{0,a} 3 & 0 \end{bmatrix}$$

Дуга $1 \xrightarrow{2,a} 2$ включается в множество L' , причем она должна быть в разных классах разбиения Π с дугой $1 \xrightarrow{0,a} 2$. Указанной паре дуг в проверочном графе $G^A(R)$ соответствует выделенная дуга $\{1, 1\} \xrightarrow{0,2,a} \{2, 2\}$.

После удаления выделенной дуги из $G^A(R)$ этот граф не удовлетворяет условиям теоремы 5.2, поэтому вычислим матрицу переходов $[A]^2$:

$$[A]^2 = \begin{bmatrix} 1 \xrightarrow{0,a} 2 \xrightarrow{1,b} 1 & 1 \xrightarrow{1,a} 3 \xrightarrow{0,b} 2 & 0 & 1 \xrightarrow{0,a} 2 \xrightarrow{0,a} 4 \\ & & & 1 \xrightarrow{1,a} 3 \xrightarrow{1,c} 4 \\ 0 & 2 \xrightarrow{1,b} 1 \xrightarrow{0,a} 2 & 2 \xrightarrow{1,b} 1 \xrightarrow{1,a} 3 & 0 \\ & 2 \xrightarrow{0,a} 4 \xrightarrow{1,b} 2 & 2 \xrightarrow{0,a} 4 \xrightarrow{0,a} 3 & \\ 3 \xrightarrow{0,b} 2 \xrightarrow{1,b} 1 & 3 \xrightarrow{1,a} 4 \xrightarrow{1,b} 2 & 3 \xrightarrow{1,a} 4 \xrightarrow{0,a} 3 & 3 \xrightarrow{0,b} 2 \xrightarrow{0,a} 4 \\ 4 \xrightarrow{1,b} 2 \xrightarrow{1,b} 1 & 4 \xrightarrow{0,a} 3 \xrightarrow{0,b} 2 & 0 & 4 \xrightarrow{1,b} 2 \xrightarrow{0,a} 4 \\ & & & 4 \xrightarrow{0,a} 3 \xrightarrow{1,a} 4 \end{bmatrix}$$

Удалим дуги из элементов (3,1) и (4,1) на основании пункте 2, так как слово bb не принадлежит событию $a(ba^*) * a$. Пути элемента (1,4) удовлетворяют условию, сформулированному в пункте 1. Максимальный вес (равный 5) среди всех дуг, входящих в состав этих путей, имеет дуга $2 \xrightarrow{0,a} 4$. Включаем эту дугу в множество L' и удаляем из $[A]^2$ все пути, содержащие ее. После этого преобразования матрица примет следующий вид:

$$[A]^2 = \begin{bmatrix} 1 \xrightarrow{0,a} 2 \xrightarrow{1,b} 1 & 1 \xrightarrow{1,a} 3 \xrightarrow{0,b} 2 & 0 & 1 \xrightarrow{1,a} 3 \xrightarrow{1,c} 4 \\ 0 & 2 \xrightarrow{1,b} 1 \xrightarrow{0,a} 2 & 2 \xrightarrow{1,b} 1 \xrightarrow{1,a} 3 & 0 \\ 0 & 3 \xrightarrow{1,a} 4 \xrightarrow{1,b} 2 & 3 \xrightarrow{1,a} 4 \xrightarrow{0,a} 3 & 0 \\ 0 & 4 \xrightarrow{0,a} 3 \xrightarrow{0,b} 2 & 0 & 4 \xrightarrow{0,a} 3 \xrightarrow{1,a} 4 \end{bmatrix}$$

Удаление дуги $2 \xrightarrow{0,a} 4$ из графа автомата A соответствует удалению выделенных дуг

$\{2, 3\} \xRightarrow{0,1,a} \{4, 4\}$; $\{1, 2\} \xRightarrow{0,2,a} \{2, 4\}$; $\{1, 2\} \xRightarrow{1,0,a} \{3, 4\}$
в проверочном графе $G^A(R)$, а также дуги $\{1, 2\} \xrightarrow{0,0,a} \{2, 4\}$.

Легко проверить, что после удаления названных дуг граф $G^A(R)$ удовлетворяет условиям теоремы 5.2.

Таким образом, изменив выходные сигналы при переходах $1 \xrightarrow{2,a} 2$ и $2 \xrightarrow{0,a} 4$ в графе автомата A , например, следующим образом: $1 \xrightarrow{2,a1} 2, 2 \xrightarrow{0,a1} 4$, мы удовлетворяем требованию, чтобы дуги $1 \xrightarrow{2,a1} 2, 1 \xrightarrow{0,a} 2$, принадлежали разным классам разбиения.

Поскольку число полученных классов разбиения равно 2, то для кодирования символов классов достаточно одной двоичной переменной. Это означает, что у автомата A достаточно вывести одну дополнительную выходную контрольную точку. Таблицей автомата с контрольными точками в силу сказанного выше является табл. 5.3.

Таблица 5.3.

s\х	0	1	2	3
1	2, a 0 3, a 0	2, a 1 1, c 0		
2	4, a 1 1, b 0	2, c 0 3, c 0		
3	2, b 0 4, a 0	3, d 0 2, f 0		
4	3, a 0 2, b 0	4, f 0 3, d 0		

Итак, мы получим еще один вариант автомата с контрольными точками,

являющегося P -БПИ-автоматом, где событие Q представлено выражением $R = a(ba^*) * a$.

Вопросы и упражнения

1. Дайте определение автомата БПИ на заданном регулярном множестве его входных слов.
2. Сформулируйте основную идею метода решения задачи преобразования автомата в автомат БПИ на регулярном множестве слов.
3. Сформулируйте две вспомогательных задачи, к решению которых сводится проблема исследуемого преобразования автомата в автомат БПИ.
4. Опишите кратко основные идеи методов решения двух вспомогательных задач.

Эксперименты по контролю функции выходов инициального автомата

Исследуется задача контроля функции инициального автомата, которая с математической точки зрения сводится к задаче построения специального обхода графа этого автомата. Приведен критерий существования такого обхода, который реализуется в процессе подачи на автомат так называемого характеристического слова. Доказываются различные утверждения, касающиеся оценок длин кратчайших обходов графа и характеристических слов.

Контроль функции выходов инициального автомата с использованием простого безусловного эксперимента

Известно, что ДУ с памятью, описываемое математической моделью конечного автомата Мили, всегда может быть реализовано в виде, изображенном на [рис.6.1](#). Здесь блок C реализует функцию выходов и является комбинационным устройством, а блок B представляет собой запоминающее устройство.

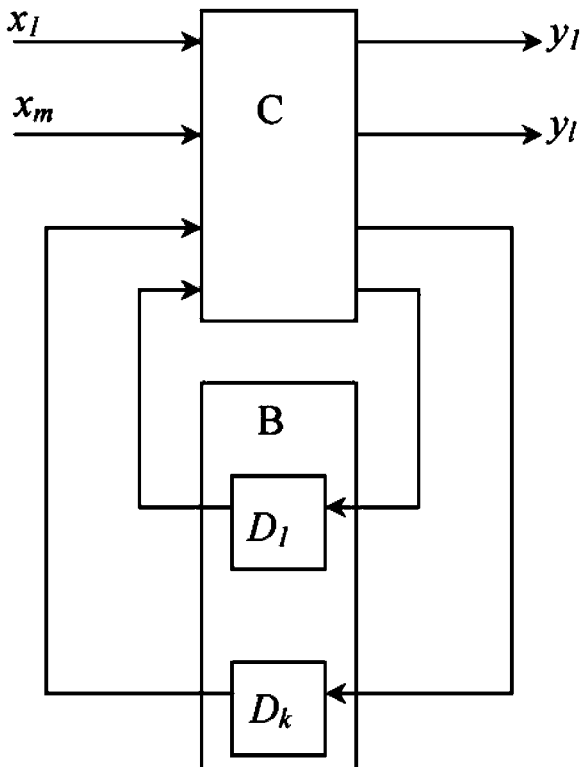


Рис. 6.1.

В общем случае причинами неисправности ДУ могут быть как неисправности в блоке C (блоке B), так и неисправности, возникшие в обоих блоках.

В этой лекции рассматриваются эксперименты, ориентированные на

контроль функционирования блока C , т. е. функции выхода автомата. Мы займемся исследованием этого частного случая экспериментов потому, что для проверки памяти ДУ, т. е. блока B , разработаны хорошо зарекомендовавшие себя специальные методы. Конечно, для проверки функционирования блока C можно применить контрольные эксперименты общего вида, однако в рассмотренной ситуации методы их построения не учитывают ее специфики и будут порождать тесты, длина которых заведомо больше той, которая минимально необходима.

Естественно, что методы контроля зависят от класса рассматриваемых автоматов и применяемых средств. Ниже исследуются методы контроля для инициальных и неинициальных автоматов, а в качестве средств используются простые и кратные безусловные эксперименты. Характерной особенностью всех рассматриваемых здесь задач контроля является их сводимость к построению различных обходов графов автомата, т. е. путей по графу, проходящих через все его дуги.

Поскольку в рамках теории графов рассматриваются более общие конструкции, чем автоматные графы, полезно попытаться найти решение задачи построения обхода графа, не обязательно являющегося автоматным. Из решения этой более общей задачи решение соответствующей автоматной задачи будет следовать как частный случай.

В связи с этим при исследовании рассматриваемых ниже задач будем придерживаться следующей схемы: после постановки задачи контроля она будет сводиться к задаче построения соответствующего обхода графа. Далее в терминах теории графов будут получены условия существования решения, алгоритмы построения решения и оценки длины получаемых обходов.

При решении задач контроля автоматов обычно принимается предположение, что в процессе проведения эксперимента дополнительных неисправностей в автомате не возникает. В рассматриваемых ниже задачах это предположение также остается в силе.

Рассмотрим следующую задачу. Пусть в распоряжении экспериментатора находится один экземпляр автомата Мили A , у

которого известны входной алфавит $X = \{x_1, \dots, x_m\}$, выходной алфавит $Y = \{y_1, \dots, y_l\}$, множество состояний $S = \{s_0, s_1, \dots, s_{n-1}\}$, где s_0 - выделенное состояние, называемое начальным, и функция переходов $\delta : S \times X \rightarrow S$. Требуется построить простой безусловный эксперимент, позволяющий распознать функцию выходов λ автомата A или эквивалентного ему автомата.

Умение решать сформулированную задачу позволит решать и задачу контроля функции выходов инициального автомата. В самом деле, если нам удастся распознать функцию выходов λ' автомата, с которым производится эксперимент, то сравнение ее с функцией λ автомата A , играющего роль исправного, позволит сделать вывод об исправности λ' в исследуемом автомате. Если же изменения, появившиеся в функционировании блока λ , таковы, что исследуемый автомат эквивалентен A , то естественно игнорировать эти изменения и считать исследуемый автомат исправным. Именно этим мотивируется требование распознаваемости функции выходов с точностью до эквивалентности.

Очевидно, что функция выходов будет распознана, если в результате проведения эксперимента для любого состояния этого автомата будет определена его реакция на любой входной символ.

В соответствии с определением простого безусловного эксперимента его проведение должно включать построение соответствующего входного слова, наблюдение реакции автомата на это входное слово и, наконец, сравнение на основе полученной реакции функции выходов исследуемого автомата с эталоном и вывод заключения об исправности ее.

Контроль функции выходов исследуемого автомата может быть выполнен с помощью так называемого контролирующего автомата, связанного с исследуемым так, как это изображено на [рис.6.2](#), где A - исследуемый автомат, K - контролирующий автомат.

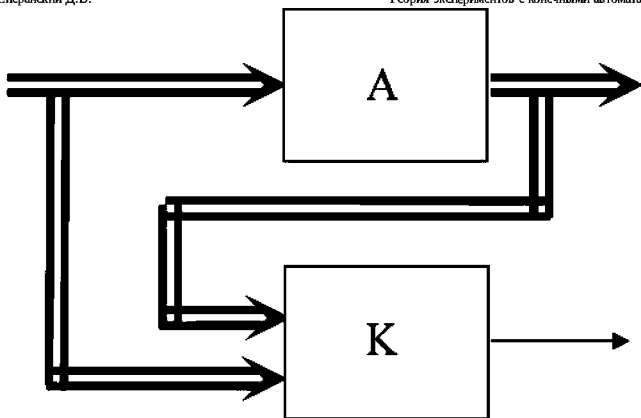


Рис. 6.2.

Автомат K фактически реализует второй и третий этап эксперимента. Вопросы синтеза подобных автоматов рассмотрены в ряде работ, например в [3], [29], поэтому мы не будем здесь на них останавливаться, а основное внимание уделим построению входного слова.

Если автомат задан в виде ориентированного графа, у которого начальной является вершина s_0 , то упомянутому входному слову в графе автомата будет соответствовать путь, начинающийся в s_0 и проходящий через все его дуги. Этот путь будем называть обходом графа. Под путем в графе понимается последовательность дуг, таких, что конечная вершина предыдущей дуги является начальной вершиной следующей. Таким образом, сформулированная нами задача эквивалентна задаче построения обхода автоматного графа.

Под графом $G(S, U)$, где S, U - множество вершин и дуг графа соответственно, мы будем понимать конечный ориентированный мультиграф. Пусть s_0 - начальная вершина этого графа. Дугу u графа, ведущую из вершины s_1 в s_2 , для удобства будем иногда обозначать (s_1, s_2, u) , чтобы явно указать ее начальную и конечную вершины.

Перейдем к определению условий существования обхода графа. Рассмотрим бинарное отношение $\tau \subset S \times S$, задаваемое следующим образом: $(s_1, s_2) \in \tau$ тогда и только тогда, когда в графе $G(S, U)$ существует путь из вершины s_1 в вершину s_2 . Отношение τ называется отношением достижимости. Используя отношение τ , построим бинарное отношение $\eta \subset S \times S$ следующим образом:

$$\eta = \tau \cap \tau^{-1}$$

Известно [62], что η является отношением эквивалентности, а его классы, называемые слоями, являются максимальными сильно связными подмножествами множества S . Каждому слою σ графа $G(S, U)$ естественным образом ставится в соответствие подграф $G'(S, U')$, где U' - множество всех дуг из U , начало и конец которых есть вершины множества σ . Иногда для простоты мы будем отождествлять слой графа с соответствующим ему подграфом, однако каждый раз из контекста будет ясно, о чем именно идет речь.

Следующее утверждение описывает класс графов, у которых существует обход.

Теорема 6.1. Для того чтобы у графа $G(S, U)$ существовал обход, необходимо и достаточно, чтобы выполнялись следующие условия:

1. для любой вершины $s \in S$ существует путь, связывающий вершину s_0 с вершиной s ;
2. в любом слое σ графа существует не более одной вершины s , а в графе $G(S, U)$ не более одной дуги (s, s', u) , таких, что $s' \notin \sigma$.

Необходимость. Предположим, что путь $(s_0, s_{i_1}, u_{i_1})(s_{i_1}, s_{i_2}, u_{i_2}) \dots (s_{i_{k-1}}, s_{i_k}, u_{i_k})$ является обходом графа $G(S, U)$. Справедливость первого условия непосредственно следует из определения обхода графа. Справедливость второго условия докажем от противного. Допустим, что существует слой σ и две различные дуги $(s_{i_\nu}, s_{i_{\nu+1}}, u_{i_{\nu+1}})$ и $(s_{i_\mu}, s_{i_{\mu+1}}, u_{i_{\mu+1}})$, такие, что $s_{i_{\nu+1}} \notin \sigma$ и

$s_{i_{\mu+1}} \notin \sigma$. Предположим, что первая из этих дуг входит в обход ранее второй, т. е. обход имеет следующую структуру:

$$p = p_1(s_{i_{\nu}}, s_{i_{\nu+1}}, u_{i_{\nu+1}})p_2(s_{i_{\mu}}, s_{i_{\mu+1}}, u_{i_{\mu+1}})p_3,$$

где $p_1 (i = 1, 2, 3)$ - некоторые отрезки пути p .

Рассмотрим множество вершин $\sigma Y \{s_{i_{\nu+1}}\}$. Поскольку σ - сильно связанное подмножество, отрезок p_1 заканчивается в вершине $s_{i_{\nu}} \in \sigma$, а отрезок $p_1(s_{i_{\nu}}, s_{i_{\nu+1}}, u_{i_{\nu+1}})$ является начальным отрезком обхода p , то между любой вершиной $s \in \sigma$ и вершиной $s_{i_{\nu+1}} \notin \sigma$ путь по графу $G(S, U)$ существует. Далее, так как отрезок $(s_{i_{\nu}}, s_{i_{\nu+1}}, u_{i_{\nu+1}})p_2$ заканчивается в вершине $s_{i_{\mu}} \in \sigma$ и σ - сильно связанное подмножество, то между вершиной $s_{i_{\nu+1}}$ и любой вершиной $s \in \sigma$ по графу $G(S, U)$ путь также существует. Отсюда следует, что $\sigma Y \{s_{i_{\nu+1}}\}$ есть сильно связанное подмножество. Это противоречит нашему предположению о том, что σ - максимальное сильно связанное подмножество. Полученное противоречие и доказывает наше утверждение.

Достаточность. Поскольку рассматриваемый граф $G(S, U)$ конечен, то конечным является и число его слоев. Предположим, что это число есть N .

Пронумеруем слои σ графа $G(S, U)$ следующим образом:

- σ_1 - слой, содержащий начальную вершину s_0 ;
- пусть $i - 1$ слоев ($i = 2, \dots, N - 1$) уже пронумеровано; слой σ получает номер i , если существует такая вершина $s \in \sigma_{i-1}$ (σ_{i-1} - слой с номером $i - 1$) и такая дуга (s, s', u) , что $s' \in \sigma$;
- номер N получает тот слой σ , все дуги которого ведут в вершины, принадлежащие σ .

Легко показать, что каждому слою графа $G(S, U)$, удовлетворяющему условию доказываемой теоремы, в соответствии со сформулированной процедурой будет присвоен единственный номер из диапазона от 1 до N .

Обозначим через $s^{(i)}$ вершину слоя σ_i , из которой выходит дуга $(s^{(i)}, \tilde{s}^{(i)}, u^{(i)})$, такая, что $\tilde{s}^{(i)} \in \sigma_{i+1}$. В силу второго условия теоремы эта дуга единственна. Рассмотрим подграф $G(\sigma_i, u_i)$ графа $G(S, U)$, где σ_i - множество вершин i -го слоя, u_i - множество всех дуг, начало и конец которых есть вершины множества σ_i . Поскольку этот подграф сильно связан, то у него всегда существует обход p_i , начинающийся в вершине $\tilde{s}^{(i-1)}$ и заканчивающийся в вершине $s^{(i)}$ ($i = 1, 2, \dots, N-1$). Через p_N обозначим обход подграфа $G(\sigma_N, U_N)$, начинающийся в вершине $\tilde{s}^{(N-1)}$. Условимся считать, что $\tilde{s}^{(0)} = s^0$.

Легко видеть, что путь $p = p_1(s^{(1)}, \tilde{s}^{(1)}, u^{(1)})p_2p_1(s^{(2)}, \tilde{s}^{(2)}, u^{(2)}) \dots p_1(s^{(N-1)}, \tilde{s}^{(N-1)}, u^{(N-1)})p_N$ является обходом графа $G(S, U)$. Достаточность условий теоремы доказана.

Хотя приведенное доказательство теоремы 6.1 и не является конструктивным, мы не будем здесь останавливаться на методах построения обходов графа в связи с тем, что в теории графов известно много различных методов построения путей между двумя вершинами (в том числе и минимальной длины), которыми можно воспользоваться при конструировании обходов. Покажем теперь, каким образом задача построения обхода графа может быть сведена к задаче построения пути, соединяющего две его выделенные вершины.

По графу $G(S, U)$ с начальной вершиной s_0 построим новый граф $O(B(U) \times S, U)$, где $B(U)$ - множество всех подмножеств U . Начальной вершиной графа $O(B(U) \times S, U)$ будем считать вершину (U, s_0) . Положим, что в этом графе из вершины (U_{i_1}, s_{i_1}) в вершину

(U_{i_2}, s_{i_2}) ведет дуга $((U, s_{i_1}), (U_{i_2}, s_{i_2}), u)$, если:

- дуга (s_{i_1}, S_{i_2}, u) графа $G(S, U)$ принадлежит U_{i_1} ;
- $U_{i_2} = U_{i_1} \setminus (s_{i_1}, s_{i_2}, u)$, где знаком \setminus обозначена теоретико-множественная разность.

Из самой конструкции графа $O(B(U) \times S, U)$ легко следует справедливость следующего утверждения.

Лемма 6.1. Путь $p = (s_0, s_{i_1}, u_{i_1})(s_{i_1}, s_{i_2}, u_{i_2}) \dots (s_{i_{k-1}}, s_{i_k}, s_{i_k})$ в графе $G(S, U)$ является его обходом тогда и только тогда, когда в графе $O(B(U) \times S, U)$ существует путь

$$((u, s_0), (u_{i_1}, s'_{i_1}), u'_{i_1}) \times ((u_{i_1}, s'_{i_1}, (u_{i_2}, s'_{i_2}), y'_{i_2}) \dots ((u_{i_{k-1}}, s'_{i_{k-1}}), (u_{i_k}, s'_{i_k}), u'_{i_k})),$$

такой, что

1. $u_{i_k} = \emptyset$;
2. для любого $\nu, 1 \leq \nu \leq k, s_{i_\nu} = s'_{i_\nu}, u_{i_\nu} = u'_{i_\nu}$.

Лемма 6.1 по существу дает алгоритмы для решения следующих задач:

1. Выяснить, существует ли обход для графа $G(S, U)$ с начальной вершиной s_0 .
2. Найти множество всех обходов графа $G(S, U)$.
3. Найти множество всех минимальных по длине обходов графа $G(S, U)$.

Эти алгоритмы сводятся к построению по графу $G(S, U)$ графа $O(B(U) \times S, U)$ и последующему его анализу. Так, первая задача равносильна выяснению существования пути между вершиной (u, s_0)

и любой вершиной вида (\emptyset, s) этого графа, вторая - нахождению всех путей графа $O(B(U) \times S, U)$ между упомянутыми вершинами и, наконец, третья - определению всех путей минимальной длины между теми же вершинами.

Перейдем к интерпретации полученных результатов для задачи, сформулированной в начале этой лекции. Сделаем это достаточно подробно, чтобы далее на этом детально не останавливаться.

Легко видеть, что при построении простого безусловного эксперимента для распознавания функции выходов инициального автомата нам достаточно ограничиться рассмотрением автомата без выходов $A = (S, X, \delta, s_0)$, а не исходного автомата Мили $A = (S, X, Y, \delta, \lambda, s_0)$, ему соответствующего.

Определение 6.1. Слово $p \in X_e^*$ будем называть характеристическим для автомата $A = (S, X, \delta, s_0)$, если для любого состояния $s \in S$ и любого входного символа $x \in X$ существует начальный отрезок q этого слова, такой, что $p = qxq'$ и $\delta(s_0, q) = s$.

Из этого определения следует, что характеристическое слово автомата в точности соответствует пути по графу этого автомата, начинающемуся в вершине s_0 и проходящему через все его дуги, т. е. обходу автоматного графа. Таким образом, исходная задача сведена нами к задаче построения характеристического слова для автомата $A = (S, X, \delta, s_0)$. Простая переформулировка теоремы 6.1 дает возможность описать класс автоматов, для которых характеристическое слово существует.

Теорема 6.1. Для того чтобы конечный инициальный автомат $A = (S, X, \delta, s_0)$ обладал характеристическим словом, необходимо и достаточно, чтобы выполнялись следующие условия:

1. для любого состояния $s \in S$ существует слово $p \in X_e^*$, такое, что $\delta(s_0, p) = s$;
2. для любого слоя σ существует не более одного состояния $s \in \sigma$ и не более одного входного символа $x \in X$, таких, что

$$\delta(s, x) \notin \sigma.$$

Оценки длины кратчайших обходов и характеристических слов

Граф $G(S, U)$ с начальной вершиной s_0 , удовлетворяющий условиям теоремы 6.1, будем называть s_0 - правильным или просто правильным.

Через p_{\min}^G обозначим кратчайший обход графа $G(S, U)$, а через $d(p_{\min}^G)$ - его длину, т. е. число дуг, входящих в обход.

По определению обход графа $G(S, U)$ представляет собой путь, содержащий все дуги из U . Отсюда следует нижняя оценка длины кратчайшего обхода:

$$d(p_{\min}^G) \geq |U| \quad (6.1)$$

Возникает вопрос, существуют ли графы, для которых эта оценка достижима, и если да, то как описать такой класс графов. Ответы на эти вопросы даются следующей теоремой.

Теорема 6.2. Для правильного графа $G(S, U)$ обход длины $|U|$ существует тогда и только тогда, когда выполняется одно из следующих условий:

1. в каждой вершине графа число исходящих дуг равно числу заходящих;
2. в начальной вершине s_0 графа число исходящих дуг на единицу больше числа заходящих, и существует такая вершина s , в которой число исходящих дуг на единицу меньше числа заходящих; в остальных вершинах графа число исходящих дуг равно числу заходящих.

Справедливость этой теоремы следует из известных результатов теории графов [25]. Очевидно, что путь (контур) в графе длины $|U|$, проходящий через все его дуги и только по одному разу, есть не что

иное, как эйлеров путь (контур). Условия сформулированной теоремы являются соответственно условиями существования эйлерова контура и пути.

Далее обход графа $G(S, U)$, начинающийся в вершине a и заканчивающийся в одной из вершин множества B , будем именовать (a, B) -обходом.

(s_0, S) -обход, т. е. обход в определенном ранее смысле, будем называть s_0 -обходом, чтобы отметить его начальную вершину. Если $B = \{b\}$, то (a, B) -обход будем называть (a, b) -обходом.

Для установления верхней оценки длины кратчайших обходов введем несколько определений и докажем ряд вспомогательных утверждений.

Через $\Delta(s)$ обозначим разность между числом заходящих и исходящих дуг вершины s графа $G(S, U)$.

Определение 6.2. Вершину s графа $G(S, U)$, у которой $\Delta(s) > 0$ ($\Delta(s) < 0$), назовем положительной (отрицательной).

Через S^+ (S^-) обозначим множество всех положительных (отрицательных) вершин графа $G(S, U)$.

Определение 6.3. Семейство M элементарных путей назовем компенсирующей системой путей графа G для (a, b) -обхода, если:

- каждая положительная (отрицательная) вершина s , отличная от a и b , является началом $\Delta(s)$ (концом $|\Delta(s)|$) путей из M , где $|K|$ - модуль K ;
- при $\Delta(a) \geq 0$ ($\Delta(a) < 0$) вершина a является началом $\Delta(a) + \omega$ (концом $|\Delta(a) + \omega|$) путей из M , где $\omega = 0$, если $a = b$, и $\omega = 1$, если $a \neq b$;

3. если $\Delta(b) > 0$ ($\Delta(b) \leq 0$), то вершина b является началом $\Delta(b) - \omega$ (концом $|\Delta(b) - \omega|$) путей из M , где ω определяется так же, как в пункте 2.

Сумму длины всех путей из M назовем длиной компенсирующей системы M и обозначим ее через $D(M)$.

Через $(D(a, b))$ обозначим длину кратчайшего (a, b) -обхода графа G . Легко убедиться, что для правильных графов компенсирующая система всегда существует.

Теорема 6.3. $D(a, b) = |U| + D(M)$, где M - компенсирующая система минимальной длины для (a, b) -обхода графа $G(S, U)$.

Доказательство. Каждому пути из M , ведущему из вершины s в вершину s' , поставим во взаимно однозначное соответствие дугу, соединяющую те же вершины. Пусть множество этих дуг есть $V = \{v_1, \dots, v_{|M|}\}$. Рассмотрим граф $G(S, UYV)$, полученный из $G(S, U)$ добавлением всех дуг множества V . Легко видеть, что в графе $G(S, UYV)$ существует эйлеров путь P , ведущий из вершины a в вершину b (см. теорему параграфа 28 из [25]). Заменяя в P каждую дугу v_i соответствующим ему путем из M , получим новый путь p' , проходящий только по дугам графа $G(S, U)$ и являющийся его (a, b) -обходом. Из сказанного следует, что $D(a, b) \leq |U| + D(M)$. Неравенство $D(a, b) \geq |U| + D(M)$ вытекает из того, что M - компенсирующая система минимальной длины. Из приведенных неравенств следует справедливость утверждения теоремы.

Будем говорить, что граф $G(S, U)$ имеет степень m , если из каждой его вершины исходит ровно m дуг.

Теорема 6.4. Для сильно связного графа $G(S, U)$, $|S| = n$ степени m и диаметра d имеет место неравенство

$$D(a, b) \leq mn + \frac{1}{2}(m-1)(2n-d-1)d + d \quad (6.2)$$

Доказательство. Поскольку при $m = 1$ теорема очевидна, то рассмотрим случай $m > 1$. Число дуг графа G , т. е. $|U|$, равно mn , но тогда в силу теоремы 6.3 достаточно доказать, что существует компенсирующая система путей M графа G для (a, b) -обхода, длина которой не превосходит величины $\frac{1}{2}(m-1)(2n-d-1)d + d$.

Пусть $s_{i_1}, s_{i_2}, \dots, s_{i_t} (s'_{i_1}, s'_{i_2}, \dots, s'_{i_t})$ - неупорядоченная последовательность вершин графа $G(S, U)$, таких, которые в соответствии с определением 6.3 являются начальными (конечными) вершинами путей компенсирующей системы M графа G . Факт существования системы M для сильно связного графа очевиден.

Если некоторая вершина является началом (концом) l путей системы M , где $l > 1$, то в последовательности $s_{i_1}, s_{i_2}, \dots, s_{i_t} (s'_{i_1}, s'_{i_2}, \dots, s'_{i_t})$ она встречается l раз.

Построение компенсирующей системы путей M будем выполнять в виде пошагового процесса.

1-й шаг. Находим кратчайший путь μ_1 среди путей, ведущих из вершины s_{i_j} в вершину s'_{i_k} , $1 \leq j, k \leq t$. Путь μ_1 является первым путем компенсирующей системы M . Предположим, что μ_1 начинается в вершине s и заканчивается в вершине s' . Вычеркиваем s и s' из соответствующих последовательностей вершин, выписанных выше.

$(\alpha + 1)$ -й шаг. Пусть уже построены пути $\mu_1, \mu_2, \dots, \mu_\alpha$ компенсирующей системы M , где $\alpha \geq 1$. Тогда последовательность начальных (конечных) вершин путей системы M содержит на данном шаге $t - \alpha$ членов. Не теряя общности, можно считать, что такая последовательность начальных (конечных) вершин суть $s_{i_1}, s_{i_2}, \dots, s_{i_{t-\alpha}} (s'_{i_1}, s'_{i_2}, \dots, s'_{i_{t-\alpha}})$. Находим кратчайший путь $\mu_{\alpha+1}$ среди путей, ведущих из вершины s_{i_j} в вершину

$s'_{i_k}, 1 \leq j, k \leq t - \alpha$. Пусть $\mu_{\alpha+1}$ является $(\alpha + 1)$ -м путем компенсирующей системы M .

Построение компенсирующей системы путей M заканчивается, когда из последовательности ее начальных (конечных) вершин будут вычеркнуты все элементы. Предположим, что по описанному алгоритму построена компенсирующая система путей M . Покажем, что при любом $k, 1 \leq k \leq d$, система эта содержит r путей длины, не меньшей r , где $r \leq (n - k)(m - 1) + 1$.

Допустим противное. Пусть $\mu_{l_1}, \dots, \mu_{l_r}$ - пути системы M длины, не меньшей k , где $r > (n - k)(m - 1) + 1$. Поскольку G сильно

связан, то каждая вершина из множества $S^- \setminus \{b\}$ является концом не более чем для $m - 1$ путей из M . Очевидно, что вершина b может являться концом не более чем для m путей системы M . Из этого следует, что среди вершин, являющихся конечными для путей $\mu_{l_1}, \dots, \mu_{l_r}$ из M длины, не меньшей k , найдется по крайней мере $n - k + 1$ различных. Тогда, если s является начальной вершиной некоторого пути $\mu_{i_j}, 1 \leq j \leq r$, среди конечных вершин путей $\mu_{j_1}, \dots, \mu_{j_r}$ найдется вершина s' , такая, что длина кратчайшего пути из s в s' не превышает $k - 1$. Это противоречит нашему предположению.

Обозначим через β_j число путей длины j в компенсирующей системе путей M графа $G(S, U)$ для (a, b) -обхода. Очевидно, что общая длина всех путей из M определяется следующим образом:

$$D(M) = \sum_{j=1}^d j * \beta_j \leq \sum_{k=1}^d [(n - k)(m - 1) + 1] = \frac{1}{2}(m - 1)(2n - d - 1)d + d$$

Покажем, что оценка (6.2) является достижимой. Рассмотрим граф на рис.6.3.

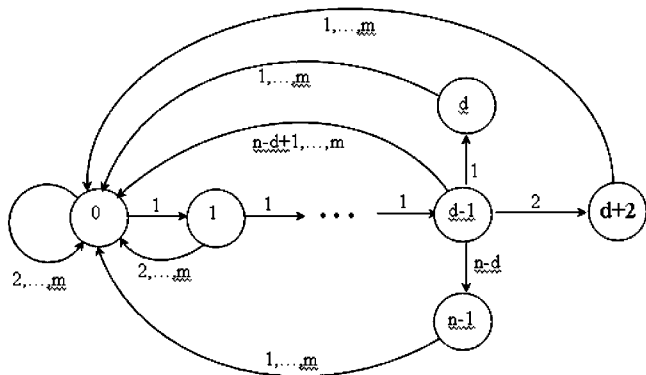


Рис. 6.3.

Из каждой вершины графа выходит m дуг с отметками $1, 2, \dots, m$. Для простоты на этом рисунке кратные дуги заменены одной дугой с кратной отметкой. Легко видеть, что для всех $n, d (1 \leq d \leq n - 1), m (m \geq n - d + 1)$ и для всех $b \in \{d, \dots, n - 1\}$ длина кратчайшего $(0, b)$ -обхода графа G точно равна оценке (6.2).

Теорема 6.5. Пусть $G(S, U), |S| = n$, есть сильно связный граф степени $m > 1$, у которого удалена одна дуга, исходящая из вершины b . Если после удаления этой дуги граф G остался сильно связным и диаметр его равен d , то для любой вершины a этого графа справедливо неравенство

$$D(a, b) \leq mn + \frac{1}{2}(m - 1)(2n - d - 1)d - 1 \quad (6.3)$$

Доказательство. При $m = 1$ сильно связный граф G представляет собой следующую конструкцию: из i -й вершины дуга ведет в $(i + 1)$ -

ю для $i = 1, 2, \dots, n - 1$, а из n -й вершины в первую. Очевидно, что удаление любой дуги из G превращает его в граф, не являющийся сильно связным. Поэтому мы покажем справедливость оценки (6.3) при $m > 1$.

Пусть G' - сильно связный граф, полученный из G удалением одной дуги, исходящей из вершины b . Добавим к этому графу петлю в ту же вершину. Полученный в результате новый граф обозначим через G^* .

Построим компенсирующую систему путей M для (a, b) -обхода графа G^* по алгоритму, изложенному при доказательстве теоремы 6.4. Рассуждения, аналогичные приведенным при доказательстве предыдущей теоремы, показывают, что компенсирующая система путей

M для (a, b) -обхода графа G^* при любом $k (1 \leq k \leq d)$ содержит r путей длины, не меньшей k , причем $r \leq (n - k)(m - 1)$. Отсюда следует, что общая длина $D(M)$ путей компенсирующей системы M графа G^* удовлетворяет неравенству

$$D(M) \leq \sum_{k=1}^d [(n - k)(m - 1)] \frac{1}{2} (m - 1)(2n - d - 1)d =$$

Тогда длина (a, b) -обхода графа G^* не превосходит величины $mn + \frac{1}{2}(m - 1)(2n - d - 1)d$. Нетрудно убедиться, что среди (a, b) -обходов графа G^* , удовлетворяющих оценке (6.3), есть такой, у которого добавленная фиктивная петля является последней в этом обходе, т. е. обход имеет следующую структуру: $\mu = \mu' u$, где u - дуга фиктивной петли в вершине b . Отсюда следует, что путь μ' является (a, b) -обходом графа G' , а длина его не превосходит величины $mn + \frac{1}{2}(m - 1)(2n - d - 1)d - 1$.

Покажем, что оценка (6.3) является достижимой. С этой целью вновь обратимся к графу, изображенному на рис.6.3. Непосредственным подсчетом легко убедиться, что для всех $n, d (1 \leq d \leq +1), m (m \geq n - d + 1)$ и $b \in \{d, \dots, n - 1\}$, если удалить одну дугу (произвольным образом) рассматриваемого графа,

длина $(0, b)$ -обхода в точности равна оценке (6.3).

Рассмотрим теперь обход графов, не являющихся сильно связными.

Пусть $G(S, U)$ - s_0 правильный граф, число слоев которого есть N . Обозначим через $G_i(S_i, U_i)$ подграф графа G , соответствующий его i -му слою. Легко видеть, что (a, B) -обход графа G существует тогда и только тогда, когда $BIS_N \neq \emptyset$ и G есть a -правильный граф.

Пусть $G(S, U), |S| = n$, есть граф степени m , являющийся a_1 -правильным. Предположим, что $G_i(S_i, U_i)$ - слои графа $G, i = 1, \dots, N$, которые занумерованы в соответствии с процедурой нумерации, описанной в разделе 6.1. Условимся считать, что $|S_i| = n_i$, а диаметр G_i есть d_i . Пусть u_i - дуга, ведущая из вершины $b_i \in S_i$ в вершину $a_{i+1} \in S_{i+1}$. Через $O_G(a, b)$ обозначим (a, b) -обход графа G . Очевидно, что $O_G(a_1, b_N) = O_{G_1}(a_1, b_1)u_1O_{G_2}(a_2, b_2) \dots u_{k-1}O_{G_N}(a_n, b_N)$. Из этого, а также теорем 6.3 и 6.5 вытекает справедливость следующего утверждения.

Теорема 6.6.

$$D(a_1, b_N) = |U| + \sum_{i=1}^N D(M_i)$$

где M_i - компенсирующая система путей минимальной длины графа G_i для (a_i, b_i) -обхода.

Следствие 1. Длина кратчайшего (a_1, a_N) -обхода (следовательно, и a_1 -обхода) графа G не превосходит величины

$$mn + \frac{1}{2}(m-1) \sum_{i=1}^N (2n_i - d_i - 1)d_i \quad (6.4)$$

Доказательство. В силу теоремы 6.4 достаточно доказать, что для каждого из графов G_i , у которого удалена дуга, ведущая из вершины b_i

в вершину a_{i+1} графа G_{i+1} ($i = 1, 2, \dots, N - 1$), существует компенсирующая система путей M_i для (a_1, b_i) -обхода, а для графа G_N - система M_N для (a_N, a_N) -обхода, для которых справедливо неравенство

$$\sum_{i=1}^N D(M_i) \leq \frac{1}{2}(m-1) \sum_{i=1}^N (2n_i - d_i - 1)d_i$$

Методом, аналогичным использованному в доказательстве теоремы 6.5, можно показать, что длина компенсирующей системы путей графа G_N для (a_N, a_N) -обхода не превышает величины

$$\frac{1}{2}(m-1)(2n_N - d_N - 1)d_N \quad (6.5)$$

Длина компенсирующей системы путей M_i для (a_i, b_i) -обхода графа G_I ($i = 1, 2, \dots, N - 1$) с одной удаленной дугой, ведущей из b_i в $a_{i+1} \in S_{i+1}$, в силу оценки (3.3) удовлетворяет неравенству

$$D(M_i) \leq \frac{1}{2}(m-1)(2n_i - d_i - 1)d_i - 1$$

Отсюда следует, что

$$D(a_1, a_N) \leq mn + \frac{1}{2} \sum_{i=1}^{N-1} [(m-1)(2n_i - d_i - 1)d_i - 1] + \\ + \frac{1}{2}(m-1)(2n_N - d_N - 1)d_N + (N-1) = mn + \frac{1}{2}(m-1) \sum_{i=1}^N (2n_i - d_i - 1)d_i$$

Замечание. На основании (6.5) получаем, что длина (a, a) -обхода n -вершинного сильно связанного графа степени m удовлетворяет неравенству

$$D(a, a) \leq mn + \frac{1}{2}(m-1)(2n - d - 1)d \quad (6.6)$$

Известно, что диаметр n -вершинного графа не превосходит величины $n - 1$. Заменяя в (3.6) d на $n - 1$, получаем

$$D(a, a) \leq \frac{1}{2}(m-1)(n-1)n \quad (6.7)$$

Таким образом, оценка (6.7), установленная в работе [36], следует из оценки (6.6).

Следствие 2. Если в графе $G(S, U)$, $|S| = n$, степени $m(a, b)$ -обход существует, то длина кратчайшего (a, b) -обхода не превосходит величины

$$mn + \frac{1}{2}(m-1) \sum_{i=1}^N (2n_i - d_i - 1)d_i + d_N \quad (6.8)$$

Это следствие непосредственно вытекает из теорем 6.5 и 6.6.

Легко показать, что оценки (6.4) и (6.8) достижимы. Для этого достаточно построить граф H , слой H_i которого есть подграф, изображенный на рис.6.3. В каждом из этих подграфов произведем следующие изменения: дугу i -го подграфа ($1 \leq i \leq N-1$), исходящую из вершины с номером $n_i - 1$ с отметкой m , заменим дугой с той же отметкой и исходящей из той же вершины, но заходящей в вершину с номером 0 ($i+1$)-го подграфа. Непосредственным подсчетом можно убедиться, что оценки (6.4) и (6.8) для графа H достижимы.

Из полученных результатов следует, что точная нижняя оценка длины кратчайших характеристических слов автомата есть оценка (6.1), а точные верхние оценки следующие: для сильно связанного автомата - (6.2) и (6.7), для остальных - (6.8).

Вопросы и упражнения

1. Приведите постановку задачи контроля функции выходов инициального автомата.
2. Приведите критерий существования обхода графа и его доказательство.
3. Дайте определение характеристического слова автомата.
4. Каков критерий существования характеристического слова автомата?
5. Каковы условия существования обхода графа минимально

возможной длины?

6. Что представляет собой компенсирующая система путей графа ?
7. Приведите верхние оценки различных типов обходов. Какие из этих оценок являются достижимыми?

Контроль функции выходов неинициального автомата с использованием простого безусловного эксперимента

Исследуется задача контроля функции выходов автомата в случае, когда неизвестно его начальное состояние. Приведен критерий разрешимости задачи. Показывается, что эта задача сводится к классической задаче распознавания автомата из известного класса. Описывается метод ее решения и приводится верхняя оценка длины контролирующего эксперимента.

Рассмотрим следующую задачу. Пусть в распоряжении экспериментатора имеется один экземпляр автомата Мили A , у которого известны входной алфавит $X = \{x_1, \dots, x_m\}$, выходной алфавит $Y = \{y_1, \dots, y_l\}$, множество состояний $S = \{s_1, \dots, s_n\}$, функция переходов δ . Требуется построить простой безусловный эксперимент, позволяющий распознать функцию выходов λ автомата A или эквивалентного ему автомата. Эта задача отличается от задачи, сформулированной в лекции 6, лишь тем, что рассматриваемый здесь автомат предполагается неинициальным. Вследствие тех же причин, которые были приведены в лекции 6, содержательный смысл рассматриваемой задачи, в частности, соответствует и контролю функции выходов неинициального автомата.

Следующая теорема описывает класс автоматов, для которого сформулированная задача разрешима.

Теорема 7.1. Для того чтобы задача распознавания функции выходов неинициального автомата (с точностью до эквивалентности) была разрешима, необходимо и достаточно, чтобы автомат был сильно связан.

Необходимость условия теоремы является очевидной, поэтому остановимся лишь на доказательстве его достаточности.

Достаточность. Построим всевозможные таблицы размерности $m \times n$, состоящие из элементов множества Y . Число таких таблиц конечно и равно $N = l^{m \cdot n}$. Очевидно, что класс автоматов $K = \{A_1, \dots, A_N\}$, у которых функция переходов совпадает с δ , содержит автомат A .

Среди автоматов класса K встречаются попарно эквивалентные. Из каждого множества всех эквивалентных автоматов произвольным образом выбираем по одному представителю и образуем из них новый класс автоматов $K' = \{A'_1, \dots, A'_r\}$, где $r < N$. По построению в классе K' отсутствуют попарно эквивалентные автоматы, а один из автоматов $A'_i \in K'$ совпадает с автоматом A или ему эквивалентен.

По предположению каждый $A_i \in K$ является сильно связным, но тогда можно показать, что построенный нами класс автоматов K' является исключительным. Известно [18], что исключительность класса является необходимым и достаточным условием распознаваемости автомата, принадлежащего этому классу, простым безусловным экспериментом. Таким образом, исходная задача сведена к задаче распознавания автомата, принадлежащего известному классу, которая в нашем случае оказывается разрешимой. Этим и заканчивается доказательство теоремы.

Из приведенного доказательства следует применимость метода (предложенного в [18]) распознавания автомата, принадлежащего известному классу, к решению рассматриваемой нами задачи. Однако в этом случае он обладает тем недостатком, что при больших l, m, n число автоматов в классе K' велико, что вызывает значительное увеличение длины соответствующего эксперимента.

В связи со сказанным мы рассмотрим иной способ проведения эксперимента, позволяющий сократить мощность рассматриваемого класса автоматов и упростить сам эксперимент. Предположим, что исследуемый неинициальный автомат A находится в состоянии s_1 . Построим для автомата $A_1 = (S, X, \delta, s_1)$ характеристическое слово, которое обозначим через p_1 , и подадим его на вход исследуемого автомата.

По наблюдаемой реакции q_1 решаем задачу распознавания функции выходов λ_1 автомата Мили, которому соответствует автомат без выходов $A_1 = (S, X, \delta, s_1)$. Вообще говоря, эта задача может и не иметь решения из-за неоднозначности, возникающей в процессе построения λ_1 . Неоднозначность этой функции означает, что s_1

заведомо не является истинным начальным состоянием исследуемого автомата A . Если же задача разрешима, то обозначим через T_1 таблицу, которая соответствует распознанной функции выходов λ_1 .

Предположим далее, что начальным состоянием исследуемого автомата перед подачей слова p_1 было s_2 . Рассмотрим автомат $A_2 = (S, X, \delta, s_2)$ и найдем такое слово p_2 , чтобы $p_1 p_2$ являлось для A_2 характеристическим словом. В частности, в качестве p_2 можно взять характеристическое слово для автомата $A'_2 = (S, X, \delta, \delta(s_2, p_1))$. Подаем p_2 на вход исследуемого автомата, и пусть наблюдаемая его реакция есть q_2 . По реакции $q_1 q_2$ решаем задачу распознавания функции выходов λ_2 автомата Милн, которому соответствует автомат без выходов A_2 . Если эта задача разрешима, то обозначим через T_2 таблицу, которая соответствует распознанной функции выходов λ_2 .

Продолжая этот процесс и далее аналогичным образом в предположении, что начальным состоянием исследуемого автомата перед подачей слова p_1 были состояния s_3, \dots, s_n соответственно, строим, если задачи оказываются разрешимыми, таблицы T'_3, \dots, T'_n , представляющие собой распознанные функции выходов $\lambda_3, \dots, \lambda_n$.

Предположим, что в результате описанного процесса построены таблицы $T'_{i_1}, \dots, T'_{i_a}$, где $a \leq n$. Рассмотрим класс автоматов $L = \{A_{i_1}, \dots, A_{i_a}\}$, функции переходов которых совпадают с функцией переходов δ исследуемого автомата, а функции выходов заданы упомянутыми таблицами. Легко видеть, что по крайней мере один из автоматов класса L совпадает с исследуемым автоматом или ему эквивалентен. Может оказаться, что в классе L есть попарно эквивалентные автоматы. В этом случае из каждого множества всех эквивалентных автоматов выберем по одному представителю с наибольшим номером и образуем из них новый класс $L' = \{A_{j_1}, \dots, A_{j_b}\}, b \leq a, j_1 \leq j_2 \leq \dots \leq j_b$.

Поскольку все автоматы из L' являются сильно связными, а по построению - попарно эквивалентными, то из уже приводимого выше

утверждения [18] следует, что класс $L' = \{A_{j_1}, \dots, A_{j_b}\}$ исключительный. Отметим, что при больших l, m, n мощность L' меньше мощности класса K' , рассматриваемого в методе из [18].

Предположим, что автомат $A_{j_k} \in L'$ был выбран из класса L в качестве представителя множества эквивалентных автоматов $\{A_{\mu_1^{(k)}}, \dots, A_{\mu_k^{(k)}}\}$, где $\mu_1^{(k)} \leq \dots \leq \mu_k^{(k)}$ и $\mu_k^{(k)} = j_k$. Пусть $s_{\mu_1^{(k)}}, \dots, s_{\mu_k^{(k)}}$ - состояния исследуемого автомата A , в которых он находился по предположению при построении таблиц $T_{\mu_1^{(k)}}, \dots, T_{\mu_k^{(k)}}$ соответственно. Образует множество состояний

$$S_{i_k} = \{s'_{\mu_1^{(k)}}, \dots, s'_{\mu_k^{(k)}}\}$$

где $s'_{\mu_1^{(k)}} = s_{\mu_1^{(k)}}$, а состояние $s'_{\mu_i^{(k)}}$, $i = 1, 2, \dots, k$, есть состояние автомата $A_{\mu_i^{(k)}}$, эквивалентное $s_{\mu_1^{(k)}}$. Если в $A_{\mu_1^{(k)}}$ существует несколько состояний, эквивалентных $s_{\mu_1^{(k)}}$, то в качестве $s'_{\mu_i^{(k)}}$, произвольным образом выбираем одно из них. Построим теперь множество состояний

$$\tilde{S}_{j_k} = \{\delta(s'_{\mu_1^{(k)}} \cdot p), \dots, \delta(s'_{\mu_k^{(k)}} \cdot p)\}$$

где $p = p_1 p_2 \dots p_n$ - слово, использованное на предыдущем этапе эксперимента для построения автоматов класса L . Образует, наконец, множество состояний $\tilde{S} = Y_{1 \leq k \leq b} \tilde{S}_{j_k}$.

Рассмотрим теперь автомат $\Delta(A_{j_k}, \dots, A_{j_b})$, состоящий из изолированных подавтоматов A_{i_k} и по терминологии [18] называемый расщепляемым. Из приведенных выше построений вытекает, что исследуемый автомат совпадает с одним из подавтоматов A_{j_k} , находящимся в данный момент в одном из состояний множества \tilde{S}_{j_k} , $= 1, \dots, b$, либо ему эквивалентен. Для завершения

эксперимента по распознаванию исследуемого автомата достаточно для расщепляемого автомата $\Delta(A_{j_1}, \dots, A_{j_b})$ с множеством допустимых начальных состояний \tilde{S} построить такое слово p' , реакция на которое позволяет определить, к какому из изолированных подавтоматов принадлежит конечное состояние. Очевидно, что этот подавтомат и является искомым.

Эта задача, как нетрудно видеть, является обобщенной задачей установки, рассмотренной в [63]. Напомним, что слово $p' \in X^*$ называется "обобщенным" установочным словом для автомата A и пары (S', ε) , если

$$\forall_{s_1, s_2 \in S'} \lambda(s_1, p') = \lambda(s_2, p') \rightarrow (\delta(s_1, p'), \delta(s_2, p')) \in \varepsilon$$

где $S' \subset S$ - некоторое выделенное подмножество состояний автомата A , интерпретируемое как множество допустимых начальных состояний A , а ε - некоторое отношение эквивалентности на множестве $S \times S$, классы которого определяют "точность" решения задачи установки. Если ε есть тождественное отношение, то "обобщенная" задача установки вырождается в обычную задачу установки.

Применительно к нашему случаю рассмотрим "обобщенную" установочную задачу, когда $S' = \tilde{S}$, а классами отношения ε являются множества состояний изолированных подавтоматов A_{j_k} ($k = 1, \dots, b$).

В [63] показано, что условие

$$\nu \subset \varepsilon \quad (7.1)$$

является достаточным для существования "обобщенного" установочного слова для автомата $A = (S, X, Y, \delta, \lambda)$, где бинарное отношение $\nu \subset S \times S$ определяется следующим образом:

$$(s_1, s_2) \in \nu \leftrightarrow \forall_{p \in X^*} \lambda(s_1, p) = \lambda(s_2, p)$$

Выше показано, что класс автоматов $L' = \{A_{j_1}, \dots, A_{j_b}\}$ является исключительным. Из этого следует, что в рассматриваемом нами случае условие (7.1) выполняется. Таким образом, знания реакции исследуемого автомата на "обобщенное" установочное слово p' достаточно для определения того изолированного подавтомата A_{j_k} расщепляемого автомата $\Delta(A_{j_k}, \dots, A_{j_b})$, которому принадлежит конечное состояние.

Итак, нами построен и обоснован алгоритм распознавания функции выходов неинициального автомата, который является более эффективным с точки зрения уменьшения объема вычислений, чем алгоритм Гилла [18].

Перейдем теперь к установлению верхней оценки длины описанного эксперимента, включающего в себя подачу на вход исследуемого автомата слова $p = p_1 p_2 \dots p_n$ и "обобщенного" установочного слова p' .

Поскольку слово p' является по построению характеристическим для автомата A , то из (6.7) следует, что длина его не превосходит величины $n + \frac{1}{2}(m-1)n(n+1)$, где $n = |S|$, $m = |X|$. Длина слов p_2, \dots, p_n тем более не превосходит этой величины. Тогда длина слова p не превосходит величины $n[n + \frac{1}{2}(m-1)n(n+1)]$.

Слово p' является "обобщенным" установочным словом для автомата $\Delta(A_{j_1}, \dots, A_{j_b})$, число состояний которого не больше n^2 , а мощность множества допустимых начальных состояний не превышает n . Очевидно, что длина слова p' не превышает длины установочного слова для того же автомата, понимаемого в обычном смысле. Тогда в силу результатов, полученных в [67], верхняя оценка длины p' есть $\frac{1}{2}n(2n-1)(n-1)$. Как следствие, из наших рассуждений вытекает следующее утверждение.

Теорема 7.2. Длина кратчайшего простого безусловного эксперимента,

позволяющего распознавать функцию выходов сильно связанного неинициального автомата $A = (S, X, Y, \delta, \lambda)$, где $|S| = n$, $|X| = m$, не превышает величины

$$n^2 + \frac{1}{2}(m-1)n^2(n+1) + \frac{1}{2}n(2n-1)(n-1)$$

Проиллюстрируем построение эксперимента на примере. Рассмотрим автомат A , функция переходов которого задана в табл. 7.1. Пусть его истинное начальное состояние (неизвестное экспериментатору) есть s_2 . В предположении, что начальным состоянием автомата A является s_1 , строим характеристическое слово $p_1 = 1010100$.

Пусть реакция автомата A на p_1 есть слово 1001100. Предполагаемые переходы и наблюдаемая реакция $(s_1 \xrightarrow{1/1} s_1 \xrightarrow{0/0} s_2 \xrightarrow{1/0} s_3 \xrightarrow{0/1} s_3 \xrightarrow{1/1} s_1 \xrightarrow{0/0} s_2 \xrightarrow{0/0} s_1)$ позволяют построить таблицу выходов T_1 . Истинные переходы автомата A при подаче слова p_1 таковы:

$$s_2 \xrightarrow{1} s_3 \xrightarrow{0} s_3 \xrightarrow{1} s_1 \xrightarrow{0} s_2 \xrightarrow{1} s_3 \xrightarrow{0} s_3 \xrightarrow{0} s_3$$

Далее, в предположении, что начальным состоянием A было s_2 , строим слово $p_2 = 1100$, позволяющее обойти дуги графа автомата A , не пройденные при подаче слова p_1 . При этом предполагаемые и истинные переходы такие:

$$s_3 \xrightarrow{1/0} s_1 \xrightarrow{1/0} s_1 \xrightarrow{0/1} s_2 \xrightarrow{0/0} s_1$$

Наблюдаемая реакция A на слово p_2 дает возможность завершить заполнение таблицы T_2 . Предположим теперь, что начальным состоянием A было s_3 . Тогда переходы A при подаче на вход слова $p_1 p_2$ должны иметь вид

$$s_3 \xrightarrow{1/1} s_1 \xrightarrow{0/0} s_2 \xrightarrow{1/0} s_3 \xrightarrow{0/1} s_3 \xrightarrow{1/1} s_1 \xrightarrow{0/0} s_2 \xrightarrow{0/0} s_3 \xrightarrow{1/0} s_1 \xrightarrow{1/0} s_1 \xrightarrow{0/1} s_2 \xrightarrow{0/0} s_1$$

Из приведенной цепочки видно, что записи в таблице T_3 ,

соответствующие состояниям s_1 при входе 0 и s_3 при входе 1, будут противоречивыми. Это означает, что s_3 не может являться начальным состоянием A . Таким образом, рассматриваемый автомат A совпадает (или эквивалентен) с автоматом A_1 , функция выходов которого задана табл. 7.2, либо с автоматом A_2 , функция выходов которого задана табл. 7.3.

Таблица

7.1.

$s \backslash x$	0	1
s_1	s_2	s_1
s_2	s_1	s_3
s_3	s_3	s_1

Таблица

7.2.

$s \backslash x$	0	1
s_1	0	1
s_2	0	0
s_3	1	1

Таблица

7.3.

$s \backslash x$	0	1
s_1	1	0
s_2	0	1
s_3	0	0

Множество \tilde{S} в рассматриваемом случае есть $\{s_1^{A_1}, s_2^{A_2}\}$. Установочной последовательностью для расщепляемого автомата $\Delta(A_1, A_2)$ является, например, 0. При подаче на вход автомата A символа 0 пусть его реакция есть 1, т. е. конечным состоянием автомата $\Delta(A_1, A_2)$ является $s_2^{A_2}$. Отсюда следует, что рассматриваемый автомат A совпадает с A_2 или эквивалентен ему.

Приведем теперь некоторые соображения, сокращающие длину эксперимента. Такое сокращение возможно, например, когда у

рассматриваемого автомата $A = (S, X, Y, \delta, \lambda)$ существует слово $p \in X^*$, такое, что $|\delta(S, p)| < |S|$, где $\delta(S, p) = Y_{s \in S} \delta(s, p)$. В этом случае после подачи слова p на вход автомата число его допустимых начальных состояний уменьшается и, соответственно, уменьшается число таблиц T_j , которые необходимо строить в процессе проведения эксперимента. Отсюда возникает задача отыскания для заданного автомата A такого слова p , чтобы $|\delta(S, p)|$ была минимальной.

Эта задача может быть решена с использованием конструкции синхронизирующего дерева для алфавита A , описанной в лекции 1. Каждый лист этого дерева связан с некоторой A -группой. При этом ее мощность может равняться либо 1, что говорит о наличии у автомата A синхронизирующей последовательности, либо некоторому числу $k > 1$.

В последнем случае путь, ведущий из вершины синхронизирующего дерева в эту A -группу, не соответствует входной последовательности, являющейся синхронизирующей. Среди A -групп упомянутого дерева, связанных с его конечными ветвями, т. е. листьями, найдем такую, число элементов которой минимально. Заметим, что, вообще говоря, таких A -групп может быть несколько. Простые рассуждения показывают, что слово p , которое соответствует пути по дереву, ведущему из корня к ветви, связанной с минимальной по мощности A -группой, дает решение поставленной задачи.

Вернемся к рассмотренному выше примеру. Выполнив для автомата A построения синхронизирующего дерева, находим, что слово $p = 11$ является для него синхронизирующей последовательностью. После подачи его на вход автомат устанавливается в состояние s_1 , и завершающий этап эксперимента таким образом сводится к подаче на автомат слова p_1 , являющегося характеристическим для A в состоянии s_1 . Общая длина эксперимента при этом равна 9, тогда как длина построенного ранее эксперимента была равна 12.

Вопросы и упражнения

1. Сформулируйте постановку задачи контроля функции выходов неинициального автомата.
2. Каков критерий разрешимости этой задачи?
3. К какой классической задаче сводится исследуемая здесь задача?
4. Опишите процедуру решения соответствующей классической задачи.
5. Сформулируйте "обобщенную" установочную задачу и приведите достаточное условие ее разрешимости.
6. Приведите верхнюю оценку длины эксперимента, контролирующего функцию выхода неинициального автомата. Является ли эта оценка достижимой?

Контроль функции выходов инициального автомата с использованием кратного безусловного эксперимента

Исследуется задача контроля функции выходов автомата с помощью кратного безусловного эксперимента. Вводятся понятия покрытия, приведенного и минимального покрытия графа множеством путей и кратности покрытия. Показано, что исходная задача редуцируется к задаче нахождения покрытия графа автомата. Приводятся критерий разрешимости задачи и оценки длин покрытия. Описывается алгоритм построения покрытия.

Рассмотрим следующую задачу. Пусть в распоряжении экспериментатора есть несколько экземпляров автомата Мили A , у которого известны входной алфавит $X = \{x_1, \dots, x_m\}$, выходной алфавит $Y = \{y_1, \dots, y_l\}$, множество состояний $S = \{s_0, \dots, s_{n-1}\}$, где s_0 - выделенное состояние, называемое начальным, и функция переходов δ . Требуется построить кратный безусловный эксперимент, позволяющий распознать функцию выходов λ автомата A или эквивалентного ему автомата.

Легко видеть, что содержательный смысл этой задачи, в частности, соответствует также и контролю функции выходов инициального автомата кратным экспериментом.

Как и в [лекции 6](#), мы обратимся к решению более общей задачи в рамках теории графов, а интересующие нас результаты будут получены как частные случаи.

Определение 8.1. Множество P путей в графе $G(S, U)$, исходящих из начальной вершины s_0 этого графа, назовем покрытием, если каждая дуга $G(S, U)$ принадлежит некоторому пути из P .

Определение 8.2. Число путей покрытия назовем его кратностью, а сумму длин всех путей покрытия - его длиной.

Перейдем к рассмотрению задачи построения покрытия графа $G(S, U)$ с начальной вершиной s_0 .

Ниже будут получены необходимые и достаточные условия разрешимости этой задачи, установлены достижимые оценки кратности и длины покрытия, предложен метод его построения.

Условимся считать, что из каждой вершины s графа G исходит "пустой" путь нулевой длины, заканчивающийся в той же вершине. Будем говорить, что вершина s графа G достижима из вершины t , если в G существует путь из вершины t в вершину s . Из этого определения следует, что каждая вершина графа достижима из самой себя. Полагаем, что каждой вершине графа G инцидентна хотя бы одна дуга. Обозначим через T множество всех тех вершин графа G , из которых исходит хотя бы одна дуга.

Теорема 8.1. Покрытие графа G существует тогда и только тогда, когда любая вершина множества T достижима из начальной.

Доказательство. Пусть P - некоторое покрытие графа G и s - некоторая вершина из T . Поскольку из вершины s исходит некоторая дуга и эта дуга принадлежит одному из путей покрытия, то вершина s достижима из s_0 . Вершина s_0 достижима из самой себя по крайней мере по "пустому" пути. Отсюда следует, что любая вершина множества T достижима из s_0 . Предположим теперь, что в графе G любая вершина $s \in T$ достижима из s_0 . Зафиксируем произвольно некоторую вершину $t \in T$. Пусть u_1, u_2, \dots, u_{k_t} - путь, ведущий из s_0 в t , а $\{\nu_1, \dots, \nu_{r_t}\}$ - множество всех дуг, исходящих из t . Построим множество путей

$P_t = \{u_1 u_2 \dots u_{k_t} \nu_1, \dots, u_1 u_2 \dots u_{k_t} \nu_{r_t}\}$. Легко видеть, что объединение множеств P_t по всем $t \in T$ является покрытием графа G . Этим и заканчивается доказательство теоремы.

Граф, для которого существует покрытие, назовем правильным. Граф $H(S, V)$ назовем сверхграфом графа $G(S, U)$, если $V \supseteq U$. Сверхграф H графа G назовем регулярным, если любая дуга из множества $V \Gamma$ оканчивается в вершине s_0 . Дуги этого множества будем называть замыкающими.

Лемма 8.1. Каждому контуру C регулярного сверхграфа H графа G ,

проходящему через все дуги графа H , соответствует покрытие графа G

Доказательство. Пусть C - контур, проходящий через все дуги регулярного сверхграфа H . По этому контуру строим множество P всех максимальных по длине путей, входящих в C , но не содержащих замыкающих дуг. Поскольку каждая замыкающая дуга оканчивается в s_0 , то путь из P начинается в s_0 . Множество P получено из контура C исключением только всех замыкающих дуг. Отсюда следует, что любая дуга графа G входит хотя бы в один путь из P , т. е. есть покрытие.

Лемма 8.2. Каждому покрытию графа G соответствует регулярный сильно связный сверхграф H графа G и контур C , проходящий через все дуги H .

Доказательство. Пусть $P = \{p_1, \dots, p_k\}$ - покрытие графа G . По этому покрытию и графу G построим граф H следующим образом: если путь $p_i \in P$ оканчивается в вершине s , то проведем из s в s_0 замыкающую дугу r_i . Поскольку G - правильный граф, то полученный граф $H(S, V)$, где $V = UY\{r_1, \dots, r_k\}$, является сильно связным регулярным сверхграфом графа G , а путь $p_1 r_1 p_2 r_2 \dots p_k r_k$ - контуром, проходящим через все дуги графа H .

Определение 8.3. Покрытие $P = \{p_1, \dots, p_k\}$ графа G назовем приведенным, если выполняются следующие условия:

1. для всех $i \neq j, 1 \leq i, j \leq k$, путь p_i не является начальным отрезком пути p_j ;
2. для всякого q , являющегося начальным отрезком пути $p_i, q \neq p_i, i = 1, \dots, k$, множество $P' = \{p_1, \dots, p_{i-1}, q, p_{i+1}, \dots, p_k\}$ не является покрытием.

Пусть P - приведенное покрытие графа G . Из пункта 2 определения 8.3 следует, что если $p \in P$ и p оканчивается дугой u , то эта дуга не принадлежит никакому другому пути из P . Из последнего факта непосредственно вытекает, что для правильного графа G всегда

существует приведенное покрытие, кратность которого не превосходит числа дуг графа. Легко видеть, что эта оценка достижима для графов с любым числом дуг и вершин. Из определения 8.3 также следует, что покрытие минимальной длины является приведенным.

Если граф сильно связан, то всегда существует контур, проходящий через все его дуги, т. е. всегда существует приведенное покрытие этого графа кратности 1.

Рассмотрим теперь не сильно связный граф. Докажем ряд вспомогательных утверждений. Используемые обозначения те же, что и в лекции 6 ч. I.

Лемма 8.3. Для правильного не сильно связного графа множество положительных вершин не пусто.

Доказательство. Пусть G - некоторый правильный не сильно связный граф, у которого S^+ пусто. Тогда множество S^- также пусто и, следовательно, $\Delta(s) = 0$ для всех $s \in S$. Поскольку правильный граф связан, то из последнего факта следует, что граф G обладает эйлеровым контуром и, стало быть, он сильно связан. Полученное противоречие доказывает лемму.

Лемма 8.4. Из любой вершины связного графа, у которого S^+ не пусто, достижима хотя бы одна положительная вершина.

Доказательство. Пусть $F(S, U)$ - связный граф, у которого S^+ не пусто. Из каждой вершины $s \in S^+$ проведем ровно $\Delta(s)$ дуг, оканчивающихся в отрицательных вершинах, таким образом, чтобы каждая отрицательная вершина t служила концом ровно $\Delta(t)$ таких дуг.

В результате получим новый граф $F'(S, U)$, у которого $V \supseteq U$ и для всех вершин $s \Delta(s) = 0$. В силу теоремы 2 из [25, с.182], граф F' обладает эйлеровым контуром C .

Рассмотрим отрезок этого контура, лежащий между двумя дугами из

множества $V\Gamma$. Этот отрезок является, очевидно, путем, ведущим из некоторой отрицательной вершины графа F' в некоторую положительную вершину. По построению графа F'' в каждую отрицательную вершину графа F' входит хотя бы одна дуга из $V\Gamma$. Следовательно, из каждой отрицательной вершины графа F' достижима некоторая положительная.

Любая положительная вершина достижима из самой себя по "пустому" пути. Рассмотрим, наконец, произвольную вершину s , у которой $\Delta(s) = 0$. Из нее достижима положительная вершина, поскольку из s выходит по крайней мере одна дуга и эта дуга содержится в пути, порожденном контуром C , из отрицательной в положительную вершину. Лемма доказана.

Рассмотрим граф G , множество S^+ которого не пусто. Из каждой вершины $s \in S'$, где $S' \subseteq S^+$, проведем в вершину s_0 ровно $\Delta(s)$ дуг. Эти дополнительные дуги обозначим через r_1, \dots, r_k , где $k = \sum_{s \in S} \Delta(s)$. Полученный граф обозначим через $H(G, S')$.

Теорема 8.2. Для правильного не сильно связанного графа G всегда существует покрытие, кратность которого не превосходит величины

$$\sum_{s \in S^+} \Delta(s) \quad (8.1)$$

Доказательство. В силу леммы 8.3 множество S^+ для правильного не сильно связанного графа G не пусто. Из теоремы 8.1 следует, что граф G связан. Построим граф $H(G, S^+)$. В силу леммы 8.4 и правильности графа G граф $H(G, S^+)$ является сильно связным регулярным сверхграфом. Зафиксируем некоторый контур C графа $H(G, S^+)$, проходящий через все его дуги. По лемме 8.1 этому контуру соответствует покрытие P , содержащее пути, полученные из C исключением всех замыкающих дуг r_1, \dots, r_k графа $H(G, S^+)$. Из правила построения этого графа следует, что между любыми двумя вхождениями замыкающих дуг в контуре C имеется хотя бы одна дуга графа G . Для завершения доказательства теоремы достаточно показать,

что в графе $H(G, S^+)$ существует такой контур C' , проходящий через все дуги этого графа, что каждая дуга T_i входит в этот контур только один раз. По построению в графе $H(G, S^+)$ для каждой вершины $s \in S^+ \Delta(s) = 0$ и положительной может быть только вершина s_0 . Поскольку из s_0 достижима любая вершина графа, то всегда можно построить компенсирующую систему графа $H(G, S^+)$, используя только дуги из U . Следовательно, искомый контур C' существует. По лемме 8.1 ему соответствует покрытие кратности (8.1), что и требовалось доказать.

Легко видеть, что оценка (8.1) достижима для всех ее значений.

Оценим снизу кратность покрытия правильного графа G . Напомним, что множество S состоит из вершин графа G , из которых не выходят дуги. По ранее сделанному предположению любая вершина графа G инцидентна хотя бы одной дуге, но тогда $S \subseteq S^+$. Пусть мощность множества S равна r . По лемме 8.2 всякому покрытию графа G соответствует некоторый регулярный сильно связный сверхграф H графа G , число замыкающих дуг которого не меньше r . Отсюда следует, что кратность любого покрытия графа G не может быть меньше величины r . Легко видеть, что эта оценка достижима для всех значений r .

Перейдем к уточнению нижней оценки кратности. Множество R^+ вершин графа G определим следующим образом: $s \in R^+$, если $s \in S^+$ и из вершины s не достижима ни одна отрицательная вершина графа G . Очевидно, что $S \subseteq S^+$ и для правильного графа $s_0 \notin R^+$. Кроме того, для правильного не сильно связного графа G множество R^+ не пусто. Доказательство этого утверждения может быть проведено по аналогии с доказательством леммы 8.3.

Покажем справедливость следующего утверждения.

Теорема 8.3. Кратность покрытия правильного графа не может быть меньше величины

$$\sum_{s \in R^+} \Delta(s) \quad (8.2)$$

Доказательство. Пусть $P = \{p_1, \dots, p_k\}$ - некоторое покрытие графа G . По этому покрытию, используя доказательство леммы 8.3, построим сильно связный регулярный сверхграф H графа G . Пусть C - контур графа H , соответствующий покрытию P . Каждую дугу графа H заменим дугами l с тем же началом и концом, где l - число вхождений дуги в контур. Полученный в результате граф H' обладает эйлеровым контуром. Поскольку для любой вершины s графа $H' * \Delta(s) = 0$, то в вершину s_0 заходит не менее $\sum_{s \in R^+} \Delta(s)$ замыкающих дуг. Таким образом, кратность покрытия графа G не может быть меньше величины (8.2).

Рассмотрим граф, изображенный на рис.8.1. Для этого графа $S^+ = \{s_3, s_7, s_8\}$, $S = \{s_8\}$, $R^+ = \{s_7, s_8\}$, $\sum_{s \in S^+} \Delta(s) = 3$, $\sum_{s \in S^+} \Delta(s) = 2$.

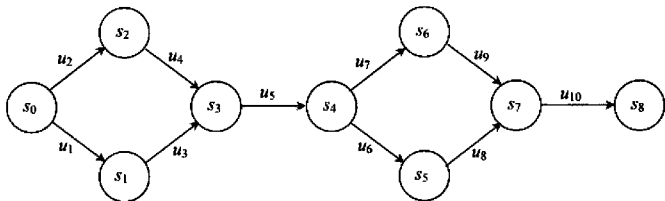


Рис. 8.1.

Множество $P = \{u_1 u_3 u_5 u_6 u_8 u_{10}, u_2 u_4 u_5 u_7 u_9\}$ является, очевидно, приведенным покрытием кратности 2, т. е. оценка (8.2) в данном примере достижима. Нетрудно показать, что (8.2) является достижимой для любого своего значения.

Следствие 1. Если для правильного не сильно связного графа G множества R^+ и S^+ совпадают, то минимальная кратность покрытия графа G равна величине (8.1).

Справедливость этого утверждения вытекает непосредственно из теорем 8.2, 8.3 и того факта, что при совпадении множеств S^+ и R^+ оценки (8.1) и (8.2) равны.

Определение 8.4. Приведенное покрытие графа, являющееся минимальным по кратности и по длине, назовем минимальным.

Покажем, что минимальное покрытие графа не всегда существует. С этой целью рассмотрим граф, изображенный на рис.8.2.

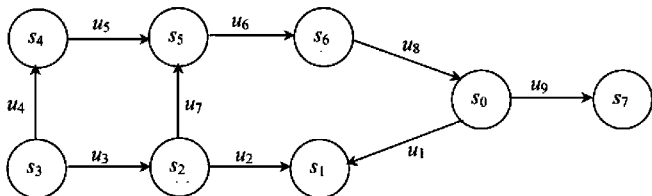


Рис. 8.2.

Покрытия $P_1 = \{u_1 u_2 u_3 u_4 u_5 u_6 u_8, u_1 u_2 u_7 u_6 u_8 u_9\}$ и $P_2 = \{u_1 u_2 u_3 u_4 u_5 u_6 u_8 u_9, u_1 u_2 u_7\}$, как легко видеть, являются приведенными покрытиями этого графа. Покрытие P_1 является минимальным по кратности (кратность его равна 1), но не минимальным по длине (длина его равна 13). Покрытие P_2 (длины 11 и кратности 2) является минимальным по длине, но не минимальным по кратности. Легко установить, что для данного графа минимального покрытия не существует.

Перейдем к отысканию условий, при которых минимальное покрытие существует. Покажем справедливость следующего вспомогательного утверждения.

Лемма 8.5. Если P есть покрытие графа G минимальной длины, то любой путь из P , не оканчивающийся в начальной вершине графа, оканчивается в некоторой положительной его вершине.

Доказательство. Пусть P - покрытие графа $G(S, U)$ минимальной

длины l . Предположим, что существует путь $p \in P$, оканчивающийся в некоторой не положительной вершине w , где $w \neq s_0$. Покрытие P соответствует регулярный сильно связный сверхграф H , способ построения которого описан при доказательстве леммы 8.3.

Преобразуем граф H следующим образом: каждую дугу $u \in U$ заменим f дугами $u(1), \dots, u(f)$ с тем же началом и концом, где f - число путей из P , содержащих дугу u . Легко видеть, что полученный в результате граф H' обладает эйлеровым контуром. Поскольку w не обладает положительной вершиной графа G и $w \neq s_0$, то в графе H эта вершина отрицательна, а в графе H' , очевидно, $\Delta(w) = 0$. Отсюда следует, что найдется такая дуга $u_1 \in U$, оканчивающаяся в вершине w , что в графе H' существуют дуги $u_1(1), \dots, u_1(f)$, где $f \geq 2$. Выполним следующее преобразование графа H' : исключим из него дугу $u_1(f)$ и одну замыкающую дугу, исходящую из w . Далее, из начальной вершины дуги $u_1(f)$, если она отлична от s_0 , проводим одну замыкающую дугу в s_0 . Полученный в результате такого преобразования новый граф H^* , как это следует из его построения, обладает эйлеровым контуром.

Число дуг графа H^* по крайней мере на единицу меньше числа дуг графа H' . Любому эйлерову контуру графа H' соответствует покрытие P^* графа G длины, не большей $l - 1$. Отсюда следует, что P не является покрытием минимальной длины. Полученное противоречие доказывает лемму.

Сформулируем достаточное условие существования у графа минимального покрытия.

Теорема 8.4. Если у правильного графа множества S^+ и R^+ совпадают, то у него существует минимальное покрытие.

Доказательство. Если у правильного графа G множество S^+ пусто, то он обладает эйлеровым контуром. Очевидно, что эйлеров контур является минимальным покрытием графа.

Предположим теперь, что множество S^+ не пусто, следовательно, S^- также не пусто. Поскольку $S^+ = R^+$, а множество R^+ состоит из вершин, из которых не достижима ни одна отрицательная вершина, то граф G заведомо не является сильно связным. В силу следствия 1 теоремы 8.3 минимальная кратность покрытия графа G в таком случае равна (8.2).

Пусть P есть покрытие графа G минимальной длины d и каждый путь из P оканчивается в вершине, не равной s_0 . Тогда по лемме 8.5 каждый путь из P оканчивается в вершине множества R^+ . Предположим, что P не минимально по кратности, т. е. его кратность k больше (8.1). Пусть покрытию P соответствует сильно связный регулярный граф H (см. доказательство леммы 8.2). По графу H построим эйлеров граф H' , заменяя каждую дугу $u \in U$ столькими дугами (с тем же началом и концом), скольким путям из P принадлежит дуга u . Поскольку k больше (8.1), то найдется вершина $t \in S_+$, из которой в графе H выходит l замыкающих дуг, где $l > \Delta(t)$. Следовательно, в графе H найдется дуга $v \in U$, входящая в вершину t и принадлежащая по крайней мере двум путям из P . Тогда в графе H' удалим одну дугу v' , соответствующую дуге v , и удалим одну замыкающую дугу, исходящую из t . Затем из начала дуги v' в вершину s_0 проведем одну дугу. Легко видеть, что полученный граф H_* является эйлеровым и его эйлерову контуру соответствует покрытие графа G длины $d - 1$. Следовательно, P не является покрытием минимальной длины. Противоречие доказывает, что P - минимально еопкрытие.

Рассмотрим, наконец, минимальное по длине покрытие $P = \{p_1, \dots, p_k\}$ графа G , у которого некоторые пути оканчиваются в s_0 . Пусть кратность покрытия P больше (8.1) и пути p_1, \dots, p_l оканчиваются в s_0 . Легко видеть, что покрытие $P' = \{p_1 p_2 \dots p_{l+1} p_{l+2}, \dots, p_k\}$ также является покрытием минимальной длины и все пути из P' оканчиваются только в положительных вершинах. Выше было показано, что кратность покрытия P' минимальна. Теорема доказана.

Перейдем к описанию алгоритма построения покрытия правильного графа G . Пути искомого покрытия обозначим через p_i . Каждой вершине $s \in T'$ поставим в соответствие множество U_s^0 всех дуг, выходящих из s . Дуги в U_s^0 занумеруем числами $1, 2, \dots$ причем если $u_i, u_j \in U_s^0$ и u_i - петля, то $i < j$.

В алгоритме будут использованы функции φ и δ , определяемые такими соотношениями: $\varphi(u)$ - номер дуги u во множестве U_s^0 , $\delta(s, p)$ - конечная вершина пути p , исходящего из вершины s .

В описанном алгоритме кроме построения покрытия производится вычисление параметра j , необходимого для оценки длины полученного покрытия.

Для краткости при описании алгоритма используем операцию присваивания "=", употребляемую в том же смысле, что и в алгоритмических языках.

Алгоритм А.

1. $P := \emptyset, i := 1, p_i := e, s := s_0, j = 0$; для всех $s \in T' : U_s := U_s^0$.
2. Если $U_s \neq \emptyset$, то перейти к пункту 3, иначе к пункту 4.
3. $p_i = p_i u$, где u - дуга из множества U_s , имеющая наименьший номер; $U_s := U_s \setminus \{u\}, s := \delta(s, u)$; если $\varphi(u) \neq 1$, то $j := j + 1$; перейти к пункту 2.
4. Если $\bigcup_{s \in S} U_s \neq \emptyset$ то перейти к пункту 5, иначе работа алгоритма завершается.
5. Если из вершины s достижима хотя бы одна вершина t , для которой $U_t \neq \emptyset$, то перейти к пункту 6, иначе к пункту 7.
6. Построить кратчайший путь Q_j , ведущий из вершины s в вершину из множества $\{t | U_t \neq \emptyset\}$, $p_i := p_i Q_j$, и перейти к пункту 3.
7. $s := s_0; P := PY\{p_i\}; i := i + 1; p_i := e$ и перейти к пункту

3.

Заметим, что в случае сильно связного графа полученное по этому алгоритму покрытие имеет кратность 1.

Оценим сверху длину покрытия, построенного с помощью алгоритма A . Пусть m_s - мощность множества $U_s, s \in T$. Множество $\{m_s\}_{s \in T}$ упорядочим отношением " \leq " и обозначим через (m_1, \dots, m_s) , где $m_i \leq m_j$, если $i < j$.

Алгоритм A можно считать пошаговым процессом, где j -м шагом будут являться построения, выполняемые алгоритмом при фиксированном значении параметра j . Путь q_j , описанный в пункте 6 алгоритма A , будем называть транслирующим. Очевидно, что длина покрытия, построенного по алгоритму, равна сумме числа дуг графа и длин всех транслирующих путей.

Оценим суммарную длину всех транслирующих путей покрытия, построенного по алгоритму A .

Из описанного алгоритма следует, что длина транслирующего пути q_j не превосходит числа таких вершин s из T , для которых на j -м шаге $U_s = \emptyset$. Поэтому для значений j от 0 до $m_1 - 1$ включительно $q_j = e$, где e - "пустой" путь. Далее, для значений j от m_1 до $m_2 - 1$ включительно длина q_j не превосходит единицы. Из описания алгоритма следует, что на последнем шаге работы алгоритма значение j равно $\sum_{i=1}^f (m_i - 1)$. Отсюда, учитывая предыдущие рассуждения, легко показать, что суммарная длина транслирующих путей, построенных по алгоритму, не превосходит величины

$$Q = \sum_{i=1}^f (i - 1)(m_i - 1) \quad (8.3)$$

Следовательно, справедливо следующее утверждение.

Теорема 8.5. Для правильного графа G всегда существует покрытие, длина которого не превышает величины

$$W + Q \quad (8.4)$$

где W - число дуг графа.

Следствие 1. Для правильного n -вершинного автоматного графа степени m всегда существует покрытие длины, не превышающей величины

$$mn + \frac{1}{2}(m - 1)n(n - 1) \quad (8.5)$$

Следствие непосредственно вытекает из теоремы 8.5, так как для автоматного графа $W = mn$ и $Q = \frac{1}{2}(m - 1)n(n - 1)$.

В [16] доказано, что для всех n и m существует n -вершинный сильно связный автоматный граф степени m , у которого имеется контур, проходящий через все дуги графа, минимальной длины, равной (8.5). Из этого факта следует, что оценка (8.5) достижима при всех n и m . Следовательно, оценки (8.4) и (8.3) также достижимы.

Перейдем теперь к автоматной интерпретации полученных результатов. Если автомат Мили задан в виде конечного ориентированного автоматного графа, то каждому пути покрытия графа ставится в соответствие некоторое входное слово, а множество слов, соответствующих всем путям покрытия, составит кратный безусловный эксперимент. Таким образом, для распознавания функции выходов заданного инициального автомата или ему эквивалентного с n состояниями и m входными сигналами может быть использован кратный эксперимент, длина которого не превышает величины (8.5), а кратность - величины (8.2). Необходимое и достаточное условие существования кратного эксперимента, а также алгоритм его построения получают простой переформулировкой теоремы 8.1 и алгоритма А.

Вопросы и упражнения

1. Приведите постановку задачи контроля функции выходов автомата с помощью кратного эксперимента.
2. Приведите определение покрытия графа и его кратности.

3. Каков критерий существования покрытия графа?
4. Дайте определение приведенного покрытия графа.
5. Приведите верхнюю и нижнюю оценки кратности для правильного не сильно связанного графа.
6. Дайте определение минимального покрытия графа.
7. Каково условие существования минимального покрытия?
8. Опишите процедуру построения покрытия правильного графа.

Преобразование автомата для упрощения функционального контроля

Описаны способы контроля автоматов с применением специальных схем, называемых схемами встроенного контроля. Предлагается одна из разновидностей такой схемы, основанная на принципе восстановления входных сигналов. В ней используются ОБПИК-автоматы порядка 1. Описывается метод преобразования произвольного автомата в автомат названного типа, базирующийся на выведении из исходного автомата дополнительных контрольных точек.

Необходимость в таком преобразовании возникает при решении проблемы повышения надежности цифровых устройств (ЦУ). Решение упомянутой проблемы в качестве обязательного этапа включает обнаружение неисправности в ЦУ, поскольку лишь после установления факта отказа ЦУ имеет смысл проводить работы по восстановлению его правильного функционирования. Существующие методы контроля ЦУ подразделяются на тестовые и функциональные. Каждый из них имеет свои достоинства и недостатки. Ниже речь будет идти о реализации функционального контроля, который требует некоторых аппаратурных затрат, что является его недостатком, но осуществляется в процессе использования ЦУ по его прямому назначению, что относится к несомненным достоинствам.

Обычно функциональный контроль реализуется так, как это изображено на рис.9.1. На этом рисунке блок СВК представляет собой схему встроенного контроля. Как видно из рисунка, рабочие воздействия X поступают как на ЦУ, так и на СВК. Помимо последовательности X на входы СВК поступают также выходные реакции Y ЦУ. На выход F СВК выводится информация о техническом состоянии (исправности или неисправности) ЦУ.

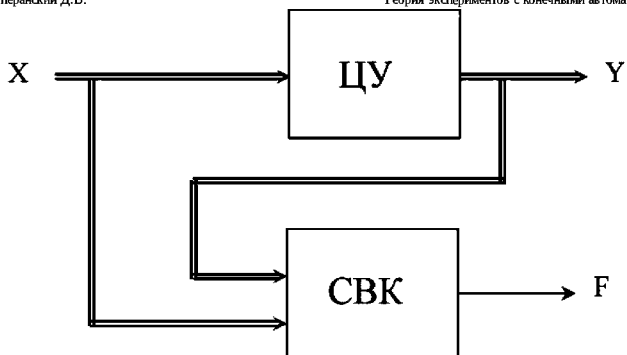


Рис. 9.1.

Один из принципов построения СВК для комбинационных устройств (КУ), используемый на практике, базируется на восстановлении входов ЦУ (рис.9.2).

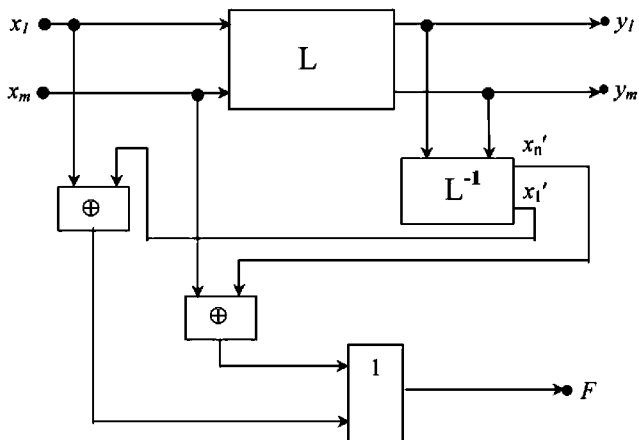


Рис. 9.2.

Известно, что для КУ блок L^{-1} существует не всегда, а необходимым и достаточным условием его существования является наличие у булевых функций, реализуемых блоком L , свойства обратимости [80].

Аналогичный принцип построения СВК может быть использован и для устройств с памятью, описываемых математической моделью автомата Мили.

В лекции 4 был рассмотрен класс ОБПИК-автоматов, которые обладают следующим свойством: по наблюдаемой реакции такого автомата на неизвестную входную последовательность длины k всегда возможно однозначное восстановление первого символа этой входной последовательности независимо от начального состояния автомата из заданного множества S_0 допустимых состояний. Отметим, что упомянутое восстановление осуществляется с помощью специальной схемы, описанной в лекции 4, которая фактически является СВК, изображенной на рис.9.2. Существенный недостаток использования упомянутой СВК при $k:0$ состоит в том, что сигнал на ее выходе F_0 возникшей неисправности в ЦУ появится не в момент ее первого проявления, а через k тактов. Понятно, что такая задержка может привести к катастрофическому отказу ЦУ. Таким образом, чтобы исключить такую ситуацию, необходимо, чтобы порядок k ОБПИК-автомата был равен 1.

В связи с изложенным возникает задача преобразования произвольного автомата в ОБПИК-автомат порядка 1, которая и будет исследована в этом разделе.

Указанное преобразование будем осуществлять за счет расширения выходного алфавита автомата, достигаемого путем выведения контрольных точек.

Рассмотрим слабоинициальный минимальный автомат Мили $A = (S, X, Y, \beta, \gamma, S_0)$, где S_0 - множество допустимых начальных состояний $1 \leq |S_0| \leq |S|$.

Как следует из теоремы 4.1, необходимое и достаточное условие принадлежности автомата классу ОБПИК порядка 1 состоит в следующем: в его проверочном графе, конструкция которого описана в лекции 4, должны отсутствовать выделенные дуги.

Легко показать, что выведение контрольных точек из автомата A , преобразующее его в ОБПИК-автомат порядка 1, эквивалентно нахождению такого разбиения $\Pi = \{K_1, \dots, K_\gamma$ множества $Q = S + X$, чтобы проверочный граф автомата $\tilde{A} = (S, X, Y \times \Pi, \beta, \Lambda)$ не имел выделенных дуг. Функция выходов в автомате \tilde{A} такова, что $\Lambda(s, x) = (\lambda(s, x), K(s, x))$, где $K(s, x)$ - код класса разбиения Π , содержащего пару (s, x) в некотором стандартном алфавите.

В формулировке рассматриваемой задачи отсутствует требование минимальности множества контрольных точек, поскольку при его наличии никакого другого метода ее решения, кроме перебора, не существует. Справедливость последнего утверждения вытекает из того, что задача поиска минимального множества контрольных точек может быть сведена к классической задаче о покрытии, относящейся к числу NP -полных проблем.

Напомним, что дуга $\{s_k, s_l\} \rightarrow \{s_i, s_j\}$ с пометкой y проверочного графа $T(A)$ автомата A называется выделенной, если

$$\exists_{x_1, x_2 \in X} \beta(s_k, x_1) = s_i \beta(s_l, x_2) = s_j \lambda(s_k, x_2) = y \rightarrow x_1 \neq x_2$$

Очевидно, что в проверочном графе выделенные дуги имеются тогда и только тогда, когда один и тот же выходной сигнал встречается не менее чем в двух различных столбцах таблицы переходов-выходов (ТПВ) автомата. Следующая лемма дает возможность оценить количество выделенных дуг в проверочном графе автомата.

Лемма 9.1. Пусть символ $y \in Y$ находится в m клетках ТПВ автомата A , а p_1, p_2, \dots, p_y - номера всех столбцов ТПВ, которые этот символ содержит. Если n_1, n_2, \dots, n_y - количество символов $y \in Y$ в

соответствующих столбцах, то число выделенных дуг в проверочном графе автомата A , помеченных символом Y , равно величине

$$C_m^2 - \sum_{i=1}^r C_{n_i}^2 \quad (9.1)$$

где $\sum_{i=1}^r C_{n_i}^2 = m$ и $C_{n_i}^2 = 0$, если $n_i/lt; 2, i = 1, 2, \dots, r$.

Доказательство приведенной формулы достаточно очевидно, поэтому мы не будем на нем останавливаться.

Вернемся теперь к рассматриваемой задаче отыскания множества выходных контрольных точек. По построению каждой выделенной дуге

$\{s_k, s_l\} \rightarrow |s_i, s_j|$ с пометкой Y графа $T(A)$ соответствует две дуги:

$s_k \xrightarrow{x_1/Y} s_i$ и $s_l \xrightarrow{x_2/Y} s_j$ графа автомата A , которые далее именуется ее составляющими. Очевидно, что если изменить выходной сигнал Y одной из этих дуг, то в проверочном графе для автомата с измененной функцией выхода эта выделенная дуга будет отсутствовать. Воспользуемся этим для удаления всех выделенных дуг в проверочном графе.

Опишем теперь способ получения разбиения Π множества $Q = S * X$. С этой целью сначала сформируем список D всех выделенных дуг в проверочном графе $T(A)$ автомата A , упорядочив его произвольным образом. На основе списка D составим множество $L = \{l_1, \dots, l_v\}$ всех попарно различных дуг графа автомата A , являющихся составляющими выделенных дуг этого списка. Далее сформируем множество $M = \{(l_{i_1}, l_{i_2}), \dots, (l_{i_a}, l_{i_b})\}$ всех таких пар дуг автомата, которые совместно не образуют выделенной дуги проверочного графа. Содержательно принадлежность пары (l_i, l_j) множеству M означает, что в автомате \tilde{A} , полученном из A в результате расширения выходного алфавита, дуги l_i и l_j могут иметь одну и ту же пометку. Вместе с тем, если \tilde{l}_i и \tilde{l}_j принадлежат разным парам из M , то их пометки обязательно должны быть различными.

В принятых обозначениях алгоритм получения искомого разбиения Π можно сформулировать следующим образом:

1. $K_1 := \{l_1\}; K_2 := K_3 = \dots = \emptyset; i := 1, j := 1.$
2. $j := j + 1.$
3. Проверка: выполняется ли неравенство $\nu < \nu$? Если нет, то переходим к пункту 4, предварительно положив $i := 1$, иначе переходим к пункту 7.
4. Для всех элементов $d \in K$ проверяется условие: принадлежит ли (d, l_j) множеству M ? Если условие выполняется для всех $d \in K_i$, то K_i пополняем дугой l_j и переходим к пункту 2, в противном случае переходим к пункту 5.
5. $i := i + 1.$
6. Проверка: $K_i = \emptyset$? Если да, то класс K_i пополняем дугой l_j и переходим к пункту 2, иначе - к пункту 4.
7. Конец работы алгоритма.

Проиллюстрируем описанный способ на примере. Пусть автомат A , у которого $S_0 = S$, задан своей ТПВ (табл.9.1), в которой имеется m клеток ($m = 6$), помеченных одним и тем же выходным символом a , причем в первом столбце имеется $n_1 = 3$ таких символов, во втором - $n_2 = 1$ и в третьем - $n_3 = 2$. В соответствии с леммой 9.1 общее число выделенных дуг этого автомата равно $C_6^2 - (C_3^2 + C_2^2 + C_1^2) = 15 - (3 + 1 + 0) = 11.$

Таблица 9.1.

s	0	1	2
1	1, a	2, b	3, a
2	2, a	3, a	1, a
3	3, a	1, b	2, c

Приведем список D этих выделенных дуг, каждая из которых имеет одну и ту же пометку \underline{a} .

1. $(1, 1) \rightarrow (1, 3)$
2. $(1, 2) \rightarrow (1, 1)$
3. $(1, 2) \rightarrow (1, 3)$
4. $(1.2) \rightarrow (3.2)$
5. $(1.2) \rightarrow (3.3)$
6. $(1.3) \rightarrow (3.3)$
7. $(2.2) \rightarrow (2.2)$
8. $(2.2) \rightarrow (2.1)$
9. $(2.2) \rightarrow (2.1)$
10. $(2.3) \rightarrow (3.3)$
11. $(2.3) \rightarrow (1.3)$

Множество всех попарно различных составляющих выделенных дуг этого списка есть

$L = \{l_1 = (1, 1), l_2 = (1, 3), l_3 = (2, 2), l_4 = (2.3), l_5 = (2.1), l_6 = (3.3)\}$.
Сформируем теперь на основе L множество пар дуг $M = \{(l_1, l_3), (l_1, l_6), (l_2, l_5), (l_3, l_6)\}$.

Применение предложенного алгоритма дает следующее разбиение Π :

$K_1 = \{l_1, l_3, l_6\}; K_2 = \{l_2, l_5\}; K_3 = \{l_4\}$.

Поскольку разбиение имеет три класса, для кодирования символов классов достаточно двух двоичных переменных. Пусть

$K_1 = 00, K_2 = 10, K_3 = 01$, тогда автомат с контрольными точками представлен табл.9.2.

Таблица 9.2.

s	0	1	2
1	1, a 0 0 2, b 0 0 3, a 1 0		
2	2, a 0 0 3, a 0 1 1, a 1 0		
3	3, a 0 0 1, b 0 0 2, c 0 0		

Предположим, что выходные сигналы a, b, c закодированы двоичными

числами 00, 01 и 10. При таком кодировании часть СВК для рассматриваемого автомата, восстанавливающая входные сигналы, должна, как легко сообразить, реализовывать две булевы функции f_1 и f_2 , заданные табл.9.3.

Таблица 9.3.

Вход.набор	f_1	f_2
0000	0	0
0001	0	1
0010	1	0
0100	0	0
1000	1	0

Отметим, что класс автоматов без потери информации конечного порядка 1 представляет собой важный частный случай, широко используемый в теории кодирования и получивший название взаимно однозначных автоматов. По аналогии, рассмотренные выше ОБПИК-автоматы порядка 1 назовем сильно взаимно однозначными (СВО).

Исходя из общего определения ОБПИК-автоматов, определение СВО-автоматов может быть сформулировано следующим образом: слабоинициальный автомат $A = (S, X, Y, \beta, \lambda, S_0)$ назовем СВО, если

$$\forall s_1, s_2 \in \Gamma\{S_0\} \forall p_1, p_2 \in X * \lambda(s_1, p_1) = \lambda(s_2, p_2) \rightarrow p_1 = p_2 \quad (9.2)$$

где $\Gamma\{S_0\}$ - множество состояний автомата A , достижимых без состояний множества S_0 .

Из этого определения непосредственно следует критерий принадлежности автомата классу СВО.

Лемма 9.2. Слабоинициальный автомат A является СВО-автоматом тогда и только тогда, когда

$$\forall s_1, s_2 \in \Gamma\{S_0\} \forall x_1, x_2 \in X * \lambda(s_1, x_1) = \lambda(s_2, x_2) \rightarrow x_1 = x_2 \quad (9.$$

Отметим существенное различие между (9.1) и (9.2): если в (9.1) используются слова из алфавита X , то в (9.2) фигурируют символы этого алфавита.

Пусть $A_1 = (S_1, X, Y_1, \beta_1, \lambda_1, S_0^{(1)})$ и $A_2 = (S_2, Y_1, Y, \beta_2, \lambda_2, S_0^{(2)})$ - два слабоинициальных автомата, тогда автомат $A = (S_1 \times S_2, X, Y, \beta, \lambda, S_0)$ с множеством начальных состояний $S_0 = S_0^{(1)} \times S_0^{(2)}$ назовем суперпозицией автоматов A_1 и A_2 , если $\beta((s_1, s_2), x) = (\beta_1(s_1, x), \beta_2(s_2, x)), \lambda((s_1, s_2), x) = \lambda_2(s_2, \lambda_1(s_1, x))$.

Нетрудно показать справедливость следующего утверждения.

Лемма 9.3. Суперпозиция СВО-автоматов A_1 и A_2 является СВО-автоматом.

Понятно, что определение суперпозиции (последовательного соединения) может быть распространено на любое конечное число автоматов, но и в этом случае лемма 9.3 остается справедливой. Таким образом, функциональный контроль любого числа последовательно соединенных СВО-автоматов может быть осуществлен по схеме, изображенной на рис.9.1.

Вопросы и упражнения

1. Что представляет собой функциональный контроль автомата?
2. Опишите принципиальную схему организации такого контроля.
3. Опишите схему контроля, основанную на восстановлении входных сигналов.
4. Сформулируйте постановку задачи организации контроля с применением ОБПИК-автоматов порядка 1.
5. Опишите процедуру преобразования произвольного автомата в ОБПИК-автомат порядка 1 путем выведения контрольных точек.

Синхронизирующие эксперименты с линейными автоматами

Описываются модели стационарных и нестационарных линейных автоматов, их структура и элементарные составляющие. Приводятся основные понятия теории экспериментов с линейными автоматами, формулы для вычисления конечных состояний и реакций автоматов. Исследуются синхронизирующие последовательности: критерий существования, их свойства, верхняя оценка длины минимальных СП. Описан метод решения задачи перевода автомата в заданное синхросостояние.

Из теории автоматов известно, что реакция автомата с памятью непредсказуема, если неизвестно его начальное состояние. Вместе с тем эта реакция может быть предсказана, если имеется информация о том, из какого состояния данный автомат стартовал. Отсюда вполне очевидна важность задачи анализа автоматов, состоящей в распознавании его начального состояния. Другой важной задачей анализа является распознавание конечного (финального) состояния автомата, поскольку знание его позволяет ввести автомат в различные режимы работы, желательные для достижения заданной цели. Напомним, что эксперименты, служащие для определения начального и конечного состояний, принято называть диагностическими и установочными соответственно. Один частный случай установочного эксперимента называется синхронизирующим.

Эта лекция посвящена исследованию некоторых из перечисленных выше экспериментов для линейных автоматов (ЛА). Предварительно мы опишем объект исследования и приведем основные определения, которые потребуются в дальнейшем.

Основные определения

Начнем с краткого описания модели ЛА, а для более детального знакомства с ней отошлем читателя к монографиям [19], [66]. ЛА является системой с конечным числом входных полюсов, к которым подводятся внешние сигналы, и с конечным числом выходных полюсов,

на которых наблюдаются сигналы реакции. Воздействия, поступающие на ЛА, прикладываются одновременно ко всем входам в дискретные моменты времени, которые для удобства представляются целыми числами. Интервалы времени между двумя такими последовательными моментами называются тактами.

Структурная схема ЛА состоит из соединения конечного числа элементарных составляющих, каждая из которых мгновенно выполняет одну из трех функций:

- сложение входных сигналов по правилам конечного поля, над которым задан ЛА;
- умножение входного сигнала на константу по правилам конечного поля, над которым задан автомат;
- задержка входного сигнала на один временной такт.

Все операции совершаются одновременно, в результате чего на выходных полюсах ЛА появляются сигналы, принимающие значения из конечного поля, над которым задан ЛА. Таким образом, ЛА может рассматриваться как "черный" ящик с некоторым числом входов и выходов.

Определим три типа компонент, называемых элементарными составляющими ЛА над конечным полем $GF(p) = \{0, 1, \dots, p-1\}$

1. Сумматор. Сумматор имеет l входов и один выход. Если на входе появились сигналы $u_1(t), u_2(t), \dots, u_l(t)$, то на выходе появляется реакция $u_1(t) + u_2(t) + \dots + u_l(t) \pmod{p}$.
2. Усилитель. Усилитель с константой a из поля $GF(p)$ имеет один вход и один выход. Если на входе действует сигнал $u(t)$, то на выходе появляется сигнал $au(t) \pmod{p}$.
3. Задержка. Задержка имеет один вход и один выход. Если входной сигнал равен $u(t)$, то на выходе появляется сигнал $u(t-1)$. Сигнал $u(t)$ обычно называют состоянием задержки в интервале времени от t до $t+1$.

При построении ЛА допускается любое соединение конечного числа элементарных составляющих с одним исключением: не должно быть ни одной замкнутой петли, не содержащей по крайней мере одну задержку. Нарушение этого правила приводит к построению соединения, в котором действует неопределенный сигнал.

Отметим, что усилитель с константой 0 или 1 обозначают соответственно разрыв и прямое соединение. Таким образом, ЛА над полем $GF(2)$ (двоичный ЛА) состоит только из сумматоров по модулю 2 и задержек. Поэтому практическая привлекательность двоичных ЛА вполне очевидна.

Усилитель с константой $p - 1 = -1$ называется инвертором; такой элемент просто изменяет знак входного сигнала.

Двоичные ЛА являются очень удобными математическими моделями реальных электронных устройств.

Условимся считать, что число входных полюсов ЛА равно l , число выходных полюсов равно m . Предполагается, что входные сигналы принимают значение из поля $GF(p) = \{0, 1, \dots, p - 1\}$, где p - простое число. Под состоянием ЛА понимается упорядоченная совокупность состояний элементов задержек (обозначим их число через n), входящих в состав ЛА. Число n обычно называют размерностью ЛА, и множество состояний обозначают через S_n .

Введем следующие обозначения:

$$\bar{u}(t) = [u_1(t), \dots, u_l(t)]', \bar{y}(t) = [y_1(t), \dots, y_m(t)]', \bar{s}(t) = [s_1(t), \dots, s_n(t)]'.$$

Здесь $\bar{u}(t)$, $\bar{y}(t)$, $\bar{s}(t)$ - входной, выходной векторы и вектор-состояние соответственно, а через t обозначен момент дискретного времени.

Функционирование ЛА \tilde{A} задается системами уравнений состояний и выходов соответственно:

$$\bar{s}(t + 1) = A\bar{s}(t) + B\bar{u}(t) \quad (10.1)$$

$$\bar{y}(t) = C\bar{s}(t) + D\bar{u}(t) \quad (10.2)$$

где $A = [a_{i,j}]_{n \times n}$, $B = [b_{i,j}]_{n \times l}$, $C = [c_{i,j}]_{m \times n}$, $D = [d_{i,j}]_{m \times l}$, называются характеристическими матрицами ЛА. Матрица A обычно называется главной (основной) характеристической матрицей. Все перечисленные матрицы состоят из элементов поля $GF(p)$.

Методом математической индукции можно доказать, что конечное состояние и выходная реакция ЛА \tilde{A} на входное слово $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ длины $k + 1$ при начальном состоянии $\bar{s}(0)$ вычисляются по формулам

$$\begin{aligned} \bar{s}(k+1) &= A^{k+1}\bar{s}(0) + A^k B\bar{u}(0) + A^{k-1} B\bar{u}(1) + \dots + A B\bar{u}(k-1) + B\bar{u}(k), \\ y(k) &= CA^k \bar{s}(0) + CA^{k-1} B\bar{u}(0) + CA^{k-2} B\bar{u}(1) + \dots + C B\bar{u}(k-1) + D\bar{u}(k) \end{aligned}$$

Последняя из этих формул носит название формулы полной реакции ЛА.

Описанную выше модель принято называть стационарным ЛА. Наряду с ней рассмотрим и так называемые нестационарные ЛА (НЛА), функционирование которых задается следующими уравнениями состояний и выходов соответственно:

$$\bar{s}(t+1) = A(t)\bar{s}(t) + B(t)\bar{u}(t), \quad (10.5)$$

$$\bar{y}(t) = C(t)\bar{s}(t) + D(t)\bar{u}(t) \quad (10.6)$$

Размерность матриц в этих формулах та же, что и в формулах (10.1) и (10.2).

По аналогии со стационарными ЛА можно доказать, что если $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(t)$ - входная последовательность НЛА и $s(0)$ - начальное состояние НЛА, то его конечное состояние и выходная реакция вычисляются по следующим формулам:

$$\bar{s}(t+1) = A(t)A(t-1) \dots A(0)\bar{s}(0) + \sum_{i=1}^{t-1} A(t)A(t-1) \dots A(i+1)B(i)\bar{u}(i) + B(t)\bar{u}(t)$$

$$\bar{y}(t) = C(t)A(t-1) \dots A(0)\bar{s}(0) + \sum_{i=0}^{t-1} C(t)A(t-1) \dots A(i+1)B(i)\bar{u}(i) + D(t)\bar{u}(t)$$

Определим теперь различные типы экспериментов, которые будут исследоваться нами. Для их проведения необходимо иметь соответствующие последовательности, подаваемые на вход автомата и позволяющие по наблюдаемой реакции находить ответ на интересующий исследователя вопрос. Для разных типов экспериментов необходимо располагать различными типами последовательностей.

Напомним теперь определения соответствующих последовательностей и сделаем это для компактности записей применительно к самой общей модели конечного детерминированного автомата Мили.

Как и ранее, под автоматом Мили понимается пятерка объектов

$$A = (S, X, Y, \delta, \lambda)$$

где S, X, Y - конечные множества состояний, входной и выходной алфавиты соответственно, а $\delta : S \times X \rightarrow S$ и $\lambda : S \times X \rightarrow Y$ - отображения, называемые функциями переходов и выходов.

Пусть $S = \{s_1, \dots, s_n\}$, $X = \{x_1, \dots, x_l\}$.

Определение 10.1. Входная последовательность $p = x_{i_1}, x_{i_2}, \dots, x_{i_n}$ называется синхронизирующей (СП), если

$$\forall s_{j_1}, s_{j_2} \in S \delta(s_{j_1}, p) = \delta(s_{j_2}, p)$$

Содержательно это определение означает, что СП P переводит автомат A в одно и то же конечное состояние независимо от того, из какого состояния он стартовал.

В этом и приведенных ниже определениях функции переходов и выходов предполагаются расширенными на входные последовательности. Это расширение понимается общепринятым в

теории автоматов смысле.

Определение 10.2. Входная последовательность $p = x_{i_1}, x_{i_2}, \dots, x_{i_n}$ называется установочной (УП), если

$$\forall s_{j_1}, s_{j_2} \in S \lambda(s_{j_1}, p) = \lambda(s_{j_2}, p) \rightarrow \delta(s_{j_1}, p) = \delta(s_{j_2}, p)$$

Поясним содержательный смысл этого определения: по наблюдаемой реакции на установочную последовательность однозначно идентифицируется конечное состояние автомата.

Очевидно, что СП можно рассматривать как частный, а точнее, как вырожденный случай УП, поскольку подача СП вызывает перевод автомата в известное конечное состояние, хотя и не требует при этом наблюдения его реакции.

Определение 10.3. Входная последовательность $p = x_{i_1}, x_{i_2}, \dots, x_{i_n}$ называется диагностической (ДП), если

$$\forall s_{j_1}, s_{j_2} \in S \lambda(s_{j_1}, p) = \lambda(s_{j_2}, p) \rightarrow s_{j_1} = s_{j_2}$$

Обращаясь к содержательному смыслу этого определения, отметим, что знание реакции автомата на ДП позволяет однозначно идентифицировать его начальное состояние.

Понятно, что каждая ДП одновременно является и УП, поскольку, определив по реакции на ДП начальное состояние автомата, легко определить состояние, в котором он окажется после подачи ДП. Обратное, однако, неверно, т. е. не каждая УП является одновременно и ДП.

Разрешимость установочной и диагностических задач для автомата в общем случае зависит не только от самого автомата, но и от множества начальных состояний, в которых он может находиться перед проведением соответствующего эксперимента. Назовем это множество множеством допустимых начальных состояний автомата A и будем обозначать его через (A) . Обычно предполагается, что это множество совпадает со всем множеством S_n состояний автомата A . Если это не так, то соответствующий факт будет оговариваться особо.

Отметим, что для автоматов Мили, в общем случае нелинейных, в монографии А. Гилла [18] условия существования перечисленных выше последовательностей сформулированы в терминах весьма громоздкой конструкции дерева преемников и потому проверка их является очень трудоемким процессом. Теперь мы перейдем к исследованию условий существования СП для линейных автоматов.

Условия существования синхронизирующей последовательности

Приведенное выше определение СП применительно к ЛА формулируется следующим образом: входную последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ ЛА A назовем СП, если

$$\forall \bar{s}_i(0), \bar{s}_j(0) \in S_n \\ A^k \bar{s}_i(0) + A^{k-1} B \bar{u}(0) + \dots + B \bar{u}(k-1) = A^k \bar{s}_j(0) + A^{k-1} B \bar{u}(0) + \dots + B \bar{u}(k-1)$$

Перенося в (10.9) правую часть равенства влево, получим

$$\forall \bar{s}_i(0), \bar{s}_j(0) \in S_n A^k (\bar{s}_i(0) - \bar{s}_j(0)) = [0] \quad (10.10)$$

Символом $[0]$ здесь и далее обозначается нулевая матрица или нулевой вектор подходящей размерности.

Поскольку в (10.10) $\bar{s}_i(0)$ и $\bar{s}_j(0)$ есть произвольные состояния, то (10.10) эквивалентно предикату

$$\forall \bar{s} \in S_n A^k \bar{s} = [0] \quad (10.11)$$

Теорема 10.1. Для того чтобы ЛА A имел СП длины k , необходимо и достаточно, чтобы $A^k = [0]$.

Доказательство. Соотношение

$$A^k \bar{s} = [0] \quad (10.12)$$

можно интерпретировать как систему линейных алгебраических уравнений (СЛАУ) относительно переменных s_1, s_2, \dots, s_n , являющихся координатами вектор-столбца \bar{s} . Пусть ранг матрицы A^k равен r , где $r \leq n$. Из алгебры известно, что в этом случае число свободных неизвестных $s_{r+1}, s_{r+2}, \dots, s_n$ в системе (10.12) равно $n - r$. Тогда переменные s_1, s_2, \dots, s_r будут выражаться через свободные переменные и, следовательно, число решений системы равно p^{n-r} . Поскольку СП для ЛА существует тогда и только тогда, когда решением системы (10.12) является любой вектор $\bar{s} \in S_n$, где $|S_n| = p^n$, то из сравнения p^n и p^{n-r} вытекает, что r должно равняться нулю, т. е. ранг матрицы A^k равен 0. Из последнего равенства следует, что $A^k = [0]$.

Определение 10.4. Линейный автомат A будем называть синхронизируемым, если у него существует СП.

Из теоремы 10.1 вытекает следующее следствие.

Следствие. Для того чтобы ЛА A был синхронизируем, необходимо, чтобы его главная характеристическая матрица A была вырожденной.

Доказательство. Пусть ЛА синхронизируем, тогда существует такое целое k , что $A^k = [0]$. Это означает, что матрица A^k является вырожденной. Из алгебры известно, что определитель произведения матриц равен произведению определителей. Отсюда следует, что $|A^k| = |A|^k$, где $|A^k|$ и $|A|$ - определители матриц A^k и A соответственно. Поскольку $A^k = [0]$, то $|A^k| = 0$, но $|A^k| = |A|^k$, следовательно $|A|^k = 0$. Отсюда следует, что $|A| = 0$, т. е. матрица A является вырожденной.

Теорема 10.2. Если для ЛА A существует хотя бы одна СП длины k , то для этого автомата синхронизирующими являются любые входные последовательности длины k и более.

Доказательство. Из теоремы 10.1 следует, что для ЛА, имеющего СП

$\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ длины k , конечное состояние ЛА после подачи СП выражается по формуле (10.3) следующим образом:

$$\bar{s}(k) = A^{k-1}B\bar{u}(0) + A^{k-2}B\bar{u}(1) + \dots + AB\bar{u}(k-2) + B\bar{u}(k-1)$$

Поскольку правая часть последнего равенства не зависит от начального состояния $\bar{s}(0)$ ЛА, это означает, что при любом начальном состоянии некоторая фиксированная входная последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ переводит ЛА в одно и то же конечное состояние. Этот факт остается справедливым и для любой другой фиксированной входной последовательности. Иными словами, любая входная последовательность длины k является для этого ЛА синхронизирующей. Ясно, что и любая входная последовательность большей длины также будет синхронизирующей.

Заметим, что последняя теорема свидетельствует о принципиальном различии между общими автоматами Мили и линейными автоматами с точки зрения возможности их синхронизации. Если для некоторого автомата Мили СП длины k существует, то общее их число может быть невелико. В частности, автомат Мили может иметь одну единственную СП. В то же время для ЛА существование одной СП длины k автоматически влечет существование p^k СП такой же длины.

Обратимся теперь к вопросу о том, как практически установить существование у заданного ЛА СП и если таковая существует, то как найти длину минимальной СП. Из теории автоматов известно, что установочная задача для автомата с ν состояниями и с ν допустимыми начальными состояниями всегда может быть решена с помощью простого безусловного эксперимента длины l , где $l \leq (\nu - 1)^2$. Задача синхронизации автомата является, очевидно, частным случаем установочной задачи. Используя этот факт, для установления факта существования СП необходимо вычислять последовательно степени A^k главной характеристической матрицы ЛА для $k = 1, 2, \dots$, до тех пор, пока на очередном шаге она не выродится в нулевую. Очевидно, что наименьшее целое k такое, для которого $A^k = [0]$, равно длине минимальной СП. Если же в процессе вычисления окажется, что

$A^k \neq [0]$ при $k = (\nu - 1)^2$, где $\nu = 2^n$, n - размерность ЛА, то СП для этого ЛА не существует и на этом процесс вычисления степеней матриц прекращается.

Упомянутая выше верхняя граница длины минимальной СП, равная величине $(2^n - 1)^2$, достаточно велика. Укажем один частный вид ЛА, для которого эта верхняя граница существенно ниже.

Назовем квадратную матрицу верхней (нижней) треугольной, если все ее элементы, лежащие на главной диагонали и ниже (выше) нее, равны нулю.

Теорема 10.3. Если главная характеристическая матрица ЛА размерности n является верхней (нижней) треугольной, то длина минимальной СП для этого ЛА не превосходит n .

Доказательство. Проведем его для верхней треугольной матрицы, которую обозначим через A :

$$\begin{bmatrix} 0 & a_{12} & a_{13} & \dots & a_{1,n-1} & a_{1,n} \\ 0 & 0 & a_{23} & \dots & a_{2,n-1} & a_{2,n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-2,n-1} & a_{n-2,n} \\ 0 & 0 & 0 & \dots & 0 & a_{n-1,n} \\ 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Условимся нумеровать диагонали матрицы, параллельные ее главной диагонали и расположенные выше нее, в порядке убывания числа элементов в них. Тогда диагональ, содержащая элементы $a_{12}, a_{23}, \dots, a_{n-2,n-1}, a_{n-1,n}$, получит номер 1, а диагональ, содержащая единственный элемент a_{1n} , - номер $n - 1$. Непосредственными вычислениями можно убедиться, что в матрице A^2 все элементы первой диагонали равны 0, в матрице A^3 все элементы первой и второй диагонали также равны 0. Методом индукции можно доказать, что в матрице A^k наряду с ее нулевой нижней треугольной подматрицей диагонали с номерами $1, 2, \dots, k - 1$ также содержат только нули. Из последнего

утверждения следует, что по крайней мере в матрице A^n все элементы равны 0. Тогда на основании теоремы 10.1 рассматриваемый ЛА имеет СП, длина которой не превосходит n .

Для нижней треугольной матрицы соответствующее утверждение доказывается аналогично.

Ниже будет показано, что в действительности длина минимальной СП не превосходит n для произвольного синхронизируемого ЛА размерности n , а не только для частного вида ЛА, фигурирующего в теореме 10.3.

Состояние ЛА A , в котором он оказывается после подачи некоторой его СП, назовем синхросостоянием.

Обозначим через (A) множество всех возможных синхросостояний ЛА. Ясно, что в общем случае попарно различные СП могут переводить ЛА как в различные, так и в совпадающие синхросостояния, т. е.

$$|(A)| \leq p^n.$$

Рассмотрим следующую задачу. Пусть задан синхронизируемый ЛА A , синхросостояние $\bar{s} \in (A)$ и пусть k - длина минимальной СП этого автомата. Требуется найти такую входную последовательность длины k , которая переводит ЛА A в синхросостояние \bar{s} .

Для рассматриваемого ЛА выражение (10.13) можно переписать в виде

$$A^{k-1}B\bar{u}(0) + A^{k-2}B\bar{u}(1) + \dots + AB\bar{u}(k-2) + B\bar{u}(k-1) = \bar{s}$$

Далее (10.14) будем рассматривать как СЛАУ (неоднородных) относительно неизвестных

$u_1(0), \dots, u_l(0), \dots, u_1(k-1), \dots, u_l(k-1)$, общее число которых равно $l \times k$.

Пусть $Q = [q_{ij}]_{n \times lk}$ есть матрица системы (10.14), а \bar{Q} есть расширенная (добавлением к матрице Q столбца \bar{s}) матрица той же системы и $Q = r$. Необходимым и достаточным условием

совместности системы (10.14) является, как известно из алгебры, условие $\text{rank } Q = \bar{Q}$ (теорема Кронекера - Капелли). Понятно, что для выполнения этого условия вектор-столбец \bar{s} должен быть линейной комбинацией r линейно независимых столбцов матрицы Q . Иными словами, только для вектор-столбцов \bar{s} , представляющих собой такие комбинации, соответствующая СП может быть найдена. Ясно, что если для некоторого \bar{s} система (10.14) оказывается несовместной, это означает, что заданное \bar{s} не принадлежит множеству синхросостояний (A) .

Напомним, как могут быть найдены решения системы (10.14) в случае ее совместности. С этой целью выберем в матрице Q r линейно независимых строк и в (10.14) оставим лишь те уравнения, коэффициенты которых вошли в выбранные строки. В левых частях этих уравнений оставляем такие r неизвестных, что определитель из коэффициентов при них отличен от нуля. Остальные неизвестные в этих уравнениях объявляем свободными и переносим в правые части уравнений. Варьируя значения свободных переменных (все они являются элементами поля $GF(p)$) и вычисляя значения остальных неизвестных (например, по правилу Крамера), получим все решения системы (10.14).

Проиллюстрируем сказанное на примере. Пусть ЛА над полем $GF(2)$ задан характеристическими матрицами:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix},$$

$$b = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

где $n = 4, l = 2$, тогда $\bar{u} = [u_1, u_2]'$.

По теореме 10.3 этот ЛА синхронизируем и длина его минимальной СП равна 4. Система (10.14) в нашем случае в матричной форме имеет вид

$$A^3 B \bar{u}(0) + A^2 B(1) + AB \bar{u}(2) + B \bar{u}(3) = \bar{s} \quad (10.15)$$

Вычислим матрицы $A^3 B$, $A^2 B$, AB :

$$A^3 B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{bmatrix},$$

$$A^2 B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$AB = \begin{bmatrix} 0 & 0 \\ 1 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}$$

Используя эти матрицы, перепишем систему уравнений (10.15) в координатной форме, где неизвестными являются $u_1(0), u_2(0), u_1(1), u_2(1), u_1(2), u_2(2), u_1(3), u_2(3)$ - координаты векторов $\bar{u}(0), \bar{u}(1), \bar{u}(2), \bar{u}(3)$:

$$u_1(3) + u_2(3) = s_1,$$

$$u_1(2) + u_2(2) + u_1(3) + u_2(3) = s_2,$$

$$u_1(1) + u_2(1) + u_1(3) + u_2(3) = s_3,$$

$$u_1(0) + u_2(0) + u_1(1) + u_2(1) + u_1(2) + u_2(2) + u_1(3) + u_2(3) = s_4.$$

Здесь s_1, s_2, s_3, s_4 есть координаты вектора \bar{s} .

Матрица этой системы имеет вид

$$Q = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}$$

Вычисления показывают, что $Q = 4$. Перенумеруем столбцы матрицы Q слева направо начиная с 1. Выберем в Q четыре линейно независимых столбца, например, с номерами 1, 3, 5, 7, которым соответствуют неизвестные $u_1(0), u_1(1), u_1(2), u_1(3)$. Легко проверить, что определитель, составленный из названных столбцов, отличен от нуля. Оставим эти перечисленные неизвестные в левых частях уравнений системы (10.16), а остальные переменные объявим свободными и перенесем в правые части уравнений:

$$\begin{aligned} u_1(3) &= u_2(3) + s_1, \\ u_1(2) + u_1(3) &= u_2(2) + u_2(3) + s_2, \\ u_1(1) + u_1(3) &= u_2(1) + u_2(3) + s_3, \\ u_1(0) + u_1(1) + u_1(2) + u_1(3) &= u_2(0) + u_2(1) + u_2(2) + u_2(3) + s_4. \end{aligned}$$

После простых и очевидных преобразований последняя система примет вид

$$\begin{aligned} u_1(3) &= u_2(3) + s_1, \\ u_1(2) &= u_2(2) + s_1 + s_2, \\ u_1(1) &= u_2(1) + s_1 + s_3, \\ u_1(0) &= u_2(0) + s_1 + s_2 + s_3 + s_4. \end{aligned} \tag{10.17}$$

Выберем, например, состояние $\bar{s} = (1, 1, 1, 1)'$ и проверим, является ли оно синхросостоянием. Если да, то найдем минимальную по длине СП, переводящую заданный ЛА в это синхросостояние. Условимся также среди всех минимальных по длине СП найти такую, которая имеет минимальный вес. Весом СП назовем число, которое равно сумме координат входных символов, составляющих эту СП.

Поскольку в нашем примере $s_1 = s_2 = s_3 = s_4 = 1$, то система (10.17) примет вид:

$$u_1(3) = u_2(3) + 1; u_1(2) = u_2(2); u_1(1) = u_2(1); u_1(0) = u_2(0).$$

Для получения искомой СП минимального веса положим $u_2(3), u_2(2), u_2(1), u_2(0)$ равными нулю. Тогда искомая СП, переводящая заданный ЛА из любого состояния в состояние $(1, 1, 1, 1)'$, такова: $(0, 0)'$, $(0, 0)'$, $(0, 0)'$, $(1, 0)'$.

Вопросы и упражнения

1. Опишите состав и устройство линейного автомата.
2. Приведите системы уравнений состояний и выходов, описывающих функционирование стационарных (нестационарных) линейных автоматов.
3. Приведите формулы для определения конечного состояния и реакции стационарного (нестационарного) линейного автомата при подаче на него последовательности входных символов.
4. Дайте определения синхронизирующих, установочных и диагностических последовательностей и поясните их содержательный смысл.
5. Сформулируйте критерий существования синхронизирующих последовательностей для стационарных линейных автоматов.
6. Охарактеризуйте свойства синхронизирующих последовательностей.
7. Сформулируйте постановку задачи перевода линейного автомата в заданное синхросостояние.
8. Опишите метод решения задачи перевода линейного автомата в заданное синхросостояние.
9. ЛА над полем $GF(2)$ задан следующими характеристическими матрицами:

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix},$$

$$b = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix},$$

$C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ является ли для этого ЛА (диагностическая) последовательность.

10. ЛА над полем $GF(2)$ имеет следующие характеристические матрицы:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$$

Покажите, что этот ЛА является синхронизируемым и найдите длину минимальной СП.

Проверьте, является ли состояние $s = [1, 1, 1]'$ этого ЛА синхросостоянием, и если да, то найдите минимальную СП, переводящую ЛА в это состояние.

Установочные и диагностические эксперименты со стационарными и нестационарными линейными автоматами

Приводятся условия существования установочных и диагностических последовательностей для стационарных линейных автоматов, сформулированные в терминах характеристических матриц автоматов. Исследуются свойства последовательностей. Аналогичные результаты приводятся для нестационарных линейных автоматов.

Условия существования установочной последовательности

Сформулированное ранее определение УП для автомата Мили применительно к линейному автомату принимает следующий вид: входную последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ ЛА A назовем УП, если

$$\begin{aligned} \forall \bar{s}(0), \hat{s}(0) \in S_n \bigwedge_{d=0}^k [CA^d \bar{s}(0) + CA^{d-1} B \bar{u}(0) + \dots + CB \bar{u}(d-1) + D \bar{u}(d) = \\ = CA^d \hat{s}(0) + CA^{d-1} B \bar{u}(0) + \dots + CB \bar{u}(d-1) + D \bar{u}(d)] \rightarrow \\ \rightarrow [A^{k+1} \bar{s}(0) + A^k B \bar{u}(0) + \dots + AB \bar{u}(k-1) + B \bar{u}(k) = \\ = A^{k+1} \hat{s}(0) + A^k B \bar{u}(0) + \dots + AB \bar{u}(k-1) + B \bar{u}(k)] \end{aligned}$$

Знак $\bigwedge_{d=0}^k$ в (11.1) означает конъюнкцию $k+1$ выражений, стоящих после этого знака и получаемых при изменении индекса d от 0 до k .

Отметим также, что допустимое множество состояний ЛА при построении УП предполагается совпадающим со всем множеством S_n состояний ЛА.

Теорема 11.1. Для ЛА A УП длины $k+1$ существует тогда и только тогда, когда

$$(\forall \bar{s} \in S_n \bigvee_{d=0}^k CA^d \bar{s} \neq [0]) \vee A^{k+1} = [0] \quad (11.2)$$

Доказательство. Вначале отметим, что знак $\bigwedge_{d=0}^k$ в (11.2) означает дизъюнкцию $(k + 1)$ выражений и имеет аналогичный со знаком $\bigwedge_{d=0}^k$ в (11.1) смысл.

Преобразуем (11.1), перенеся в левую часть обоих равенств все слагаемые, участвующие в них, и произведем соответствующие сокращения:

$$\forall \bar{s}(0), \hat{s}(0) \in S_n \wedge d = 0^k (CA^d(\bar{s}(0) - \hat{s}(0)) = [0]) \rightarrow A^{k+1}(\bar{s}(0) - \hat{s}(0)) = [0]$$

Используя известный факт, что высказывание $x \rightarrow y$ равносильно $\bar{x} \wedge y$, получаем, что последний предикат эквивалентен следующему:

$$\forall \bar{s}(0), \hat{s}(0) \in S_n \wedge_{d=0}^k (CA^d(\bar{s}(0) - \hat{s}(0)) \neq [0]) \wedge A^{k+1}(\bar{s}(0) - \hat{s}(0)) = [0]$$

Поскольку $\bar{s}(0)$ и $\hat{s}(0)$ - произвольные состояния из S_n , то очевидно, что $\bar{s}(0) - \hat{s}(0)$ пробегает все состояния из S_n . Тогда равенство $A^{k+1}(\bar{s}(0) - \hat{s}(0)) = [0]$ равносильно равенству $A^{k+1} = [0]$ и, следовательно, предикат (1.20) эквивалентен предикату

$$\forall \bar{s} \in S_n (\bigwedge_{d=0}^k CA^d \bar{s} \neq [0]) \wedge A^{k+1} = [0]$$

Теорема 11.2. Если для ЛА \tilde{A} , у которого характеристическая матрица C невырожденная, существует хотя бы одна УП длины $k + 1$, то для этого автомата установочными являются любые входные последовательности длины $k + 1$ и более.

Доказательство. Напомним, что аналогичное утверждение относительно синхронизирующих последовательностей, являющихся частным случаем УП, было доказано выше. Поэтому теорему 11.2 достаточно доказать для собственно установочных последовательностей, не являющихся СП.

Доказательство проведем от противного. Предположим, что некоторая последовательность p длины $k + 1$ является УП, но наряду с ней

существует последовательность $\bar{p} = \bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ той же длины, которая для рассматриваемого ЛА установочной последовательностью не является. Это означает, что у ЛА существует два таких различных состояния \bar{s}_1 и \bar{s}_2 , что, стартуя в них, он выдает одинаковые выходные реакции $\bar{y}(0), \bar{y}(1), \dots, \bar{y}(k)$, но переходит в различные конечные состояния \bar{s}_1^f и \bar{s}_2^f . Выпишем реакции и конечные состояния ЛА, соответствующие различным начальным состояниям.

$$\left\{ \begin{array}{l} \bar{y}(0) = C\bar{s}_1 + D\bar{u}(0), \\ \bar{y}(1)CA\bar{s}_1 + CB\bar{u}(0) + D\bar{u}(1), \\ \dots\dots\dots \\ \bar{y}(k) = CA^k\bar{s}_1 + CA^{k-1}B\bar{u}(0) + \dots + CB\bar{u}(k-1) + D\bar{u}(k), \\ \bar{s}_1^f = A^{k+1}\bar{s}_1 + A^k B\bar{u}(0) + \dots + AB\bar{u}(k-1) + B\bar{u}(k) \end{array} \right. \\ \left\{ \begin{array}{l} \bar{y}(0) = C\bar{s}_2 + D\bar{u}(0), \\ \bar{y}(1) = CA\bar{s}_2 + CB\bar{u}(0) + D\bar{u}(1), \\ \dots\dots\dots \\ \bar{s}_2^f = A^{k+1}\bar{s}_2 + A^k B\bar{u}(0) + \dots + AB\bar{u}(k-1) + B\bar{u}(k) \end{array} \right. \quad (11.5)$$

Учитывая, что $\bar{s}_1^f \neq \bar{s}_2^f$, из (11.4) и (11.5) получаем

$$\left\{ \begin{array}{l} C\bar{s}_1 = C\bar{s}_2 \\ CA\bar{s}_1 = CA\bar{s}_2 \\ \dots\dots\dots \\ CA^k\bar{s}_1 = A^{k+1}\bar{s} \end{array} \right. \quad (11.6)$$

Последнее равенство в (11.6) равносильно

$$A^{k+1}\bar{s} = \bar{b} \quad (11.7)$$

где $\bar{s} \neq [0]$ и $\bar{b} \neq [0]$.

Соотношение (11.7) можно интерпретировать как матричную форму записи системы линейных неоднородных уравнений относительно неизвестных, являющихся координатами вектор-столбца $\bar{s} = (s_1, s_2, \dots, s_n)'$. Поскольку эта система имеет ненулевое решение $(\bar{s}_1 - \bar{s}_2)$, тогда, как известно из алгебры [33], ее определитель $|A^{k+1}| \neq 0$. В силу того, что $|A|^{k+1} = |A^{k+1}| \neq 0$, определитель $|A|$ также отличен от нуля, т. е. матрица A невырожденная. Отсюда вытекает, что для нее существует обратная матрица A^{-1} , также невырожденная. Умножая слева последнее неравенство в (11.6) на A^{-1} , получим

$$A^k \bar{s}_1 \neq A^k \bar{s}_2$$

где матрица A^k - невырожденная. Поскольку по условию теоремы матрица C невырожденная, умножив на нее обе части последнего неравенства, получим

$$CA^k \bar{s}_1 \neq CA^k \bar{s}_2$$

что противоречит предпоследнему равенству в (11.6). Полученное противоречие доказывает теорему.

Доказанная теорема говорит о принципиальном различии между автоматами Мили и линейными автоматами с точки зрения идентификации их конечных состояний с помощью УП. В общем случае, если автоматы Мили имеют УП длины k , то число таких последовательностей меньше числа всех последовательностей этой же длины. В то же время для ЛА, заданного над полем $GF(p)$, как утверждает теорема, существование одной УП длины k влечет существование p^k таких УП.

Таким образом, для ЛА задача построения УП фактически сводится к задаче нахождения такого натурального числа k , при котором УП длины k существует. В то же время для автоматов Мили в общем случае задача построения УП далеко не тривиальна и требует применения специально разработанных для этого методов.

Заметим, что поиск целого числа k , являющегося длиной минимальной УП для ЛА, осуществляется точно так же, как и поиск длины минимальной СП, что было описано в предыдущем разделе.

Для иллюстрации теоремы 11.2 рассмотрим ЛА над полем $GF(2)$, заданный следующими характеристическими матрицами:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Обозначим состояния этого ЛА через $\bar{s}_1 = [0, 0]'$, $\bar{s}_2 = [0, 1]'$, $\bar{s}_3 = [1, 0]'$, $\bar{s}_4 = [1, 1]'$ и его выходные реакции через $\bar{y}_1 = [0, 0]'$, $\bar{y}_2 = [0, 1]'$, $\bar{y}_3 = [1, 0]'$, $\bar{y}_4 = [1, 1]'$.

Заметим, что рассматриваемый ЛА не имеет СП, поскольку его основная характеристическая матрица A невырожденная. Приведенная ниже таблица показывает, что для этого ЛА УП являются все входные последовательности длины 1.

Входное слово	Начальные состояния	Реакция	Конечные состояния
$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\bar{s}_1, \bar{s}_2, \bar{s}_3, \bar{s}_4$	$\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4$	$\bar{s}_1, \bar{s}_4, \bar{s}_3, \bar{s}_2$
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\bar{s}_1, \bar{s}_2, \bar{s}_3, \bar{s}_4$	$\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4$	$\bar{s}_3, \bar{s}_2, \bar{s}_1, \bar{s}_4$
$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\bar{s}_1, \bar{s}_2, \bar{s}_3, \bar{s}_4$	$\bar{y}_2, \bar{y}_1, \bar{y}_4, \bar{y}_3$	$\bar{s}_1, \bar{s}_4, \bar{s}_3, \bar{s}_2$

$$\begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \bar{s}_1, \bar{s}_2, \bar{s}_3, \bar{s}_4 \quad \bar{y}_2, \bar{y}_1, \bar{y}_4, \bar{y}_3 \quad \bar{s}_3, \bar{s}_2, \bar{s}_1, \bar{s}_4$$

Теорема 11.3. Если для n -мерного ЛА существуют УП, то минимальная их длина не превосходит величины n .

Доказательство. В [19] доказано следующее утверждение: $CA^k\bar{s} = [0]$ для любого $k \geq 1$ тогда и только тогда, когда $K\bar{s} = [0]$, где диагностическая матрица

$$K = \begin{bmatrix} C \\ CA \\ CA^2 \\ \dots \\ CA^{n-1} \end{bmatrix}$$

Очевидно, что это равносильно следующему предикату:

$$\exists k \geq CA^k\bar{s} \neq [0] \leftrightarrow K\bar{s} \neq [0] \quad (11.8)$$

Последнее утверждение фактически обосновывает достаточность проверки равенства $CA^k\bar{s} = [0]$ только для $k \leq n - 1$. Если при этих значениях $kCA^k\bar{s} = [0]$, то и при $k \geq n$ это равенство также остается справедливым. Отсюда и вытекает справедливость теоремы.

Заметим, что поскольку СП для ЛА является частными случаями УП, то приведенная в теореме 10.6 верхняя оценка длины УП является одновременно и верхней оценкой длины для СП.

Условия существования диагностической последовательности

Теорема 11.4. Для того чтобы n -мерный ЛА A имел ДП длины t , необходимо и достаточно, чтобы ранг матрицы

$$K_t = \begin{bmatrix} C \\ CA \\ CA^2 \\ \dots \\ CA^{t-1} \end{bmatrix}$$

был равен n .

Доказательство. Выпишем реакции ЛА A , стартующего в неизвестном начальном состоянии $\bar{s}(0)$, на входную последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(t-1)$ длины t :

$$\bar{y}(0) = C\bar{s}(0) + D\bar{u}(0)$$

$$\bar{y}(1) = CA\bar{s}(0) + CB\bar{u}(0) + D\bar{u}(1),$$

.....

$$\bar{y}(t-1) = CA^{t-1}\bar{s}(0) + CA^{t-2}B\bar{u}(0) + \dots + CB\bar{u}(t-2) + D\bar{u}(t-1)$$

Простыми и очевидными преобразованиями эти равенства всегда можно привести к виду

$$C\bar{s}(0) = [Z_0],$$

$$CA\bar{s}(0) = [Z_1], \quad (11.9)$$

.....

$$CA^{t-1}\bar{s}(0) = [Z_{t-1}]$$

где $[Z_i], i = 1, 2, \dots, t-1$ - некоторые конкретные вектор-столбцы размерности m , где m - число выходов ЛА.

Существование ДП длины t для рассматриваемого ЛА A означает, что СЛАУ (11.9) относительно неизвестных $s_1(0), \dots, s_n(0)$, являющихся координатами вектора $\bar{s}(0)$, должна иметь единственное решение, которое и соответствует искомому начальному состоянию ЛА. Из алгебры известно, что СЛАУ обладает единственным решением тогда и только тогда, когда ранг этой системы равен числу неизвестных. Отсюда вытекает справедливость теоремы.

Теорема 11.5. Если у ЛА существует хотя бы одна ДП длины k , то для этого ЛА диагностическими являются любые входные последовательности длины k и более.

Справедливость этой теоремы вытекает из того, что матрица K_t , фигурирующая в теореме 11.4, не зависит от входных слов.

Что касается процедуры поиска минимального значения k , при котором для заданного ЛА существует ДП, то она полностью совпадает с процедурой поиска минимальной длины СП и УП, описанной выше.

Теорема 11.6. Если для n -мерного ЛА существует ДП, то минимальная их длина не превышает величины n .

Справедливость этой теоремы следует из того, что каждая ДП является одновременно и УП, а для УП такая верхняя оценка установлена в предыдущем подразделе.

В [18] было установлено, что в общем случае минимальность автомата Мили является необходимым, но не достаточным условием существования для него ДП. Покажем, что для ЛА имеет место следующее утверждение.

Теорема 11.7. Если ЛА минимален, то у него существует ДП.

Доказательство. Проведем его в предположении, что множество всех состояний S_n ЛА и множество его допустимых начальных состояний совпадают.

Напомним, что в [19] была доказана справедливость следующего утверждения: состояния \bar{s}_1 и \bar{s}_2 ЛА эквивалентны тогда и только тогда, когда $K\bar{s}_1 = K\bar{s}_2$. С учетом этого утверждения минимальность ЛА означает

$$\forall s_1, s_2 \in S_n K\bar{s}_1 = K\bar{s}_2 \rightarrow \bar{s}_1 = \bar{s}_2$$

Последнее означает, что СЛАУ $K\bar{s} = [0]$ в случае минимальности ЛА должна иметь только нулевые решения. Из алгебры известно, что это выполняется только тогда, когда ранг матрицы K системы равен n , т.

е. числу неизвестных этой системы. Отсюда в силу теоремы 11.4 вытекает справедливость доказываемой теоремы.

Отыскание ДП для заданного ЛА обычно называют диагностической задачей. Известно [18], что возможность решения диагностической задачи для автомата Мили зависит от множества его допустимых начальных состояний, а также от применяемых для этого средств. В [18] показано, что наиболее мощным средством для решения диагностических задач являются кратные эксперименты, меньшими возможностями обладают простые безусловные эксперименты. Ниже будет показано, что для линейных автоматов упомянутая иерархия разрешающих возможностей перечисленных типов экспериментов не существует.

Теорема 11.8. Для любого минимального ЛА и любого множества его допустимых начальных состояний диагностическая задача всегда разрешима с помощью простого безусловного диагностического эксперимента.

Доказательство. Пусть у минимального ЛА множество допустимых начальных состояний совпадает со всем множеством его состояний. Напомним, что автомат называется определенно диагностируемым, если существует такое натуральное число k , что все входные последовательности длины k являются для него диагностическими. Из теоремы 11.5 следует, что любой ЛА является определенно диагностируемым. Покажем, что любое входное слово длины n , где n - размерность ЛА, различает любые два состояния минимального ЛА. Предположим противное. Пусть существуют два таких состояния ЛА, которые не различаются некоторым словом длины n . С учетом теоремы 11.6 это означает, что рассматриваемый минимальный ЛА не является определенно диагностируемым, что противоречит теореме 11.5. Полученное противоречие доказывает наше утверждение.

Эта теорема фактически означает, что простой безусловный эксперимент с ЛА обладает большими возможностями, чем аналогичный эксперимент с автоматом Мили, поскольку, как известно из [18], в последнем случае диагностическая задача разрешима не для всякого множества допустимых начальных состояний, мощность которого более двух.

Обратимся теперь к простым условным диагностическим экспериментам. Известно [18], что для автоматов Мили существуют диагностические задачи для множества допустимых начальных состояний более чем с двумя элементами, которые не разрешимы простым безусловным, но разрешимы простым условным экспериментом. Из теоремы 11.8 следует, что такая ситуация не может иметь места для минимальных ЛА. Другими словами, все диагностические задачи, разрешимые простыми условными экспериментами, могут быть разрешены и простыми безусловными. Таким образом, для ЛА разрешающие возможности обоих типов диагностических экспериментов просто совпадают.

Остановимся теперь на кратных диагностических экспериментах. В [18] доказано, что для любого автомата диагностическая задача для произвольного множества допустимых начальных состояний всегда разрешима кратным безусловным экспериментом. Тогда, как это следует из теоремы 11.8, для минимальных ЛА разрешающие возможности кратного безусловного и простого экспериментов совпадают.

Подведем итог: для минимальных ЛА ни простые условные, ни кратные безусловные диагностические эксперименты не привносят ничего нового по сравнению с простыми безусловными диагностическими экспериментами с точки зрения принципиальной возможности решения диагностических задач.

Сделанный вывод, однако, не означает, что условные и кратные диагностические эксперименты являются бесполезными средствами при решении диагностических задач. Нетрудно построить такие примеры диагностических задач, решения которых с помощью условных и кратных экспериментов достигаются последовательностями меньшей длины по сравнению с длиной безусловного эксперимента.

Синхронизирующие, установочные и диагностические эксперименты с нестационарными линейными автоматами

Конкретизируем определение 1.1 синхронизирующей последовательности применительно к нестационарному ЛА:

последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(t)$ является СП для НЛА \tilde{A} , если

$$\begin{aligned} \forall s_1(0), s_2(0) \in (\tilde{A}) * A(t)A(t-1) \dots A(0)s_1(0) + \\ + \sum_{i=0}^t A(t) \dots A(i+1)B(i)\bar{u}(i) = A(t)A(t-1) \dots A(0)s_2(0) + \\ + \sum_{i=0}^t A(t)A(t-1) \dots A(i+1)B(i)\bar{u}(i) \end{aligned}$$

В этом определении (\tilde{A}) означает множество допустимых начальных состояний ЛА.

Теорема 11.9. Для того чтобы входная последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(t)$ была СП для НЛА \tilde{A} , необходимо и достаточно, чтобы

$$\forall s_1(0), s_2(0) \in (\tilde{A}) A(t)A(t-1) \dots A(0)[s_1(0) - s_2(0)] = [0] \quad (11.10)$$

Доказательство получается путем переноса в левую часть всех членов равенства в приведенном определении СП.

Следствие 1. Если $(\tilde{A}) = S_n$, то необходимым и достаточным условием существования СП длины $t+1$ для НЛА является выполнение равенства

$$A(t)A(t-1) \dots A(0) = [0] \quad (11.11)$$

Доказательство. В силу произвольности начальных состояний $s_1(0)$ и $s_2(0)$ разность $\bar{s}(0) = s_1(0) - s_2(0)$ пробегает все множество состояний S_n , следовательно, (11.10) можно переписать так:

$$\forall \bar{s}(0) \in S_n A(t)A(t-1) \dots A(0)\bar{s}(0) = [0]$$

Понятно, что последнее справедливо тогда и только тогда, когда справедливо (11.11).

Следствие 2. Если $(\tilde{A}) = S_n$ и для НЛА существует хотя бы одна СП длины t , то для него синхронизирующей является любая входная

последовательность длины t и более.

Справедливость этого утверждения вытекает из того, что условие (11.11) не зависит от входной последовательности.

Исследуем теперь вопрос об оценке длины СП для НЛА.

Рассмотрим НЛА \tilde{A} со следующими главными характеристическими матрицами:

$$A(i) = E, i = \overline{1, t-1}, A(t) = 0$$

где E - единичная матрица.

Очевидно, что для этого НЛА, как это вытекает из (11.11), СП имеет длину $t+1$ и для него не существует СП меньшей длины. В силу произвольности параметра t отсюда следует, что в общем случае длина минимальной СП для НЛА не ограничена сверху. Напомним, что в отличие от НЛА для стационарных ЛА верхняя граница длины минимальных СП, как было показано выше, не превосходит величины n , где n - размерность ЛА.

Поскольку в общем случае задание НЛА требует перечисления бесконечных последовательностей характеристических матриц, что не всегда можно сделать конструктивно, рассмотрим специальный класс НЛА, описываемый конечными множествами таких матриц. НЛА этого класса назовем периодическими и потребуем, чтобы периодическими были все его характеристические матрицы. Последнее означает, что существует такая целая положительная константа λ , что $A(t+\lambda) = A(t)$, $B(t+\lambda) = B(t)$ и т. д.

Перейдем теперь к исследованию условий существования СП периодических НЛА. Построим по периодической НЛА \tilde{A} стационарный ЛА, обозначаемый как $A_c \tilde{m}$, у которого функция переходов имеет вид

$$\bar{s}(t+1) = A\hat{s}(t)$$

где

$$\hat{A} = A(\lambda - 1)A(\lambda - 2) \dots A(0)$$

Теорема 11.10. СП для периодической НЛА \tilde{A} существует тогда и только тогда, когда она существует для стационарного ЛА $A_{c\tilde{m}}$

Доказательство.

Необходимость. Пусть для НЛА \tilde{A} существует СП минимальной длины $\lambda k + t + 1$ где $t < \lambda$. Тогда по следствию 1 из теоремы 11.9 должно выполняться равенство

$$A(\lambda k + t)A(\lambda k + t - 1) \dots A(\lambda)A(\lambda - 1) \dots A(0) = [0]$$

В силу периодичности матрицы $A(t)$ последнее равенство эквивалентно равенству

$$A(t)A(t - 1) \dots A(0)\hat{A}^k = [0]$$

Отсюда вытекает, что $\hat{A}^k = [0]$ но по теореме 1.1 это есть необходимое и достаточное условие существования СП для стационарного ЛА $A_{c\tilde{m}}$

Достаточность. Пусть для линейного автомата $A_{c\tilde{m}}$ существует СП длины $k + 1$ тогда должно выполняться условие $A_{c\tilde{m}}^{k+1} = [0]$. Отсюда следует, что

$$\underbrace{[A(\lambda - 1) \dots A(0) \dots A(\lambda - 1(A(0)))]}_{k+1 \text{ раз}} = \\ = A((k + 1)\lambda - 1) \dots A(k\lambda)A(k\lambda - 1) \dots A((k - 1)\lambda) \dots A(\lambda - 1) \dots A(0)$$

Тогда, в силу справедливости (1.28), для НЛА \tilde{A} существует СП длина $(k + 1)\lambda$

Теорема 11.11. Если λ - период главной характеристической матрицы $A(t)$ НЛА \tilde{A} , то длина минимальной СП не превосходит величины λn , где n - размерность НЛА.

Справедливость этой теоремы вытекает из того, что длина минимальной СП стационарного ЛА размерности n , как было показано выше, не превосходит величины n .

Конкретизируем теперь определение 1.2 установочной последовательности применительно к нестационарному ЛА: последовательность $\bar{u}(0), \text{baru}(1), \dots, \bar{u}(t)$ является УП для НЛА \tilde{A} , если

$$\forall \bar{s}_1(0), \bar{s}_2(0) \in (\tilde{A})(\bigwedge_{k=0}^t \bar{y}_1(k) = \bar{y}_2(k)) \rightarrow (\bar{s}_1(t+1) = \bar{s}_2(t+1))$$

где $\bigwedge_{k=0}^t$ - символ конкатенации $k+1$ равенств, $\bar{y}_i(k), \bar{s}_i(k)$ - выходная реакция и состояние автомата в момент времени, стартующего из состояния $\bar{s}_i(0), i = 1, 3$

Теорема 11.12. Для того чтобы входная последовательность $\bar{u}(0), \text{baru}(1), \dots, \bar{u}(t)$ являлась УП для НЛА \tilde{A} , необходимо и достаточно, чтобы

$$\begin{aligned} & \forall \bar{s}_1(0), \bar{s}_2(0) \in (\tilde{A}) \\ & \exists k \in [0:t](C(k)A(k-1) \dots A(0)[\bar{s}_1(0) - \bar{s}_2(0)] \neq [0]) \vee (11.12) \\ & \vee (A(t)A(t-1) \dots A(0)[\bar{s}_1(0) - \bar{s}_2(0)] = [0]) \end{aligned}$$

Доказательство. Перепишем приведенное выше определение УП для НЛА в терминах характеристических матриц:

$$\begin{aligned} & \forall \bar{s}_1(0), \bar{s}_2(0) \in (\tilde{A}) \\ & (\bigwedge_{k=0}^t (C(k)A(k-1) \dots A(0)\bar{s}_1(0) + \\ & + \sum_{i=0}^{k-1} C(k)A(k-1) \dots A(i+1)B(i)\bar{u}(i) + D(k)\bar{u}(k) = \\ & = C(k)A(k-1) \dots A(0)\bar{s}_2(0) + \sum_{i=0}^{k-1} C(k)A(k-1) \dots A(i+1)B(i)\bar{u}(i) + D(k)\bar{u}(k))) \rightarrow \\ & \rightarrow \bar{s}_1(t+1) = \bar{s}_2(t+1) \end{aligned}$$

Выполнив преобразования выражения, стоящего после квантора общности, получим

$$\begin{aligned} & \forall \bar{s}_1(0), \bar{s}_2(0) \in (\tilde{A}) \\ & (\bigwedge_{k=0}^t (C(k)A(k-1) \dots A(0)\bar{s}_1(0) = C(k)A(k-1) \dots A(0)\bar{s}_2(0))) \rightarrow \\ & \rightarrow (A(t)A(0)\bar{s}_1(0) = A(t) \dots A(0)\bar{s}_2(0)) \end{aligned}$$

Учитывая, что $(x \rightarrow y) \leftrightarrow \bar{x} \vee y$ и $\overline{x \wedge y} \leftrightarrow \bar{x} \vee \bar{y}$, из последнего соотношения получим

$$\begin{aligned} & \forall \bar{s}_1(0), \bar{s}_2(0) \in (\tilde{A}) \\ & (\bigwedge_{k=0}^t C(k)A(k-1) \dots A(0)[\bar{s}_1(0) - \bar{s}_2(0)] = n\epsilon[0]) \vee \\ & \vee (A(t) \dots A(0)[\bar{s}_1(0) - \bar{s}_2(0)] = [0]) \end{aligned}$$

Очевидно, что последний предикат эквивалентен предикату, приведенному в формулировке теоремы.

Заметим, что если $(\tilde{A}) = S_n$, то разность $[\bar{s}_1(0) - \bar{s}_2(0)]$ пробегает все множество состояний S_n рассматриваемого НЛА *Atilde* и в этом случае условие (11.12) принимает следующий вид:

$$\forall \bar{s}(0) \in S_n \exists k \in [0 : t] (C(k)A(k-1) \dots A(0)\bar{s}(0) \neq [0]) \vee (A(t) \dots A(0) = [0])$$

Следствие. Если для НЛА \tilde{A} существует хотя бы одна УП длины t , то для него установочной является любая входная последовательность длины t и более.

Справедливость этого утверждения вытекает из того, что предикат (11.13) не зависит от входной последовательности.

Поскольку СП есть частный случай УП, то в общем случае длина минимальной УП для НЛА есть величина, не ограниченная сверху. Что касается периодического НЛА \tilde{A} , то для оценки длины минимальной УП справедлив аналог теоремы 1.14, т. е. эта длина не превосходит величины $n\lambda$, где n - размерность НЛА, λ - период матрицы $A(t)$.

Обратимся теперь к исследованию условия существования для НЛА диагностической последовательности.

Определение ДП для НЛА можно представить так: последовательность

$\bar{u}(0), \bar{u}(1), \dots, \bar{u}(t)$ является ДП для НЛА \tilde{A} , если

$$\begin{aligned} \forall \bar{s}_1(0), \bar{s}_2(0) \in (\tilde{A}) \\ (\wedge_{k=0}^t y_1(t) = y_2(t) \rightarrow s_1(0) = s_2(0)) \end{aligned}$$

Используемые здесь обозначения совпадают с теми, что были приведены выше в определении УП.

По аналогии со стационарным ЛА введем в рассмотрение следующую матрицу, которую будем называть диагностической матрицей для НЛА:

$$K_t = \begin{bmatrix} C(0) \\ C(1)A(0) \\ C(2)A(1)A(0) \\ \dots\dots\dots \\ C(9t)A(t-1)A(t-2)\dots a(0) \end{bmatrix}$$

Теорема 11.13. Для того чтобы для НЛА \tilde{A} размерности n , у которого $(\tilde{A}) = S_n$, входная последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(t)$ являлась ДП, необходимо и достаточно, чтобы $K_t = n$.

Доказательство. В терминах характеристических матриц приведенное только что определение ДП запишется следующим образом:

$$\begin{aligned} \forall \bar{s}_1(0), \bar{s}_2(0) \in S_n \\ (\wedge_{k=0}^t (C(k)A(k-1)\dots A(0)\bar{s}_1(0) + \sum_{i=0}^{k-1} C(k)A(k-1)\dots A(i+1)B(i)\bar{u}(i) + \\ D(k) = \\ = C(k)A(k-1)\dots A(0)\bar{s}_2(0) + \sum_{i=0}^{k-1} C(k)A(k-1)\dots A(i+1)B(i)\bar{u}(i) + D(k)\bar{u}(k))) \rightarrow \\ \rightarrow \bar{s}_1(0) = \bar{s}_2(0) \end{aligned}$$

Выполнив простые преобразования и сокращения, в результате получим

$$\begin{aligned} \forall \bar{s}_1(0), \bar{s}_2(0) \in S_n \\ (\wedge_{k=0}^t (C(k)A(k-1)\dots A(0)\bar{s}_1(0) = \wedge_{k=0}^t C(k)A(k-1)\dots A(0)\bar{s}_2(0))) \rightarrow \bar{s}_1(0) = \\ \bar{s}_2(0) \end{aligned}$$

Обозначив разность $\bar{s}_1(0) - \bar{s}_2(0)$ через $\bar{s}(0)$, последний предикат можно переписать в следующем виде:

$$\forall \bar{s}_1(0), \bar{s}_2(0) \in S_n$$

или, что все равно,

$$\forall \bar{s}(0) \in S_n \\ K_t \bar{s}(0) = [0]$$

Последнее соотношение, стоящее под знаком квантора общности, можно трактовать как систему линейных однородных алгебраических уравнений относительно координат вектора $\bar{s}(0)$. Существование ДП для НЛА равносильно тому, что соответствующая система имеет единственное решение. Из алгебры известно, что необходимым и достаточным условием для этого является выполнение равенства $K_t = n$.

Следствие. Если для НЛА существует ДП длины t , то для него диагностической является любая входная последовательность длины t и более. Справедливость этого утверждения вытекает из того, что условие теоремы 11.13 не зависит от входной последовательности.

Понятно, что всякая ДП для НЛА одновременно является и УП для этого автомата. Вместе с тем, в общем случае не всякая УП для НЛА одновременно является и ДП. По этой причине все сказанное выше о верхней границе длины минимальной УП в полной мере относится и к верхней границе ДП.

Подведем некоторые итоги исследования экспериментов как для стационарных, так и нестационарных автоматов.

Представленные в первых двух разделах лекции результаты свидетельствуют о том, что специфика линейных автоматов существенно упрощает построение теории экспериментов для них. Так, эта специфика дает возможность значительно понизить верхние оценки длин минимальных экспериментов всех типов по сравнению с соответствующими оценками, известными для автоматов (в общем случае нелинейных) Мили. Кроме того, эта специфика позволяет свести

задачу построения рассмотренных экспериментов, в общем случае весьма сложную и трудоемкую, к значительно более простой задаче установления факта существования таких экспериментов. Решение же последней задачи требует лишь вычисления произведения некоторых характеристических матриц, либо степеней матриц и их рангов. Иными словами, условия существования экспериментов исследованных нами типов для линейных автоматов достаточно легко проверяются. Отметим еще одно важное обстоятельство: идентификация финальных и начальных состояний после проведения соответствующих типов экспериментов в случае линейных автоматов сводится к решению систем линейных алгебраических уравнений, для чего имеется хорошо разработанный математический аппарат.

Вопросы и упражнения

1. Сформулируйте критерии существования установочной последовательности для линейного стационарного автомата.
2. При каких ограничениях на характеристическую матрицу S линейного стационарного автомата все входные последовательности некоторой фиксированной длины являются для него установочными?
3. Какова верхняя оценка длины минимальной установочной последовательности для стационарного линейного автомата?
4. Сформулируйте критерии существования диагностической последовательности для линейного стационарного автомата.
5. Какова верхняя оценка длины минимальной диагностической последовательности для стационарного линейного автомата?
6. Является ли условие минимальности стационарного ЛА достаточным для существования у него диагностической последовательности?
7. Приведите доказательство того, что для стационарного ЛА с помощью простого безусловного эксперимента разрешима любая диагностическая задача.
8. Дайте определение периодического линейного нестационарного автомата.
9. Сформулируйте критерии существования синхронизирующей, установочной и диагностической последовательностей для периодического линейного нестационарного автомата.

Эксперименты в пространстве обобщенных состояний и с линейными автоматами с запаздыванием

Введены понятия обобщенного состояния линейного автомата, обобщенных синхронизирующих, установочных и диагностических последовательностей. Для всех таких последовательностей сформулированы критерии их существования. Рассмотрены различные типы автоматов с запаздыванием (обыкновенные, по управлению, по состоянию), и для них получены критерии существования всех типов упомянутых выше последовательностей.

Эксперименты с линейными автоматами в пространстве обобщенных состояний

Наряду с традиционным понятием состояния ЛА \tilde{A} размерности n , представляющего собой n -мерный вектор, введем понятие обобщенного состояния. С этой целью зафиксируем некоторое целое положительное число μ , такое, что $1 \leq \mu n$, и вектор-столбец

$$s(t) = [s_1(t), \dots, s_\mu, \underbrace{x, \dots, x}_{n-\mu\mu}, \dots] \quad (12.1)$$

где x - символ неопределенности, будем называть обобщенным состоянием (ОС) линейного автомата. Предполагается, что каждый символ неопределенности x может принимать любое значение из поля $GF(p)$. Таким образом, каждому ОС линейного автомата соответствует некоторое подмножество S_n всех состояний ЛА. Далее через \tilde{S}_μ будем обозначать пространство обобщенных состояний ЛА. Понятно, что число обобщенных состояний в \tilde{S}_μ равно p^μ .

Вообще говоря, обобщенное состояние ЛА можно определить иначе, например, положив

$$\tilde{s}(t) = [s_1(t), \dots, s_{i_1-1}(t), \dots, x, s_{i_1+1}(t), \dots, s_{i_n-\mu-1}(t), x, s_{i_n-\mu+1}(t), \dots, s_n(t)]'$$

где $n - \mu$ штук значений координат $s_{i,j}(t)$, $j = 1, \dots, \mu$, номера которых выбраны произвольным образом, замещены символом неопределенности x . Легко показать, что путем подходящих перестановок строк и столбцов характеристических матриц линейный автомат \hat{A} всегда может быть преобразован в такой изоморфный ему ЛА \tilde{A} , у которого ОС (12.2) отображается при этом в ОС (12.1). Поэтому для простоты записи, но не теряя при этом общности, примем первый вариант (12.1) определения ОС.

Далее условимся о следующих обозначениях: если H есть матрица, то через $[H]_a$ обозначим матрицу с a штук строками, совпадающими с первыми a строками матрицы H . Аналогичное обозначение используем и для вектора: $[\bar{s}]_a$ - это вектор размерности a , у которого координаты совпадают с a первыми координатами вектора \bar{s} .

ЛА с заданной величиной μ , определяющей для ЛА пространство обобщенных состояний, условимся обозначать как μ -ЛА.

Введем для μ -ЛА понятие обобщенной диаграммы переходов (ОДП). ОДП для μ -ЛА над полем $GF(p)$ назовем ориентированный граф, содержащий p^μ вершин, взаимно однозначно соответствующих каждому ОС этого автомата. Две вершины \bar{a} и \bar{b} в ОДП соединяются дугой, ведущей из \bar{a} в \bar{b} , помеченной входным символом u μ -ЛА, если существуют по крайней мере два таких "традиционных" состояния \bar{s} и $\bar{\xi}$ из S_n , что $[s]_\mu = \bar{a}$, $[\xi]_\mu = \bar{b}$, $\bar{\xi} = A\bar{s} + B\bar{u}$.

Отметим различие между "традиционной" диаграммой переходов ЛА [19] и ОДП: в "традиционной" диаграмме каждой ее дуге, помеченной входным символом \bar{u} , соответствует единственный выходной символ (реакция ЛА на вход \bar{u} в состоянии \bar{s}), а в ОДП некоторой дуге может соответствовать несколько выходных символов. Назовем ОДП сильно связной, если в ней для любой пары вершин \bar{a} и \bar{b} существует путь, ведущий из \bar{a} в \bar{b} . Соответствующий этой ОДП μ -ЛА будем называть μ -сильно связным.

Понятно, что любому пути в ОДП из \bar{a} в \bar{b} соответствует некоторое

входное слово μ -ЛА, возможно не одно, переводящее его из ОС \bar{a} в ОС \bar{b} .

По аналогии с [19] μ -ЛА назовем k -управляемым, если для любых двух его обобщенных состояний \bar{a} и \bar{b} существует входное слово длины k , переводящее его из \bar{a} в \bar{b} . Очевидно, что k -управляемый μ -ЛА является μ -сильно связным.

Теорема 12.1. μ -ЛА является μ -управляемым тогда и только тогда, когда ранг матрицы

$$[L_k]_\mu = [A^{k-1}B, A^{k-2}B, \dots, AB, B]_\mu$$

равен μ .

Доказательство. Пусть $[\bar{s}]_\mu$ и $[\tilde{s}]_\mu$ - различные произвольные ОС рассматриваемого μ -ЛА. По определению k -управляемости существует такая входная последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ длины k , которая переводит μ -ЛА из $[\bar{s}]_\mu$ в $[\tilde{s}]_\mu$. Тогда на основании формулы (1.3) можно записать:

$$[\bar{s}]_\mu = [A^k \bar{s} + A^{k-1}B\bar{u}(0) + \dots + AB\bar{u}(k-2) + B\bar{u}(k-1)]_\mu$$

Отсюда следует, что

$$[\bar{s} - A^k \bar{s}]_\mu = [A^{k-1}B\bar{u}(0) + \dots + AB\bar{u}(k-2) + B\bar{u}(k-1)]_\mu = [L_k[\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)]]'_\mu$$

Поскольку $\bar{s} - A^k \bar{s}$ может быть любым n -мерным вектором, $[\bar{s} - A^k \bar{s}]_\mu$ также может быть любым μ -мерным вектором. В силу этого последнее равенство можно записать в виде

$$[L_k \hat{u}]_\mu \quad (12.3)$$

где $\hat{u} = [\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)]'$, и интерпретировать его как СЛАУ относительно неизвестных, являющихся координатами вектора

\hat{u} .

Понятно, что существование входной последовательности, переводящей μ -ЛА из произвольного ОС $[\tilde{s}]_\mu$ в произвольное ОС $[\tilde{s}]_\mu$, эквивалентно существованию решений выписанной выше системы при любых правых частях. Из алгебры известно [33], что необходимым и достаточным условием для этого является равенство ранга системы (12.3) величине μ . Заметим, что при $\mu = n$, где n - размерность ЛА, μ -сильно связность совпадает с общепринятым понятием сильно связности ЛА. Очевидно, что если ЛА размерности n является сильно связным, то он будет μ -сильно связным при любом $\mu < n$. Обратное, однако, в общем случае неверно. Действительно, рассмотрим ЛА над полем $GF(2)$, заданный характеристическими матрицами

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$$

Путем вычислений можно убедиться, что

$$L_3 = [A^2 B, AB, B] = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

и $L_3 = 2$. Тогда по теореме 12.1 этот ЛА не является сильно связным.

Положим $\mu = 2$, тогда

$$[L_3]_2 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

и $\text{rank}[L_3]_2 = 2$. По теореме 12.1 рассматриваемый μ -ЛА является 2-сильно связным. Таким образом, исходный ЛА не является сильно связным, однако в пространстве обобщенных состояний он становится 2-сильно связным.

Определим теперь понятие синхронизирующей последовательности

для μ -ЛА в пространстве обобщенных состояний.

Входную последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ μ -ЛА назовем обобщенной СП (ОСП), если

$$\forall \bar{s}, \bar{s} \in \tilde{S}_\mu [A^{k+1}\bar{s} + A^k B\bar{u}(0) + \dots + B\bar{u}(k)]_\mu = [A^{k+1}\bar{s} + A^k B\bar{u}(0) + \dots + B\bar{u}(k)]_\mu$$

Понятно, что при $\mu = n$, где n - размерность ЛА, приведенное определение совпадает с определением обычной СП для этого же автомата. Переносим в (12.4) правую часть равенства влево и выполнив сокращения, получим

$$\forall \bar{s}, \bar{s} \in S_\mu [A^{k+1}\bar{s} - A^k\bar{s}]_\mu = [0]_\mu \quad (12.5)$$

где $[0] =_\mu$ - нулевой вектор размерности μ . Поскольку в (12.5) \bar{s} и \bar{s} могут быть любыми ОС из \tilde{S}_μ , то (12.5) эквивалентно

$$\forall \bar{s} \in S_n [A^{k+1}\bar{s}]_\mu = [0]_\mu \quad (12.6)$$

Из предиката (12.6) вытекает справедливость следующего утверждения.

Теорема 12.2. Для того чтобы μ -ЛА \tilde{A} имела ОСП длины $k + 1$, необходимо и достаточно, чтобы

$$[A^{k+1}]_\mu = [0]_\mu$$

Как видим, эта теорема является аналогом теоремы 1.1 для обычного ЛА, т. е. ее естественным обобщением.

По аналогии с разделом 1.2 μ -ЛА условимся называть обобщенно синхронизируемым, если для него существует ОСП.

Рассмотрим ЛА над полем $GF(2)$, заданный следующими характеристическими матрицами:

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Поскольку $A \neq [0]$, то для этого ЛА, как это следует из теоремы 1.1, СП длины 1 не существует. Положим $\mu = 1$ и рассмотрим соответствующий 1-ЛА. Понятно, что состояние этого автомата в момент времени $t + 1$ не зависит от состояния в момент t , так как B есть нулевая матрица. Более того, для этого ЛА

$$s(t + 1) = [0s_2(t)s_2(t)]'$$

Из последнего равенства следует, что

$$[s(t + 1)]_1 = [0]_1$$

Последнее означает: любая входная последовательность длины 1 переводит 1-ЛА в одно и то же состояние $[0]_1$, т. е. любая такая последовательность является ОСП.

Таким образом, нами построен пример ЛА, не имеющего обычной СП некоторой фиксированной длины, но в то же время имеющего обобщенную СП такой длины.

Почти дословное повторение доказательства теоремы 1.2 позволяет установить справедливость ее аналога для μ -ЛА.

Теорема 12.3. Если для μ -ЛА существует хотя бы одна ОСП длины k , то обобщенными СП для него являются любые входные последовательности длины k и более.

По аналогии с обобщенными СП введем понятия обобщенных УП и ДП для μ -ЛА \tilde{A} , полагая, что множеством его допустимых начальных состояний является множество, которое мы обозначим как $(\tilde{A}0$.

Входную последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ μ -ЛА A назовем обобщенной УП (ОУП), если

$$\begin{aligned}
& \forall \bar{s}(0), \tilde{s}(0) \in \mathit{Init}(\tilde{A}) \wedge_{d=0}^k [CA^d \bar{s}(0) + CA^{d-1} B\bar{u}(0) + \dots + CB\bar{u}(d-1) + D\bar{u}(d) = \\
& = CA^d \tilde{s}(0) + CA^{d-1} B\tilde{u}(0) + \dots + CB\tilde{u}(d-1) + D\tilde{u}(d)] \rightarrow \\
& \rightarrow [A^{k+1} \bar{s}(0) + A^k B\bar{u}(0) + \dots + AB\bar{u}(k-1) + B\bar{u}(k)]_\mu = \\
& = [A^{k+1} \tilde{s}(0) + A^k B\tilde{u}(0) + \dots + AB\tilde{u}(k-1) + B\tilde{u}(k)]_\mu
\end{aligned}$$

Входную последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ μ -ЛА \tilde{A} назовем обобщенной ДП (ОДП), если

$$\begin{aligned}
& \forall \bar{s}(0), \tilde{s}(0) \in \mathit{Init}(\tilde{A}) \wedge_{d=0}^k [CA^d \bar{s}(0) + CA^{d-1} B\bar{u}(0) + \dots + CB\bar{u}(d-1) + D\bar{u}(d) = \\
& = CA^d \tilde{s}(0) + CA^{d-1} B\tilde{u}(0) + \dots + CB\tilde{u}(d-1) + D\tilde{u}(d)] \rightarrow [\bar{s}(0)]_\mu = [\tilde{s}(0)]_\mu
\end{aligned}$$

Заметим, что содержательный смысл обобщенных УП и ДП тот же, что и обычных, но применительно к обобщенным состояниям. Ниже приводятся теоремы, которые содержат естественные обобщения результатов для обычных ЛА на ЛА, рассматриваемые в пространстве обобщенных состояний. Доказательства некоторых из этих утверждений опущены, поскольку могут быть получены за счет незначительного изменения уже приводившихся ранее доказательств.

Теорема 12.4. Для того чтобы μ -ЛА имел обобщенную УП длины $k+1$, необходимо и достаточно, чтобы

$$\forall \bar{s} \in \mathit{Init}(\tilde{A}) (\exists k \geq 0 CA^k \bar{s} \neq [0]) \vee (\exists k \geq 0 [A^{k+1} \bar{s}]_\mu = [0]_\mu)$$

Здесь через $\Upsilon(\tilde{A})$ обозначено множество состояний, представляющее собой всевозможные разности состояний из множества (\tilde{A}) .

Теорема 12.5. Если у μ -ЛА \tilde{A} размерности n ранг характеристической матрицы C равен n , то для него существует обобщенная УП длины 1.

Доказательство. Пусть \bar{s} и \tilde{s} - два различных начальных состояния μ -ЛА A из множества допустимых. Через $\bar{y}(0)$ и $\tilde{y}(0)$ обозначим реакции этого автомата на входной символ $\bar{u}(0)$. Пусть $[\bar{s}_k]_\mu$ и $[\tilde{s}_k]_\mu$ - состояния μ -ЛА после подачи $\bar{u}(0)$. Если $\bar{u}(0)$ - обобщенная УП, то это означает, что

$$\forall \bar{s}, \tilde{s} \in (\tilde{A})\bar{y}(0) = \tilde{y}(0) \rightarrow [\bar{s}_k]_\mu = [\tilde{s}_k]_\mu$$

Предположим, что $\bar{u}(0)$ не является обобщенной УП, тогда

$$\exists \bar{s}, \tilde{s} \in (\tilde{A})\bar{y}(0) = \tilde{y}(0) \wedge [\bar{s}_k]_\mu \neq [\tilde{s}_k]_\mu$$

что эквивалентно

$$\exists \bar{s}, \tilde{s} \in (\tilde{A})C(\bar{s} - \tilde{s}) = [0] \wedge [\bar{s}_k]_\mu \neq [\tilde{s}_k]_\mu \quad (12.7)$$

Первый множитель в (12.7) можно интерпретировать как систему линейных однородных уравнений относительно n неизвестных, являющихся координатами вектор-состояния. Эта система должна иметь ненулевое решение, поскольку по предположению $\bar{s} \neq \tilde{s}$. Из алгебры известно, что необходимым и достаточным условием для этого является выполнение неравенства

$$C < n$$

что противоречит условию теоремы. Отсюда следует ложность предиката (12.7) и, следовательно, справедливость теоремы.

Теорема 12.6. Если для μ -ЛА размерности n существует хотя бы одна обобщенная УП длины k , то для этого автомата обобщенными УП являются любые входные последовательности длины k и более.

Введем следующие обозначения: через K_t^μ и $K_t^{(\mu)}$ соответственно обозначим подматрицы матрицы

$$K_t = [C, CA, \dots, CA^{t-1}]'$$

содержащие первые и соответственно последние μ ее столбцов.

Теорема 12.7. Для того чтобы μ -ЛА A размерности n имел обобщенную ДП длины t , необходимо и достаточно, чтобы ранг матрицы K_t^μ был равен μ и $K_t^{(\mu)} = [0]$.

Эта теорема является естественным обобщением теоремы 1.7. Два

последующих утверждения являются аналогами теорем 1.8 и 1.9.

Теорема 12.8. Если для μ -ЛА существует хотя бы одна обобщенная ДП длины k , то для него обобщенными являются любые входные последовательности длины k и более.

Теорема 12.9. Если для μ -ЛА размерности n существуют обобщенные ДП, то длина таких минимальных последовательностей не превосходит n .

Заметим, что аналоги последней теоремы справедливы как для обобщенных СП, так и для обобщенных УП.

Эксперименты с линейными автоматами с запаздыванием

Выше были рассмотрены ЛА, в правых частях уравнений переходов и выходов которых фигурировали векторы и матрицы со значениями только в момент времени t . Вместе с тем многие реальные процессы и устройства требуют для своего описания использования уравнений с запаздыванием.

Ниже будет использована классификация произвольных дискретных систем (ДС) с запаздыванием, предложенная в [17]. Очевидно, что эта классификация годится и для рассматриваемого нами частного случая ДС - линейных автоматов, заданных над полем $GF(p)$. В приведенных ранее обозначениях упомянутая классификация ДС такова:

Обыкновенная ДС - поведение ее описывается уравнением

$$\bar{s}(t + \alpha) + A_1 \bar{s}(t + \alpha - 1) + \dots + A_\alpha \bar{s}(t) = B_0 \bar{u}(t + \beta) + \dots + B_\beta \bar{u}(t), \beta < \alpha$$

Для того чтобы движение обыкновенной ДС было однозначно определено при $t \geq 0$, для любого входа $\bar{u}(t)$ должны быть заданы начальные условия $\bar{s}(0), \bar{s}(1), \dots, \bar{s}(\alpha - 1)$ и $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(\beta - 1)$

ДС с запаздыванием по состоянию - поведение ее описывается уравнением

$$\bar{s}(t+1) = A\bar{s}(t) + A_1\bar{s}(t-h) + B\bar{u}(t) \quad (12.9)$$

Для однозначного определения процесса $\bar{s}(t)$ при любом входе $\bar{u}(t)$ задаются начальные условия $\bar{s}(0), \bar{s}(-1), \dots, \bar{s}(-h)$, где h - некоторое положительное целое число.

ДС с запаздыванием по управлению - поведение ее описывается уравнением

$$\bar{s}(t+1) = A\bar{s}(t) + B\bar{u}(t) + B_1\bar{u}(t-h) \quad (12.10)$$

Для однозначности определения процесса $\bar{s}(t)$ при $t \geq -$ при любом входе $\bar{u}(t)$ должны быть заданы начальные условия $\bar{s}(0)$ и $\bar{u}(0), \bar{u}(-1), \dots, \bar{u}(-h)$, где h - некоторое положительное целое число.

В этом разделе будут рассмотрены вопросы существования СП, УП и ДП и методы их построения для линейных автоматов с запаздыванием, уравнения переходов которых представлены формулами (12.8)-(12.10).

Для получения критериев существования различных типов экспериментов для ЛА с запаздыванием используем следующую идею, изложенную в [71]. Попытаемся из фазового пространства состояний исходной ДС с запаздыванием по состоянию перейти в новое фазовое пространство состояний большей размерности с тем, чтобы уравнение состояний при этом трансформировалось в уравнение состояний ЛА без запаздывания. Аналогичное преобразование сделаем и в случае ЛА с запаздыванием по управлению, переходя от исходного входного вектора к вектору большей размерности с целью получения уравнений ЛА без запаздывания.

Из приведенных выше описаний ДС можно заключить, что участвующие в них векторы $\bar{s}(t)$, вообще говоря, не всегда представляют собой состояние системы в общепринятом смысле, т. е.

ту минимальную по объему информацию, которую необходимо знать в текущий момент времени t для однозначного продолжения процесса $\bar{s}(t)$ при известном входе $\bar{u}(t)$. Учитывая сказанное, каждую из перечисленных выше ДС с запаздыванием мы опишем в виде ЛА, определив соответствующим образом для ДС различных типов уравнение ее состояний, а затем для формулировки условий существования СП для этой ДС воспользуемся соответствующим критерием, полученным для ЛА без запаздывания.

Рассмотрим вначале обыкновенную ДС, заданную уравнением (12.8). Состояние $\bar{z}(t + \alpha)$ этой ДС в произвольный момент времени $t + \alpha$ определим как вектор $[\bar{s}(t), \bar{s}(t + 1), \dots, \bar{s}(t + \alpha + 1)]'$. Тогда состояние, в которое она переходит при подаче входного сигнала $\bar{u}(t + \beta)$, есть вектор $\bar{u}(t + \alpha + 1) = [\bar{s}(t + 1), \dots, \bar{s}(t + \alpha)]'$. Таким образом, если эту ДС описывать в виде ЛА, то размерность такого ЛА есть величина $n = m\alpha$ а начальное состояние есть $\bar{z}(0) = [\bar{s}(0), \bar{s}(1), \dots, \bar{s}(\alpha + 1)]'$. В качестве входного вектора $\bar{u}(t + \beta)$ этой ДС теперь будем использовать вектор $\bar{u}(t + \beta) = [\bar{u}(t), \bar{u}(t + 1), \dots, \bar{u}(t + \beta)]'$. С учетом изложенного, если описывать рассматриваемую ДС в виде ЛА, уравнение ее переходов состояний примет вид

$$\bar{z}(t + \alpha + 1) = A * \bar{z}(t + \alpha) + B * \bar{u}(t + \beta) \quad (12.11)$$

Матрицы $A*$ и $B*$ имеют блочную структуру, которая описывается ниже. В роли блоков в матрице $A*$ выступают матрицы размерности $m \times m$, а число "блочных" строк и столбцов ее равно величине α . Структура этой матрицы имеет следующий вид (E_m - единичная матрица размером $m \times m$):

$$A* = \begin{bmatrix} [0] & E_m & [0] & \dots & [0] \\ [0] & [0] & E_m & \dots & [0] \\ \dots & a \dots & \dots & \dots & \dots \\ [0] & [0] & [0] & \dots E_m & \\ \alpha & -A_{\alpha-1} & -A_{\alpha-2} & \dots - A_1 & \end{bmatrix}$$

В роли блоков матрицы B^* выступают матрицы размерности $m \times l$. Число "блочных" строк в B^* равно α , а число "блочных" столбцов - $\beta + 1$. Структура этой матрицы такова:

$$B^* = \begin{bmatrix} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \dots & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \dots & \dots & \dots & \dots & \dots \\ \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \dots & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ B_\beta & b_{\beta-1} & B_{\beta-2} & \dots & B_0 \end{bmatrix}$$

Непосредственной проверкой можно убедиться, что уравнение (12.11) с такими характеристическими матрицами задает линейный автомат, эквивалентный, с точки зрения функционирования, обыкновенной ДС, описанной разностным уравнением (1.1), откуда следует справедливость следующего утверждения:

Теорема 1.1 Для того чтобы ЛА с запаздыванием, описанный уравнением (12.8), имел СП длины k , необходимо и достаточно, чтобы выполнялось условие $(A^*)^k$.

Рассмотрим теперь ДС с запаздыванием по состоянию. Состояние ее в произвольный момент времени t определим как вектор $\bar{z}(t) = [\bar{s}(t-h), \bar{s}(t-h+1), \dots, \bar{s}(t)]'$ Тогда следующее состояние после подачи очередного входа будет $\bar{z}(t+1) = [\bar{s}(t)(h+1), \dots, \bar{s}(t), \bar{s}(t+1)]'$ Следовательно, при описании этой ДС в виде ЛА размерность последнего равна величине $n = m(h+1)$, а начальное состояние есть $\bar{z}(0) = [\bar{s}(0), \bar{s}(-1), \dots, \bar{s}(-h)]$, компонентами которого являются заданные начальные условия. Учитывая сказанное, при описании ДС с запаздыванием по состоянию в виде ЛА уравнение переходов состояний примет вид

$$\bar{z}(t+1) = A^* \bar{z}(t) + B^* \bar{u}(t)$$

где матрицы A^* и B^* так же, как и в предыдущем случае, имеют блочную структуру. В роли блоков в матрице A^* выступают матрицы

размерности $m \times m$, а число ее "блочных" строк и столбцов равно $h + 1$:

$$A^* = \begin{bmatrix} [0] & E_m & [0] & \dots & [0] \\ [0] & [0] & E_m & \dots & [0] \\ \dots & \dots & \dots & \dots & \dots \\ [0] & [0] & [0] & \dots & E_m \\ A_1 & [0] & [0] & \dots & A \end{bmatrix}$$

В роли блоков в матрице B^* выступают матрицы размерности $m \times l$, "блочных" строк в ней одна, а число столбцов - $h + 1$:

$$B^* = [[0] \dots [0]B]'$$

Можно проверить, что уравнение (10.1) с приведенными характеристическими матрицами задают ЛА, эквивалентный, с точки зрения функционирования, ДС с запаздыванием по состоянию, описанную уравнением (12.9). Отсюда следует, что для ДС такого типа справедлив аналог теоремы 10.1.

Рассмотрим, наконец, ДС с запаздыванием по управлению. При описании в виде ЛА состояние последнего в момент времени t отождествим с вектором $\bar{s}(t)$ из (12.10), а входным вектором ЛА теперь будет являться вектор

$$\bar{u}(t) = [\bar{u}(t), \bar{u}(t-1), \dots, \bar{u}(t-h)]'$$

Тогда очевидно, что основная характеристическая матрица соответствующего ЛА просто совпадает с матрицей A из (1.1), а в роли матрицы B в формуле (1.1) будет выступать "блочный" вектор

$$B^* = [B[0] \dots [0]B_1]'$$

Из изложенного следует, что условие существования СП для ДС с запаздыванием по управлению просто совпадает с соответствующим условием для ЛА, описанного уравнением (1.1).

Подводя итог, отметим, что каждой из перечисленных выше типов ДС

можно поставить в соответствие такой ЛА без запаздывания, который эквивалентен соответствующей ДС по поведению. Отсюда следует, что ранее доказанные для ЛА без запаздывания утверждения оказываются справедливыми для ЛА с запаздыванием. В частности, для них имеют место аналоги теорем 1.2, 1.3, следствие из теоремы 1.1 и т. д.

Кроме того, для решения задачи о переводе ЛА с запаздыванием в заданное синхросостояние остается пригодным метод, описанный в разделе 1.2 лекции 1. Проиллюстрируем это на примере применительно к обыкновенной ДС над полем $GF(2)$ при $\alpha = 1$ и $\beta = 1$, заданной уравнением

$$\bar{s}(t+2) + A_1 \bar{s}(t+1) + A_2 \bar{s}(t) = B_0 \bar{u}(t+1) + B_1 \bar{u}(t)$$

где $\bar{s}(t) = [s_1(t), s_2(t)]'$ и $\bar{u}(t) = [u_1(t)]'$, т. е. $m = 2, l = 1$, а характеристические матрицы ДС имеют следующий вид:

$$A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, A_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, B_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, B_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Вычисления показывают, что эквивалентный рассматриваемой системе ЛА имеет следующие характеристические матрицы:

$$A^* = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, B^* = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

При этом состоянием ЛА и его входным вектором являются векторы:

$$\bar{s}(t) = [s_1(t), s_2(t), s_1(t+1), s_2(t+1)]', \dots, \bar{u}(t) = [u_1(t), u_1(t+1)]'$$

Пусть задано синхросостояние $\bar{s} = [0, 1, 0, 1]'$ и требуется найти такую входную последовательность, которая переводит ДС из произвольного начального состояния в синхросостояние \bar{s} .

Легко проверить, что $(A^*)^3 = [0]$ откуда вытекает, что эквивалентный

рассматриваемой системе ЛА синхронизируем и длина минимальной СП равна 3. Для нахождения искомой СП выпишем матричное уравнение вида (1.14) для нашего примера:

$$(A^*)^2 B * \tilde{u}(0) + A * B * \tilde{u}(1) + B * \tilde{u}(2) = \tilde{s}$$

С учетом приведенных выше матриц A^* и B^* это уравнение принимает следующий вид:

$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} u_1(0) \\ u_1(1) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \times \begin{bmatrix} u_1(1) \\ u_2(1) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \times \begin{bmatrix} u_1(2) \\ u_1(3) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ - \\ 1 \end{bmatrix}$$

Перейдем от этого уравнения к равносильной системе, записанной в координатной форме:

$$u_1(1) + u_1(1) = 0, u_1(2) = 1, u_1(1) + u_1(2) = 0, u_1(3) = 1,$$

где "+" - операция по модулю 2.

Легко проверить, что эта система совместна и ее решение таково:

$$u_1(0) = x, u_1(1) = 1, u_1(2) = 1, u_1(3) = 0,$$

где x означает безразличное значение (0 или 1). Следовательно, для рассматриваемой ДС в качестве СП, переводящей ее из произвольного начального состояния в заданное синхросостояние \tilde{s} , могут быть использованы две следующие входных последовательностей:

1. $u_1(0) = 1, u_1(1) = 1, u_1(2) = 1, u_1(3) = 0,$
2. $u_1(0) = 0, u_1(1) = 1, u_1(2) = 1, u_1(3) = 0.$

Коснемся теперь критериев существования УП и ДП для ДС с запаздыванием.

Поскольку, как показано выше, каждый из рассматриваемых типов ДС соответствует эквивалентной ей ЛА без запаздывания, то критерий

существования УП и ДП для ДС с запаздыванием - это теорема 1.4 и 2.1, в которых вместо матриц A и C должны фигурировать матрицы A^* и C^* эквивалентных ЛА. Кроме того, для ДС с запаздыванием оказываются справедливыми аналоги теорем 10.5, 11.1 и 11.3.

Вопросы и упражнения

1. Дайте определение обобщенного состояния автомата.
2. Укажите различие между классической диаграммой автомата и диаграммой автомата в пространстве обобщенных состояний.
3. Сформулируйте критерий управляемости обобщенного автомата в матричной форме.
4. Может ли автомат, не являющийся сильно связным в классическом понимании, оказаться сильно связным в пространстве обобщенных состояний?
5. Дайте определения обобщенных синхронизирующих, установочных и диагностических последовательностей.
6. Сформулируйте критерии существования таких последовательностей.
7. Являются ли свойства обобщенных последовательностей аналогами соответствующих свойств обычных последовательностей тех же типов для "классических" линейных автоматов?
8. Приведите классификацию линейных автоматов с запаздыванием.
9. Изложите основополагающую идею для получения критериев существования всех типов последовательностей для всех типов автоматов с запаздыванием.

Синхронизация и устойчивость дискретных линейных систем

Для состояния линейного автомата определяются понятия равновесия и асимптотической устойчивости. Приводится критерий существования у свободного ЛА асимптотически устойчивого состояния. Рассмотрена задача стабилизации ЛА и проблема ее разрешимости. Для дискретных линейных систем над полем R введено понятие ϵ -синхронизирующей последовательности и приводится критерий ее существования.

Понятие устойчивости систем и связанные с ним другие понятия первоначально были введены и исследованы для непрерывных систем [40], [41] и позже были распространены на класс дискретных линейных систем (ДЛС) [64]. Что касается ДЛС, то их задание возможно над различными полями: либо над полем $GF(p)$, как это сделано в случае линейных автоматов, либо над полем R вещественных чисел, либо над полем Z комплексных чисел. Ниже основное внимание будет сосредоточено на линейных автоматах, заданных над полем $GF(p)$, и ДЛС, заданных над полем R .

Напомним предварительно некоторые определения из [64], которые понадобятся в дальнейшем. Поскольку понятие устойчивости системы связано только с ее движением в фазовом пространстве состояний, то в описании систем нас будет интересовать только уравнение, определяющее траекторию систем.

Определение 13.1. ЛА \tilde{A} , заданный над полем $GF(p)$ уравнением

$$\bar{s}(t+1) = A\bar{s}(t) + B\bar{u}(t)$$

называется свободным, если $\bar{u}(t) = [0]$ для любого t .

Определение 13.2. Состояние \bar{s} свободного ЛА \tilde{A} называется состоянием равновесия, если для любого t

$$\bar{s} = \bar{s}(t)$$

Из этих определений следует, что состояние равновесия свободного ЛА должно удовлетворять равенству

$$\bar{s} = A\bar{s} \quad (13.1)$$

Из (13.1) вытекает справедливость следующего утверждения.

Теорема 13.1. Нулевое состояние является состоянием равновесия любого свободного ЛА.

Рассматривая (13.1) как СЛАУ относительно неизвестных s_1, \dots, s_n , являющихся компонентами вектора \bar{s} , нахождение состояния равновесия ЛА сводится к решению этой системы.

Понятно, что число состояний равновесия ЛА равно числу решений системы (13.1). Если определитель $|A - E|$ этой системы равен 0, где E - единичная матрица, то ЛА имеет единственное (нулевое) состояние равновесия, в противном случае их существует конечное число, но больше одного.

В качестве примера рассмотрим ЛА над полем $GF(2)$ со следующей главной характеристической матрицей

$$A = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Система (13.1) в координатной форме в данном случае примет вид

$$\begin{aligned} s_1 + s_2 + s_3 + s_4 &= 0, \\ s_1 + s_2 + s_3 + s_4 &= 0, \\ s_3 + s_4 &= 0, \\ s_3 + s_4 &= 0 \end{aligned} \quad (13.2)$$

Здесь операция "+" понимается как сложение по модулю 2.

Поскольку ранг матрицы A равен 2, то система (13.2) имеет только два

линейно независимых уравнения. Выберем в качестве таковых первое и третье уравнения. Считая переменные s_2 и s_4 свободными, выразим через них остальные переменные:

$$s_1 = s_2 + s_3 + s_4 = s_2,$$

$$s_3 = s_4..$$

Придавая значения 0 и 1 свободным переменным, получаем искомые состояния равновесия:

$$[0, 0, 0, 0]', [0, 0, 1, 1]', [1, 1, 0, 0]', [1, 1, 1, 1]'$$

Определение 13.3. Состояние равновесия \bar{s} свободного ЛА \tilde{A} назовем асимптотически устойчивым, если

$$\forall \bar{s} \in S_n \exists k \in \mathbb{N} \forall t > k \rightarrow \bar{s}(t) = \bar{s}$$

Через \mathbb{N} здесь обозначено множество натуральных чисел.

Содержательно последнее определение означает, что после отклонения ЛА от состояния равновесия он возвращается в это состояние после подачи на вход нулевой последовательности некоторой подходящей длины.

В рассмотренном выше примере ЛА ни одно из четырех состояний равновесия не является асимптотически устойчивым. В самом деле, вычисления показывают, что если этот ЛА стартует в состояниях

$[1, 0, 0, 0]'$ или $[0, 1, 0, 0]'$, то при нулевой входной последовательности эти состояния переходят только друг в друга.

Аналогичная ситуация имеет место и с состояниями $[0, 1, 1, 1]'$ и $[1, 0, 1, 1]'$, $[0, 0, 0, 1]'$ и $[1, 1, 1, 0]'$, $[0, 0, 1, 0]'$ и $[1, 1, 1, 0]'$, $[0, 1, 0, 1]'$ и $[0, 1, 1, 0]'$, $[1, 0, 0, 1]'$ и $[1, 0, 1, 0]'$.

Найдем условие, при котором состояние равновесия является асимптотически устойчивым. Из определения 13.3 и формулы (1.3) следует, что для асимптотически устойчивого состояния равновесия \bar{s} должно существовать такое натуральное число k , что

$$\bar{s} = A^k \bar{s}(0) = A^k \tilde{s}(0)$$

где $\bar{s}(0)$, $\tilde{s}(0)$ - различные произвольные начальные состояния ЛА.

Вычитая одно равенство из другого, получим

$$A^k[\bar{s}(0) - \tilde{s}(0)] = [0]$$

В силу произвольности состояний $\bar{s}(0)$ и $\tilde{s}(0)$ их разность может быть любым ненулевым вектором \hat{s} . Поэтому последнее равенство примет вид

$$A^k \hat{s} = [0] \quad (13.3)$$

и оно должно выполняться для любого вектора $\hat{s} \in S_n$. Из этого факта вытекает, что однородная СЛАУ относительно координат вектора \hat{s} имеет не единственное решение и таковыми для нее являются любые ненулевые вектора. Последнее обстоятельство позволяет сформулировать два следующих утверждения.

Теорема 13.2. Если свободный ЛА над полем $GF(p)$ имеет асимптотически устойчивое состояние, то оно единственно и совпадает с нулевым.

Теорема 13.3. Для того чтобы свободный ЛА над полем $GF(p)$ имел асимптотически устойчивое состояние равновесия, необходимо и достаточно, чтобы существовало такое натуральное k , для которого

$$A^k = [0]$$

Из сопоставления последней теоремы с теоремой 1.1 вытекает справедливость следующего утверждения.

Теорема 13.4. Для того чтобы ЛА над полем $GF(p)$ имел асимптотически устойчивое состояние, необходимо и достаточно, чтобы он был синхронизируемым.

Напомним, что синхронизируемым называется такой ЛА, у которого существует СП.

Таким образом, для свободного ЛА над полем $GF(p)$ наличие асимптотически устойчивого состояния эквивалентно его синхронизируемости.

Попробуем провести параллель между асимптотической устойчивостью для ДЛС над полем R и асимптотической устойчивостью ЛА над полем $GF(p)$. С этой целью напомним некоторые понятия из [64].

Состояние \bar{s} свободной ДЛС асимптотически устойчиво, если

$$\lim_{t \rightarrow \infty} [\bar{s}(t) - \bar{s}] = [0]$$

Областью асимптотической устойчивости ДЛ называется множество M всех таких его состояний, что траектории ДЛС, начинающиеся в любом состоянии из M , при достаточно длинной нулевой входной последовательности заканчиваются в состоянии равновесия.

Состояние равновесия ДЛС устойчиво в большом, если существует только одно состояние равновесия и если областью асимптотической устойчивости является все пространство состояний.

ДЛС над полем R по существу можно считать аналогом линейного автомата над полем $GF(p)$. Используя для свободных ЛА над полем $GF(p)$ аналоги только что приведенных понятий, на основании теоремы 13.3 можно сделать вывод, что для них понятия асимптотической устойчивости и устойчивости в большом просто совпадают. Для ДЛС, как известно [64], это тоже справедливо.

Что касается устойчивости ДЛС над полем R , то она включает устойчивость по начальным условиям и устойчивость к внешним возмущениям.

Аналог этого понятия для ЛА введем следующим образом.

Определение 13.4. ЛА над полем $GF(p)$ назовем устойчивым, если в

случае, когда для любого t входной сигнал $\bar{u}(t) = [0]$,

$$\begin{aligned} \exists t_0 \in N \\ t \geq t_0 \rightarrow \bar{s}(t) = [0] \end{aligned}$$

Это определение по существу есть аналог устойчивости ЛА только по начальным условиям. Вторая же составляющая понятия устойчивости, устойчивость к внешним возмущениям, для ЛА над полем $GF(p)$ всегда имеет место, поскольку множества состояний и входов для ЛА являются ограниченными.

Как правило, для выполнения своих функций система должна постоянно находиться в некотором определенном состоянии фазового пространства, из которого она может быть выведена возникающими внешними возмущениями. В этом случае появляется необходимость в возвращении системы в упомянутое состояние всякий раз, когда она из него выводится. Такая задача, называемая задачей стабилизации, обычно решается путем организации в системе подходящей обратной связи.

Напомним математическую постановку задачи стабилизации: для ЛА \tilde{A} , траектория которого описывается уравнением (1.1) и который находится в начальном состоянии $\bar{s}(0)$, требуется выбрать обратную связь $\bar{u}(t) = -P\bar{s}(t)$ так, чтобы обеспечить устойчивость замкнутой системы

$$\bar{s}(t+1) = [A - BP]\bar{s}(t) \quad (13.4)$$

Для ДЛС, заданной над полем R , известно следующее утверждение [38] если пара матриц A, B является невырожденной, то всегда можно выбрать матрицу P коэффициентов обратной связи $\bar{u}(t) = -P\bar{s}(t)$ так, чтобы замкнутая система (13.4) была устойчивой.

Напомним, что пара матриц A, B называется невырожденной, если

$$\text{rank}(A, AB, A^2B, \dots, A^{n-1}B) = n$$

где n - размерность ДЛС.

Покажем, что в отличие от ДЛС над полем R задача стабилизации для ЛА над полем $GF(p)$ в общем случае не всегда разрешима.

Пусть ЛА над полем $GF(3)$ задан следующими характеристическими матрицами:

$$A = \begin{bmatrix} 0 & 2 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Для этого автомата $n = 4, l = 1$, следовательно, матрица P из (13.4) имеет вид

$$P = [p_1, p_2, p_3, p_4]$$

Выполнив вычисления по правилам поля $GF(3)$, построим матрицу

$$[A, AB, A^2B, A^3B] = \begin{bmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 2 & 2 & 1 \end{bmatrix}$$

Можно проверить, что определитель этой матрицы равен 1, т. е. отличен от нуля, тогда ее ранг равен 4. Таким образом, рассматриваемая пара матриц является невырожденной.

Легко сообразить, что для свободного ЛА над полем $GF(p)$ понятие устойчивости, введенное в определении 13.4, совпадает с понятием асимптотической устойчивости (см. определение 13.3 и теорему 13.2). Тогда в силу теорем 13.3 и 13.4 нам достаточно установить, что матрица

$[A - BP]$ для нашего примера не может быть сделана нильпотентной ни при каком выборе матрицы P , откуда и будет следовать неразрешимость задачи стабилизации.

Поскольку размерность ЛА в рассматриваемом примере равна 4, то для доказательства невозможности обращения в нулевую матрицы $[A - BP]^k$ при любом k достаточно установить, что $[A - BP]^4 \neq [0]$.

Вычисления показывают, что

$$[A - BP] = \begin{bmatrix} -p_1 & 2 - p_2 & -p_3 & -p_4 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ 2 & 0 & 1 & 1 \end{bmatrix},$$

$$[A - BP]^2 = \begin{bmatrix} p_1 - p_2 - 2p_4 + 2 & p_1 p_2 - 2p_1 - p_3 & p_1 p_3 - 2p_2 - p_3 - p_4 + 1 & p_1 p_2 - p_2 - p_4 + 2 \\ 2 - p_1 & 1 - p_2 & -p_3 & 1 - p_4 \\ 1 & 1 & 0 & 1 \\ 2 - 2p_1 & 2 - 2p_2 & 2 - 2p_3 & 1 - 2p_4 \end{bmatrix}$$

Из-за громоздкости мы здесь не приводим всю матрицу $[A - BP]^4$ (ее можно получить, возведя в квадрат последнюю матрицу), а выпишем лишь элементы ее третьей строки:

$$\begin{aligned} & p_1 - p_2 - 2p_4, \\ & p_1 p_2 - 2p_1 - p_3, \\ & p_1 p_3 - p_3 - 2p_2 - p_4, \\ & p_1 p_4 - p_2 - p_4 + 1. \end{aligned}$$

Приравняв все эти элементы нулю, получим нелинейную систему уравнений, из которой найдем p_i ($i = \overline{1, 4}$). Выполнив перебор, можно убедиться, что полученная система имеет единственное решение

$$p_1 = 0, p_2 = 2, p_3 = 0, p_4 = 2.$$

Понятно, что матрица $[A - BP]^4$ будет нулевой лишь в том случае, когда при приведенных значениях p_i в нуль обратятся все остальные элементы матрицы. Можно убедиться путем вычислений, что элемент (4,1) матрицы $[A - BP]^4$ таков:

$$(2 - p_1)(p_1^2 + 2 - p_2 - p_4) + (2 - 2p_2)(2 - p_1) + 2 - 2p_3 + (1 - 2p_4)(2 - 2p_1)$$

Подставив в это выражение приведенные выше значения p_i и выполнив вычисления по правилам поля $GF(3)$, получим значение 2.

Отсюда следует, что матрица $[A - BP]^4$ не может быть сделана нулевой за счет подходящего выбора матрицы P . Таким образом, для рассматриваемого ЛА над полем $GF(3)$ задача стабилизации с помощью обратных связей неразрешима.

Вместе с тем заметим, что если с помощью приведенных выше матриц A, B задать ДЛС над полем R , то вычисления в поле R дают следующую матрицу

$$[A, AB, A^2B, A^3B] = \begin{bmatrix} 1 & 0 & 2 & 4 \\ 0 & 1 & 2 & 6 \\ 0 & 0 & 1 & 3 \\ 0 & 2 & 2 & 7 \end{bmatrix}$$

Определитель этой матрицы равен 1, следовательно, ее ранг равен 4.

Таким образом, пара тех же матриц A, B над полем R также является невырожденной, но тогда, как это следует из приведенного выше утверждения из [38], задача стабилизации для соответствующей замкнутой системы над полем R оказывается разрешимой.

Понятие синхронизирующей последовательности, введенное в разделе 1.2 лекции 1 для ЛА над полем $GF(p)$, может быть распространено и на ДЛС, заданной над полем R .

Введем следующие обозначения. Пусть $\hat{u}(t) = \bar{u}(0), \bar{u}(1), \dots, \bar{y}(t)$ - входная последовательность длины $t + 1$, $\bar{s}^i(0)$ - начальное состояние ДЛС ($i = 1, 2, \dots$), $\bar{s}^f(\bar{s}^i(0), \hat{u}(t))$ - конечное состояние ДЛС после подачи на ее вход последовательности $\hat{u}(t)$, если она стартует из состояния $\bar{s}^{-i}(0)$.

Для ДЛС над полем R можно ввести понятие синхронизирующей последовательности, почти дословно повторив ее определение, приведенное в разделе 1.2. Легко доказать, что в этом случае для ДЛС справедлив аналог теоремы 1.1.

Введем теперь некоторую модификацию понятия СП.

Определение 13.5. Бесконечную последовательность $\{\tilde{u}(t)\} = \tilde{u}(0), \tilde{u}(1), \dots, \tilde{u}(t)$ назовем ε -синхронизирующей (ε -СП) для ДЛС над полем R с множеством допустимых начальных состояний $Init(\tilde{A})$, если

$$\forall \bar{s}^1(0), \bar{s}^2(0) \in Int(\tilde{A}) \forall \varepsilon > 0 \exists t_0 \in \mathbb{N} t > t_0 \rightarrow |\bar{s}^1(\bar{s}^1(0), \hat{u}(t)) - \bar{s}^2(\bar{s}^2(0), \hat{u}(t))| < \varepsilon,$$

где $\hat{u}(t)$ - начальный отрезок длины $t + 1$ входной последовательности $\{\tilde{u}(t)\}$.

Поясним содержательно различие между СП и ε -СП. После подачи СП (она имеет конечную длину) ДЛС оказывается в одном и том же известном конечном состоянии. В то же время после подачи начального отрезка ε -СП ДЛС переводится хотя и в различные конечные состояния, но расстояние между ними может быть сделано меньше любого заранее заданного числа ε за счет подходящего выбора длины упомянутого начального отрезка.

Теорема 13.5. Для того чтобы ДЛС \tilde{A} , у которой множество допустимых начальных состояний $Init(\tilde{A})$ ограничено, имела ε -СП, достаточно, чтобы все собственные числа λ_ν ее характеристической матрицы A лежали внутри единичного круга

$$|\lambda_\nu| < 1 \quad (13.5)$$

Доказательство. Проведем его для случая, когда все собственные числа матрицы A различны. Как следует из определения 13.5, существование ε -СП $\{\tilde{u}(t)\}$ эквивалентно тому, что

$$\lim_{t \rightarrow \infty} |\bar{s}^f(0), \hat{u}(t)) - \bar{s}^f(\bar{s}^2(0), \hat{u}(t))| = 0$$

В силу аналога формулы (1.3) для ДЛС этот предел эквивалентен

$$\lim_{t \rightarrow \infty} |A^{t+1}(\bar{s}^1(0) - \bar{s}^2(0))| = 0$$

По условию теоремы $|\bar{s}^1(0)| < M$ и $|\bar{s}^2(0)| > M$, где $M = const$, но тогда и для состояния $\bar{s}(0) = \bar{s}^1(0) - \bar{s}^2(0)$ выполняется неравенство $|\bar{s}^1(0) - \bar{s}^2(0)| > M$.

Рассмотрим ДЛС \tilde{A}_* , заданную уравнением переходов

$$\tilde{s}(t+1) = A\tilde{s}(t) \quad (13.6)$$

Очевидно, что существование ε -СП для ДЛС \tilde{A} эквивалентно выполнению следующего равенства для ДЛС \tilde{A}_* :

$$\lim_{t \rightarrow \infty} |\tilde{s}(t+1)| = \lim_{t \rightarrow \infty} |A^{t+1}\tilde{s}(0)| = 0$$

Введем замену $\tilde{s}(t) = Sz(t)$, где S - невырожденная матрица, столбцами которой являются собственные векторы матрицы A , причем

$$S^{-1}AS = \text{diag}\{\lambda_\nu\}$$

где $\text{diag}\{\lambda_\nu\}$ - диагональная матрица. Если $|\bar{s}(t)| < M$, то, очевидно, и $|\bar{z}(t)| < M_1$, где $M_1 = const$. Из уравнения (13.6) следует, что

$$z(t+1) = S^{-1}ASz(t)$$

или в координатной форме $z_\nu(t+1) = \lambda_\nu z_\nu(t)$.

Учитывая, что

$$z_\nu(t+1) = \lambda_\nu^t z_\nu(0) \rightarrow |z_\nu(t+1)| = |\lambda_\nu^t| |z_\nu(0)|$$

в силу условия (13.5) и того факта, что $|z_\nu(t)| < M_1$, получаем

$$\lim_{t \rightarrow \infty} |z_\nu(t+1)| = 0$$

Отсюда вытекает, что

$$\lim_{t \rightarrow \infty} |\bar{s}(t+1)| = \lim_{t \rightarrow \infty} |S(t+1)| = 0$$

а это, как было упомянуто выше, эквивалентно существованию ε -СП для ДЛС \tilde{A} .

Заметим, что эта теорема остается справедливой и в случае кратных собственных чисел, но доказательство ее немного усложняется.

Следствие. Если у ДЛС \tilde{A} над полем R все собственные числа характеристической матрицы A лежат внутри единичного круга, то любая бесконечная входная последовательность $\{\tilde{u}(t)\}$ является для нее ε -СП.

Справедливость следствия вытекает из того, что в условиях теоремы 13.5 входная последовательность $\{\tilde{u}(t)\}$ является произвольной.

Коснемся теперь связи между устойчивостью ДЛС над полем R и свойством ε -синхронизации. Напомним [38], что свободная ДЛС над полем R называется устойчивой по начальным условиям, если

$$\lim_{t \rightarrow \infty} \bar{s}(t) = [0]$$

ДЛС, для которой из условия $|\bar{u}(t)| \leq M = \text{const}$ следует ограниченность $|\bar{s}(t)|$, называется устойчивой к внешним возмущениям. При наличии обоих указанных свойств ДЛС называется устойчивой.

Понятно, что из устойчивости ДЛС по начальным условиям вытекает справедливость равенства

$$\lim_{t \rightarrow \infty} |\bar{s}(t)| = 0$$

Последнее означает, что бесконечная нулевая входная последовательность является для ДЛС ε -СП.

Известно [38], что условие (13.5) есть достаточное условие устойчивости. Таким образом, из теоремы 13.5 следует, что для свободной ДЛС в случае выполнения условия (13.5) понятие ε - синхронизации и устойчивости равносильны.

Из существования СП следует существование ε -СП (она состоит из двух подпоследовательностей, первая из которых совпадает с СП, а вторая, бесконечная, может выбираться произвольно), однако обратная импликация места не имеет.

Состояние ДЛС \tilde{A} , в которое она переходит после подачи на ее входы СП (ε -СП), назовем синхросостоянием. Ясно, что в общем случае различные СП (ε -СП) могут переводить ДЛС как в различные, так и совпадающие синхросостояния.

В разделе 1.2 лекции 1 была рассмотрена задача о переводе ЛА над полем $GF(p)$ в заданное синхросостояние с помощью СП. Точно такая же задача может быть сформулирована и для ДЛС над полем R , и метод ее решения аналогичен методу, предложенному для случая ЛА.

Сформулируем подобную задачу для ДЛС, когда в качестве "переводящей" последовательности используется ε -СП.

Пусть заданы ДЛС \tilde{A} размерности n над полем R , синхросостояние \bar{s} и величина ε . Требуется найти входную последовательность конечной длины (начальный отрезок ε -СП), которая переводит ДЛС в такое состояние \tilde{s} , что $|\tilde{s} - \bar{s}| < \varepsilon$.

Пусть множество $Init(\tilde{A})$ допустимых начальных состояний ДЛС \tilde{A} ограничено и для любого

$$\tilde{s} = [s_1^*, \dots, s_n^*] \in Init(\tilde{A})$$

выполняется неравенство

$$|\tilde{s}| \leq M \quad (13.7)$$

Очевидно, что если

$$|s_i^*| \leq \frac{M}{\sqrt{n}}, i = 1, \dots, n$$

то неравенство (13.7) заведомо будет выполнено. Теперь найдем такое минимальное число k , чтобы выполнялось неравенство

$$|A^k \tilde{S}^*| \leq \varepsilon$$

$$\text{где } \tilde{s}^* = \left[\frac{M}{\sqrt{n}}, \dots, \frac{M}{\sqrt{n}} \right].$$

Используя аналог формулы (1.3) для ДЛС \tilde{A} , выпишем следующее равенство:

$$A^{k-1} B \bar{u}(0) + A^{k-2} B \bar{u}(1) + \dots + A B \bar{u}(k-2) + B \bar{u}(k-1) = \bar{s} - A^k \bar{s}^*$$

Его можно рассматривать как СЛАУ относительно координат векторов входной последовательности $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$. Понятно, что решение этой системы будет давать искомую входную последовательность.

Проиллюстрируем изложенное на примере. Пусть ДЛС над полем R с одним входным каналом задана характеристическими матрицами

$$A = \begin{bmatrix} 0.01 & 0.05 & 0.00 \\ 0.05 & 0.001 & 0.02 \\ 0.00 & 0.02 & 0.0001 \end{bmatrix}, B = \begin{bmatrix} 2.0 \\ 3.0 \\ 4.0 \end{bmatrix}$$

Путем вычислений можно убедиться, что собственными значениями матрицы A являются числа $\lambda_1 = 0.01, \lambda_2 = 0.001, \lambda_3 = 0.0001$. Отсюда на основании теоремы 13.5 можно утверждать, что рассматриваемая ДЛС является синхронизируемой.

Пусть требуется перевести эту ДЛС в синхросостояние

$\bar{s} = [0.4, 1.2, -0.3]'$, и пусть точность, с которой это надо осуществить, равна $\varepsilon = 0.001$. Предположим, что для любого $\tilde{s} \in \text{Init}(\tilde{A})$ выполняется неравенство $|\tilde{s}| \leq 1.3$. Отсюда получаем, что $\tilde{s}^* = [0.75, 0.75, 0.75]'$.

Определим теперь минимальную начальную длину отрезка ε -СП, обеспечивающую перевод рассматриваемой ДЛС в заданное синхросостояние с точностью $\varepsilon = 0.001$. В соответствии с изложенным вычислим

$$A^2 = \begin{bmatrix} 0.0026 & 0.00055 & 0.001 \\ 0.00055 & 0.002901 & 0.000022 \\ 0.001 & 0.000022 & 0.0004 \end{bmatrix}, A^2 \tilde{s}^* = \begin{bmatrix} 0.00312 \\ 0.006312 \\ 0.001066 \end{bmatrix}, |A^2 \tilde{s}^*| = 0.007141$$

Поскольку $|A^2 \tilde{s}^*| > \varepsilon = 0.001$, вычислим

$$A_3 = \begin{bmatrix} 0.0054 & 0.000151 & 0.000011 \\ 0.0002 & 0.000031 & 0.00 \\ 0.00011 & 0.000058 & 0.00 \end{bmatrix}, A^3 \tilde{s}^* = \begin{bmatrix} 0.00016 \\ 0.000173 \\ 0.000052 \end{bmatrix}, |A^3 \tilde{s}^*| = 0.000242$$

Поскольку $|A^3 \tilde{s}^*| < \varepsilon = 0.001$, заданная точность достигнута. Таким образом, начальный отрезок искомой ε -СП, обеспечивающий перевод рассматриваемой ДЛС в синхросостояние $[0.4, 1.2 - 0.3]'$ с заданной точностью $\varepsilon = 0.001$, будет иметь длину $k = 3$. Исходя из этого, система (13.8) примет вид

$$A^2 B \bar{u}(0) + A B \bar{u}(1) + B \bar{u}(2) = \bar{s} - A^3 \bar{s}^*$$

В координатной форме эта система такова:

$$\begin{aligned} 0.01085u(0) + 0.17u(1) + 2.0u(2) &= 0.399839, \\ 0.00989u(0) + 0.1834u(1) + 3.0u(2) &= 1.119827, \\ 0.00366u(0) + 0.0604u(1) + 4.0u(2) &= -0.299048. \end{aligned}$$

Решая эту систему, получим следующий результат:

$$u(0) = -507.625899, u(1) = 36.675231, u(2) = -0.163412.$$

Итак, полученная входная последовательность переводит рассматриваемую ДЛС из любого состояния множества $Init(\tilde{A})$, состоящего из всевозможных трехмерных векторов, длины которых не превосходят величины $M = 1.3$, в состояния, отстоящие от заданного синхросостояния $\bar{s} = [0.4, 1.2, -0.3]'$ не более чем на 0.001.

Сделаем замечание, касающееся методов решения системы уравнений вида (13.8). Отметим, что правая часть этой системы является не точной, а приближенной величиной. Это же относится и к элементам различных степеней матриц и произведений матриц, поскольку при их вычислении производятся округления. Поэтому для решения системы (13.8) целесообразно пользоваться приближенными методами. Сейчас известен широкий спектр подобных методов [4], каждый из которых характеризуется различными условиями сходимости, и поэтому для каждой конкретной системы вида (13.8) необходимо подбирать такой, который гарантирует сходимость. Что касается практической реализации, то приближенные методы к тому же, как правило, менее трудоемки, чем точные.

Приведенные выше результаты относились к стационарным ЛА. Что касается нестационарных ЛА, то полные аналоги этих результатов оказываются справедливыми и для них.

Так, определив для НЛА состояние равновесия и асимптотически устойчивое состояние равновесия, как это делалось для стационарных ЛА (определения 13.2 и 13.3 соответственно), легко доказать, что теоремы 13.1 и 13.2 остаются справедливыми и для НЛА. Формулировка теоремы 13.3 для НЛА претерпевает незначительное изменение.

Теорема 13.3. Для того чтобы свободный НЛА \tilde{A} над полем $GF(p)$ имел асимптотически устойчивое состояние равновесия, необходимо и достаточно, чтобы существовало такое целое положительное t , что

$$A(t)A(t-1)\dots A(0) = [0]$$

Заметим, что для НЛА над полем $GF(p)$ асимптотическая устойчивость и устойчивость в большом просто совпадают, как это было и для стационарных ЛА.

По аналогии с определением 13.5 для стационарных ЛА можно ввести понятие ε -СП для нестационарного ЛА.

Среди всего множества НЛА рассмотрим далее такой класс НЛА, у которых характеристические матрицы $A(t)$ являются периодическими (с периодом λ), т. е. для них выполняется равенство $A(t + \lambda) = A(t)$, где λ - целое положительное число. Построим стационарный ЛА \tilde{A}_c , согласованный с НЛА \tilde{A} , задав функцию переходов ЛА \tilde{A}_c следующим образом:

$$\tilde{s}(t + 1) = \hat{A}\tilde{s}(t)$$

где $\hat{A} = A(\lambda - 1)A(\lambda - 2) \dots A(0)$. Напомним [68], что спектральным радиусом $\rho(A)$ матрицы A называют число $\rho(A) = \max\{|\lambda|; \lambda - \text{собственные значения матрицы } A\}$.

Теорема 13.6. Для того чтобы периодический НЛА \tilde{A} с ограниченным множеством $Init(\tilde{A})$ допустимых начальных состояний имел ε -СП, достаточно, чтобы при любом t выполнялось неравенство

$$\rho(A) < 1 \quad (13.9)$$

Доказательство. Как следует из определения ε -СП, ее существование равносильно выполнению неравенства

$$\lim_{t \rightarrow \infty} |s^f(0), \hat{u}(ty)) - \bar{s}^f(\bar{s}^2(0), \hat{u}(t)) = 0| \quad (13.10)$$

Здесь использованы те же обозначения, что и в определении 13.5. Если матрица $A(t)$ периодическая, то по индукции можно показать, что формула (1.7) примет вид

$$\bar{s}(t+1) = A(t - \mu[t \mu])A(t - \mu[t \mu] - 1) \dots A(0)\hat{A}\bar{s}(0) + A(t - \mu[t \mu]) \dots \\ \dots A(1)\hat{A}^{[t \mu]}B(0)\bar{u}(0) + \dots + A(t)B(t-1)\bar{u}(t-1) + B(t)\bar{u}(t),$$

где $[t \mu]$ - целая часть числа $t \mu$.

Из (13.11) следует, что (13.10) эквивалентно соответственно

$$\lim_{t \rightarrow \infty} |A(t - \mu[t \mu]) \dots A(0)\hat{A}^{[t \mu]}(\bar{s}^1(0) - \bar{s}^2(0))| = 0 \quad (13.12)$$

Согласно лемме 5.6.10 из [68], существует по крайней мере одна матричная форма, для которой справедливы оценки $\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon$ при заданном ε . Отсюда следует, что

$$\|\hat{A}\| = \|A(\mu - 1)a(0)\| \leq \|A(\mu - 1)\| \dots \|A(0)\| \leq \prod_{i=1}^{\mu-1} (\rho(A_i) + \varepsilon_i)$$

Поскольку по условиям теоремы $\rho(A(i)) < 1$ для $i = \overline{1, \mu}$, при соответствующем выборе величин ε_i каждый из сомножителей в правой части последнего неравенства, а следовательно, и вся его правая часть, может быть сделана меньше единицы. Из этого следует (см. лемму 5.6.11 из [68]), что

$$\lim_{t \rightarrow \infty} \hat{A}^k = [0] \ 0$$

т. е. все элементы матрицы \hat{A}^k стремятся к нулю при $k \rightarrow \infty$.

Условимся говорить, что матрица G не превосходит матрицу H (G и H имеют одинаковую размерность), и писать $G \leq H$, если каждый элемент матрицы G не превосходит соответствующего элемента матрицы H . Положим

$$M = \max_{1 \leq m \leq \mu} |q_{ij}^{(m)}|, \quad i, j = \overline{1, n}$$

где $q_{ij}^{(m)}$ - элементы матрицы $A(m-1)A(m-2) \dots A(0)$. Тогда при любом t

$$A(t - \mu[t \mu]) \dots A(0) \hat{A}^{[t \mu]}(s^{\bar{1}}(0) - s^{\bar{2}}(0)) \leq D(M) \hat{A}^{[t \mu]}(s^{\bar{1}}(0) - s^{\bar{2}}(0))$$

На основании (13.13) и (13.14) и того факта, что длина вектора $s^{\bar{1}}(0) - s^{\bar{2}}(0)$ ограничена в силу ограниченности $|s^{\bar{1}}|$ и $|s^{\bar{2}}|$ согласно условиям теоремы, вытекает справедливость (13.12).

Следствие. Для периодического НЛА \tilde{A} , у которого при любом t спектральный радиус матрицы $A(t)$ меньше 1, а множество $Init(\tilde{A})$ ограничено, любая бесконечная входная последовательность $\{\tilde{u}(t)\}$ является ε -синхронизирующей.

Можно показать, что по аналогии со стационарными ЛА для нестационарных ЛА при выполнении условий теоремы 13.6 понятие ε -синхронизируемости и асимптотической устойчивости равнозначны.

Заметим, наконец, что решение задачи о переводе НЛА в заданное синхросостояние можно осуществить тем же самым методом, что описан для стационарного ЛА.

Вопросы и упражнения

1. Дайте определения состояния равновесия и асимптотически устойчивого состояния свободного ЛА.
2. Какова связь между синхронизируемостью ЛА и асимптотической устойчивостью состояний?
3. Приведите постановку задачи стабилизации ЛА.
4. Всегда ли разрешима эта задача для ЛА с парой ее невырожденных характеристических матриц?
5. Дайте определение ε -синхронизирующей последовательности для дискретной линейной системы над полем \mathbb{R} и поясните ее содержательный смысл.
6. Сформулируйте достаточное условие существования ε -синхронизирующей последовательности для дискретной линейной системы над полем \mathbb{R} .
7. Дайте определение синхросостояния дискретной линейной системы.
8. Опишите метод решения задачи поиска входной

последовательности, переводящей ДЛС из любого состояния в ε -окрестность состояний и выходов ЛА.

Эксперименты по распознаванию неисправностей линейных автоматов

Рассматривается задача построения тестовой последовательности для заданной неисправности в стационарных и нестационарных ЛА, обнаруживающей эту неисправность. Описаны простые в реализации методы построения такой тестовой последовательности для λ -определенных и синхронизируемых линейных автоматов. Для произвольных ЛА предложен метод построения тестовой последовательности, основанный на наличии у любого ЛА конечной памяти.

В процессе эксплуатации в цифровых устройствах могут возникнуть различные неисправности. Для обеспечения правильного функционирования этих устройств необходимо проводить их проверку с целью обнаружения возможных неисправностей. Один из способов обнаружения неисправностей заключается в прерывании режима штатной работы устройства и подачи на него специально построенной входной последовательности, называемой тестом. Тест должен обладать следующим свойством: реакция проверяемых устройств на него должна быть различной в зависимости от того, является ли это устройство исправным или содержит неисправность. Следовательно, процесс обнаружения неисправностей представляет собой эксперимент, проводимый над устройством.

Эта лекция посвящена проблеме обнаружения неисправностей в цифровых системах, описываемых моделями линейных автоматов. Ранее известные методы построения тестов для ЛА, описанные в работах [2], [27], содержали жесткие требования, касающиеся как наличия информации о начальном состоянии, так и схемной реализации. Заметим, что эти требования почти никогда не выполняются.

Предлагаемые ниже методы не требуют выполнения упомянутых жестких ограничений и являются менее трудоемкими. Далее предполагается, что в ЛА могут возникать неисправности, приводящие к изменению его характеристических матриц, но не увеличивающие числа его состояний и не выводящие исходное устройство из класса ЛА.

Предполагается также, что неисправность, возникнув в ЛА в некоторый момент, сохраняется в нем и во все последующие моменты времени. Последнее означает, что из рассмотрения исключаются только сбои и перемежающиеся неисправности, которые представляют собой кратковременные самоустраняющиеся неисправности одного и того же типа, появляющиеся в устройстве через некоторые не обязательно равные промежутки времени. Что же касается "исчезающих" неисправностей, то рассматриваемое разнообразие их видов предельно широко. В самом деле, любые из встречающихся на практике неисправностей, например, типа константных, перепутывания связей, коротких замыканий, изменения реализуемой вентилем функции и т. п., в конечном счете, если они являются существенными, приводят только к изменениям характеристических матриц ЛА.

Ниже рассматривается следующая задача. Пусть задан ЛА и некоторая его неисправная модификация из множества допустимых неисправных модификаций. Требуется построить такую входную последовательность (тест), которая эту неисправность обнаруживает.

Входную последовательность назовем тестом, обнаруживающим заданную неисправность, если независимо от начальных состояний исправного ЛА и его неисправной модификации их входные реакции на эту последовательность различны.

Построенные предложенными ниже методами тесты используются для проведения экспериментов по распознаванию (обнаружению) неисправностей. Тесты подаются на вход ЛА, фиксируется реакция проверяемого устройства на этот тест и по этой реакции делается заключение об исправности или неисправности устройства.

Синтез тестов для стационарных линейных автоматов

Напомним вначале некоторые понятия, которые понадобятся в дальнейшем.

Говорят, что ЛА \tilde{A} имеет конечную память глубины μ , если в любой момент времени t выход $y(t)$ однозначно определяется входом в этот же момент и предыдущими μ входами и μ выходами, т. е. для всех t

справедливо соотношение

$$\bar{y}(t) = f(\bar{u}(t), \bar{t} - 1), \dots, \bar{u}(t - \mu), \bar{y}(t - 1), \dots, \bar{y}(t - \mu))$$

Содержательно это означает, что реакция такого ЛА в произвольный момент времени может быть предсказана только на основе знания входной последовательности и соответствующей реакции ЛА в предшествующие μ моментов времени.

Из теории ЛА известно [19], что каждый ЛА имеет конечную память глубины μ , где $\mu \leq n$ (n - размерность ЛА).

Говорят, что ЛА \tilde{A} является μ - определенным, если его выход $\bar{y}(t)$ в любой момент времени t зависит лишь от предыдущих μ входов, т. е. справедливо соотношение

$$\bar{y}(t) = f(\bar{u}(t), \bar{u}(t - 1), \dots, \bar{u}(t - \mu))$$

Перейдем теперь к описанию методов построения тестов. Пусть A_1, B_1, C_1, D_1 - характеристические матрицы ЛА \tilde{A}_1 , являющегося неисправной модификацией исходного ЛА \tilde{A} .

Рассмотрим случай, когда как исходный ЛА \tilde{A} , так и неисправный ЛА \tilde{A}_1 являются μ - определенными, однако значения параметра μ для них необязательно совпадают. Пусть $\mu = \max(\mu_1, \mu_2)$, где μ_1 и μ_2 - значения параметра μ для ЛА \tilde{A} и \tilde{A}_1 соответственно. В [19] доказано, что необходимым и достаточным условием μ - определенности ЛА является выполнение соотношения $CA^\mu = [0]$. Отсюда следует, что $CA^k = [0]$ и $C_1A_1^k = [0]$ для всех $k \geq \mu$.

Если на исправный и неисправный ЛА подана одна и та же входная последовательность $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(\mu)$, то с учетом только что приведенных равенств и формулы (1.4) полной реакции ЛА независимо от начальных состояний автоматов \tilde{A} и \tilde{A}_1 соответствующие реакции имеют вид

$$\begin{aligned}\bar{y}(\mu) &= CA^{\mu-1}B\bar{u}(0) + CA^{\mu-2}B\bar{u}(1) + \dots + CB\bar{u}(\mu-1) + D\bar{u}(\mu), \\ \bar{y}_1 &= C_1A_1^{\mu-1}B_1\bar{u}(0) + C_1A_1^{\mu-2}B_1\bar{u}(1) + \dots + C_1B_1\bar{u}(\mu-1) + D_1\bar{u}(\mu)\end{aligned}$$

Произведя вычитание, получим

$$\bar{y}(\mu) - \bar{y}_1(\mu) = [CA^{\mu-1}B - C_1A_1^{\mu-1}B_1]\bar{u}(0) + \dots + [D - D_1]\bar{u}(\mu)$$

Понятно, что заданная неисправность будет обнаружена входной последовательностью (тестом) $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(\mu)$, если и только если реакции $\bar{u}(\mu)$ и $\bar{y}_1(\mu)$ различаются, т. е. $\bar{y}(\mu) - \bar{y}_1(\mu) \neq [0]$.

Соотношение (14.1) будем рассматривать как СЛАУ относительно неизвестных, являющихся координатами вектора

$$u = [u_1(0), \dots, u_l(0), \dots, u_1(\mu), \dots, u_l(\mu)] \quad (14.2)$$

и на ее основе организуем процедуру построения теста. Обозначим через Q матрицу системы (14.1); тогда (14.1) переписывается в виде

$$Qu = y \quad (14.3)$$

где y - некоторый m - мерный ненулевой вектор.

Обозначим через I множество всех тестов, обнаруживающих заданную неисправность. Для поиска множества I необходимо варьировать правую часть системы (14.3) и при каждой конкретной правой части находить решение системы, если оно существует.

Легко сообразить, что число всевозможных правых частей системы (14.3) равно величине $p^{(\mu+1)m}$, которая даже при сравнительно небольших значениях параметров p , μ и m достаточно велика. По этой причине предлагаемый метод довольно трудоемок и мы рассмотрим другой, более эффективный метод.

Суть его состоит в том, что вместо неоднородной системы (14.3) будет рассматриваться соответствующая ей однородная система

$$Qu = [0] \quad (14.4)$$

множество решений которой обозначим через U_0 . Если U - множество всевозможных векторов вида (14.2), то очевидно, что множество U/U_0 является искомым множеством T тестов. Таким образом, при такой организации построения тестов все свелось к решению единственной однородной системы уравнений.

Пусть ранг матрицы Q равен r . Если $r = (\mu + 1)l$, то из алгебры известно, что нулевое решение будет единственным решением однородной системы (2.4), а при $r < (\mu + 1)l$ система имеет ненулевые решения, метод поиска которых известен.

Проиллюстрируем предложенный метод на примере. Пусть ЛА над полем $GF(2)$ задан следующими характеристическими матрицами ($n = 4, l = 2, m = 2$):

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$C = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Пусть неисправный автомат задан матрицами

$$A_1 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},$$

$$B_1 = B,$$

$$C_1 = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix},$$

$$D_1 = D$$

Легко проверить, что $CA^2 = [0]$ и $CA_1^2 = [0]$, т. е. оба ЛА являются μ -определенными ($\mu = 2$).

По формуле полной реакции имеем следующие результаты:

$$\bar{u}(2) = CAB\bar{u}(0) \oplus CB\bar{u}(1) \oplus D\bar{u}(2),$$

$$\bar{y}_1(2) = C_1A_1B\bar{u}(0) \oplus C_1B\bar{u}(1) \oplus D\bar{u}(2)$$

Отсюда получаем однородную линейную систему вида (14.4):

$$[CAB \oplus C_1A_1B]\bar{u}(0) \oplus [CB \oplus C_1B]\bar{u}(1) = [0]$$

Выполнив соответствующие вычисления, придем к следующей однородной системе уравнений в координатной форме:

$$u_1(0) \oplus u_2(0) = 0;$$

$$u_1(0) \oplus u_2(1) = 0$$

Легко проверить, что ранг матрицы последней системы $r = 2$, т. е. он меньше числа неизвестных ($\mu + 1$) $l = 3 \times 2 = 6$. Отсюда следует, что система имеет ненулевое решение. Считая, например, переменные $u_2(0)$ и $u_2(1)$ свободными, выразим через них оставшиеся переменные

$$u_1(0) = u_2(0); u_1(1) = u_2(1) \quad (14.5)$$

Отметим, что в рассматриваемой системе переменные $u_1(2)$ и $u_2(2)$

отсутствуют в силу специфики самой системы, т. е. выступают здесь в роли фиктивных переменных. Из сказанного вытекает, что решениями этой однородной системы являются всевозможные значения $u_1(0), u_2(0), u_1(1), u_2(1)$, удовлетворяющие соотношениям (14.5), а также произвольные значения переменных $u_1(2), u_2(2)$.

Зная множество U_0 решений рассмотренной однородной системы, построим множество T всех тестов, обнаруживающих заданную неисправность, как дополнение множества U_0 до множества всех 2^6 шук двоичных входных наборов $[u_1(0), u_2(0), u_1(1), u_1(2), u_1(2), u_2(2)]$ длины 6. Таким образом, искомое множество T составляют всевозможные двоичные входные наборы указанного типа, у которых $u_1(0) \neq u_2(0)$ и $u_1(1) \neq u_2(1)$.

Обратимся теперь к методу построения тестов для синхронизируемых автоматов. В соответствии с приведенным ранее критерием синхронизируемости для любого синхронизируемого ЛА существует такое целое число k , что $A^k = [0]$, но тогда справедливо и равенство $CA^k = [0]$ при любой матрице C . Отсюда вытекает справедливость следующего утверждения.

Теорема 14.1. Каждый синхронизируемый ЛА является одновременно и μ -определенным.

Понятно, что эта теорема обосновывает возможности применения описанного выше метода построения тестов к синхронизируемым ЛА.

Предложенный метод построения тестов предполагает, что неисправные ЛА должны сохранять свойства μ -определенности или синхронизируемости. Заметим, что это требование не является слишком ограничительным. Так, оно заведомо выполняется, если возникающие в автомате неисправности сказываются только на характеристических матрицах B, C и D . Однако не всякое изменение матрицы A указанные свойства будет сохранять.

Перейдем теперь к описанию метода построения тестов для произвольного ЛА, а не обязательно μ -определенного либо

синхронизируемого. Этот метод базируется на том факте, что любой ЛА является автоматом с конечной памятью.

Предположим, что исправный (неисправный) ЛА имеет глубину памяти μ_1 (μ_2) и пусть $\mu = \max(1, \mu_2)$. Из теории ЛА известно [19], что функции выходов ЛА \tilde{A} и \tilde{A}_1 соответственно можно представить в виде:

$$\begin{aligned} \bar{y}(t)V_0^1\bar{u}(t) + V_1\bar{u}(t-1) + \dots + V_\mu\bar{u}(t-\mu) + W_0\bar{y}(t-1) + \dots + W_\mu\bar{y}(t-\mu), \\ \bar{y}_1(t) = V_0^1\bar{u}(t) + V_1^1\bar{u}(t-1) + \dots + V_\mu^1\bar{u}(t-\mu) + W_0^1\bar{y}(t-1) + \dots + W_\mu^1\bar{y}(t-\mu), \end{aligned}$$

где $V_i(V_i^1), W_i(W_i^1)$ - матрицы соответствующей размерности. Способ приведения функции выхода ЛА к виду (14.6) описан в [19], и мы не будем здесь на нем останавливаться.

Очевидно, что входная последовательность $\bar{u}(t-\mu), u(t-\mu+1), \dots, u(t)$ является минимальным по длине тестом, если выходные реакции исправного и неисправного автоматов для всех $j = 1, \dots, \mu$ совпадают, т. е. $\bar{y}(t-j) = \bar{y}_1(t-j)$, а на последнем такте $\bar{y}(t) \neq \bar{y}_1(t)$. С учетом сказанного, произведя вычитание выражений в (14.6), получим

$$\bar{y}(t) - \bar{y}_1(t) = [V_0 - V_0^1]\bar{u}(t) + \dots + [V_\mu - V_\mu^1]\bar{u}(t-\mu) + [W_0 - W_0^1] + \dots + [W_\mu - W_\mu^1]\bar{y}(t-\mu)$$

Приравняв выражение (14.7) некоторому ненулевому вектору y , мы получим СЛАУ относительно неизвестных, являющихся координатами вектора

$$u = [u_1(t-\mu), \dots, u_1(t-\mu), \dots, u_1(t), \dots, u_1(t)]'$$

Если через Q обозначим матрицу полученной СЛАУ, то в этом обозначении она примет вид

$$Qu = y$$

Заметим, что все сказанное выше относительно поиска решений

системы (14.3) остается справедливым и для последней системы. Таким образом, построение множества всех тестов для произвольных ЛА, так же как для μ -определенных и синхронизируемых ЛА, сводится к решению системы линейных уравнений.

Проиллюстрируем предложенный метод на примере ЛА над полем $GF(2)$ с характеристическими матрицами

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix},$$

$$C = [0 \quad 1 \quad 0 \quad 1], D = [1]$$

Найдем представление функции выхода этого ЛА в виде (14.6). Поскольку минимальный многочлен матрицы A есть $m_A(x) = 1 + x + x^2$, получаем

$$CA^2\bar{s}(0) = C\bar{s}(0) + CA\bar{s}(0)$$

Тогда, используя формулы полной реакции ЛА и только что полученное равенство, вычислим

$$\begin{aligned} y(2) &= CA^2\bar{s}(0) + CABu(0) + CBu(1) + Du(0) = \\ &= C\bar{s}(0) + CA\bar{s}(0) + CABu(0) + CBu(1) + Du(1) = \\ &= [y(0) + Du(0)] + [y(1) + CBu(0) + Du(1)] + CABu(0) + CBu(1) + Du(1). \end{aligned}$$

После преобразований получаем

$$y(2) = u(2) + u(0) + u(0) + y(1) + y(0),$$

откуда для любого t

$$y(t) = u(t) + u(t-2) + y(t-1) + y(t-2).$$

Пусть неисправный ЛА задан следующими характеристическими матрицами:

$$A_1 = A, B_1 = B, C_1 = [1010], D_1 = D.$$

Проведя аналогичные вычисления для неисправного ЛА, получим

$$y_1(t) = u(t) + u(t-1) + y_1(t-1) + y_1(t-2).$$

Упомянутая выше СЛАУ для нашего примера примет вид

$$y(t) - y_1(t) = u(t-1) \oplus u(t-2)$$

Понятно, что для получения множества всех тестов, обнаруживающих заданную неисправность, необходимо найти все решения СЛАУ

$$u(t-1) \oplus u(t-2) = 1$$

Отсюда получаем, что

$$u(t-1) = 1 \oplus u(t-2) = \overline{u(t-2)}$$

Таким образом, множество всех искомых тестов составляют такие тройки двоичных наборов $(u(t), u(t-1), u(t-2))$, у которых $u(t)$ может иметь произвольное значение, а величины $u(t-1)$ и $u(t-2)$ должны быть инверсными.

Отметим следующий важный факт: предложенный метод строит тест длины $\mu + 1$, где μ - глубина памяти ЛА. Поскольку, как известно, $\mu \leq n$, где n - размерность ЛА, такой метод дает возможность строить достаточно короткие по длине тесты.

Синтез тестов для нестационарных линейных автоматов

Для нестационарных ЛА рассмотрим ту же задачу, которая исследовалась в предыдущем разделе для стационарных ЛА. При этом будем предполагать, что класс неисправностей, возникающих в НЛА, полностью идентичен соответствующему классу в случае стационарных ЛА.

По аналогии со стационарными ЛА вначале рассмотрим метод синтеза тестов для синхронизируемых НЛА. Идея этого метода ничем не отличается от случая стационарного ЛА, и поэтому мы обозначим только некоторые его ключевые моменты.

Как и ранее, предполагается, что исправный и неисправный НЛА являются синхронизируемыми, однако длины минимальных СП для них не обязательно совпадают. Пусть $k = \max(k_1, k_2)$, где k_1 и k_2 - длины минимальных СП для исправного НЛА \tilde{A} и неисправного НЛА \tilde{A}_1 соответственно. Используя формулу (1.8) полной реакции НЛА и необходимое и достаточное условие (2.11) существования СП длины $k + 1$, как и в случае стационарного ЛА приходим к выводу, что поиск искомой тестовой последовательности $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ для обнаружения заданной неисправности сводится к решению системы линейных алгебраических уравнений следующего вида

$$\sum_{i=0}^{k-1} [C(k)A(k-1) \dots A(i+1)B(i) - C_1(k)A_1(k-1) \dots A_1(i+1)B_1(i)]\bar{u}(i) + [D(k) - D_1(k)]\bar{u}(k) = y$$

Здесь y - некоторый ненулевой вектор, а матрица, снабженная индексом "1", относится к неисправному НЛА.

Множество всех искомых тестов может быть получено путем решения СЛАУ (14.8) для всевозможных ненулевых правых ее частей. Поскольку этот способ требует рассмотрения множества СЛАУ, более эффективным является другой способ, основанный на поиске решения одной однородной СЛАУ, соответствующей неоднородной системе (14.8), с последующим построением с помощью "однородного" решения искомых "неоднородных".

Вообще говоря, может оказаться, что при рассматриваемом значении k

система (14.8) не является совместной ни при каких правых частях. В этом случае увеличим k на единицу и будем строить новую систему вида (14.8). Этот процесс продолжается далее аналогичным образом до тех пор, пока либо не будет найдено множество всех тестов, либо величина k не достигнет верхней границы длины минимального теста. Такой границей, в частности, может служить верхняя оценка длины минимальной УП [67].

Проиллюстрируем предложенный метод на примере. Пусть $n = l = m = 2$ и НЛА над полем $GF(2)$ задана следующими характеристическими матрицами:

$$\begin{aligned}
 A(0) &= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \\
 A(1) &= \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \\
 A(2) &= \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \\
 A(i) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, i = 3, 4, \dots; \\
 B(0) &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\
 B(1) &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \\
 B(2) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\
 B(i) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, i = 3, 4, \dots; \\
 C(i) &= D(i) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, i = 3, 4, \dots
 \end{aligned}$$

Предположим, что неисправный НЛА задается следующими характеристическими матрицами:

$$\begin{aligned}
 A_1(0) &= \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}, \\
 A_1(1) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, A_1(2) = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}, \\
 A_1(i) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, i = 3, 4, \dots; \\
 B_1(0) &= \begin{bmatrix} 0 \\ 0 & 0 \end{bmatrix}, \\
 B_1(1) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \\
 B_1(2) &= \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \\
 B_1(i) &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, i = 3, 4, \dots; \\
 C_1(i) = D_1(i) &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, i = 3, 4, \dots
 \end{aligned}$$

Вычисления показывают, что $A(1)A(0) = [0]$ и $A_1(2)A_1(1)A_1(0) = [0]$.

Следовательно, как исправный, так и неисправный НЛА синхронизируемы и длины СП для них равны 2 и 3 соответственно. Исходя из этого, начнем построение тестов длины $k = 4$. Итак, вначале в качестве кандидата в тесты рассмотрим входную последовательность $\bar{u}(0), \bar{u}(1), \bar{u}(2), \bar{u}(3)$. Выпишем реакции на эту последовательность исправного и неисправного НЛА:

$$\bar{y}(3) = C(3)A(2)A(1)A(0)\bar{s}(0) + C(3)A(2)A(1)B(0)\bar{u}(0) + C(3)A(2)B(1)\bar{u}(1) + C(3)B(2)\bar{u}(2) + D(3)\bar{u}(3),$$

$$\begin{aligned}
 \bar{y}_1 &= C_1(3)A_1(2)A_1(1)A_1(0)s(0) + C_1(3)A_1(2)A_1(1)B_1(0)\bar{u}(0) + \\
 &+ C_1(3)A_1(2)B_1(1)\bar{u}(1) + C_1(3)B_1(2)\bar{u}(2) + D_1(3)\bar{u}(3)
 \end{aligned}$$

Выполнив соответствующие вычисления в координатной форме с учетом того, что все операции выполняются по модулю 2, получим следующие результаты:

$$\bar{y}(3) = \begin{bmatrix} u_1(1) + u_2(1) + u_1(3) + u_2(2) \\ u_1(1) + u_2(1) + u_1(2) + u_2(3) \end{bmatrix},$$

$$y_1(3) = \begin{bmatrix} u_2(2) + u_1(3) \\ u_1(1) + u_2(1) + u_1(2) + u_2(3) \end{bmatrix}$$

Используя эти векторы, построим однородную систему, соответствующую системе (14.8):

$$\bar{y}(3) + \bar{y}_1 = \begin{bmatrix} u_1(1) + u_2(1) \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Ее решениями являются все четверки векторов $\bar{u}(0), \bar{u}(1), \bar{u}(2), \bar{u}(3)$, у которых $u_1(1) = u_2(1)$. Отсюда следует, что искомое множество тестов составляет такие же четверки векторов, но для которых справедливо условие $u_1(1) \neq u_2(1)$. Легко подсчитать, что всего таких векторов 128 штук.

Предложенный метод предполагает сохранение свойства синхронизируемости у неисправного НЛА, если исправный НЛА был синхронизируем. Понятно, что это не всегда выполняется. В связи с этим рассмотрим метод синтеза тестов, ориентированный на произвольные минимальные НЛА. Этот метод есть прямой аналог метода распознавания автоматов, принадлежащих исключительно классу, описанному в [18].

Опишем кратко его суть применительно к рассматриваемой нами задаче, когда исключительный класс состоит только из двух автоматов, исправного и неисправного. При этом условимся считать, что множества их допустимых состояний совпадают с соответствующими множествами всех состояний автоматов.

Для исправного НЛА A построим минимальную УП \bar{u}_1 , которая существует в силу минимальности НЛА, и подадим ее на предъявленный для эксперимента экземпляр автомата. Естественно, что экспериментатору неизвестно, является ли этот автомат исправным или неисправным. По наблюдаемой реакции найдем конечное состояние $\bar{s}_f(\bar{A})$, в котором должен оказаться испытываемый автомат, если бы

он был исправным.

Далее построим УП \bar{u}_2 для неисправного НЛА \tilde{A}_1 и подадим ее на предъявленный для эксперимента экземпляр автомата. По наблюдаемой реакции найдем конечное состояние $\bar{s}_1(\tilde{A}_1)$, в котором должен оказаться испытываемый автомат, если бы он был исправным. Предполагается, что НЛА \tilde{A}_1 представлен своей минимальной формой и потому УП для него существует.

Напомним, что в силу теоремы 2.2 если НЛА имеет хотя бы одну УП длины k , то все другие входные последовательности длины k и более также являются для этого НЛА установочными. Это свойство позволяет "наложить" друг на друга УП \bar{u}_1 и \bar{u}_1 и тем самым сократить длину последовательности, осуществляющей установку НЛА \tilde{A} и \tilde{A}_1 . В самом деле, если выбрать произвольную входную последовательность \bar{u} длины $k = \max(k_1, k_2)$, где $k_1(k_2)$ есть длина минимальной УП для автомата $\tilde{A}(\tilde{A}_1)$, то эта последовательность будет устанавливать в известное экспериментатору состояние как автомат \tilde{A} так и автомат \tilde{A}_1 .

Для обнаружения неисправности теперь осталось ответить на вопрос, является ли испытываемый автомат автоматом \tilde{A} в состоянии $\bar{s}_1(\tilde{A})$ либо автоматом \tilde{A}_1 в состоянии $\bar{s}_f(\tilde{A}_1)$. Для ответа на этот вопрос подадим на вход испытываемого автомата входную последовательность, построенную на основе метода различающей функции. Синтез этой последовательности осуществим в виде пошагового процесса, делая попытки найти сначала последовательности длины 1, затем длины 2 и т. д. С этой целью сначала выпишем реакции автоматов \tilde{A} и \tilde{A}_1 с начальными состояниями $\bar{s}_f(\tilde{A})$ и $\bar{s}_f(\tilde{A}_1)$ на входную последовательность $\bar{u}(0)$ длины 1:

$$\begin{aligned}\bar{y}(0) &= C(0)\bar{s}_f(\tilde{A}) + D(0)\bar{u}(0), \\ y_1(0) &= C_1(0)\bar{s}_f(\tilde{A}_1) + D_1(0)\bar{u}(0)\end{aligned}$$

Теперь сформируем выражение (различающую функцию):

$$\bar{y}(0) - \bar{y}_1(0) = \bar{b}$$

где \bar{b} - некоторый произвольный ненулевой вектор. Это выражение будем интерпретировать как СЛАУ относительно координат $\bar{u}(0)$. Решая эту систему при всевозможных ненулевых правых частях, найдем все различающие последовательности длины 1. Если эта система окажется несовместной, сформируем очередную систему

$$\bar{y}(1) - \bar{y}_1(1) = \bar{b}$$

и ищем ее решение. Указанный процесс продолжается далее аналогичным образом до тех пор, пока не будет найдена различающая последовательность. Этот процесс заведомо конечен, поскольку неисправности в НЛА предполагаются существенными, т. е. проявляющими себя в процессе функционирования. Кстати, вывод о конечности процесса построения различающей последовательности следует также и из того факта, что упомянутая последовательность есть не что иное, как ДП минимального автомата для двух допустимых начальных состояний. Такая диагностическая задача, как известно [18], всегда разрешима.

Таким образом, искомый тест будет состоять из двух последовательностей, первая из которых представляет собой совмещенную УП, а вторая - различающую последовательность.

Остановимся теперь на верхней оценке длины теста, полученного описанным методом.

В разделе 2.3 лекции 2 было показано, что в общем случае длина минимальной УП для НЛА не ограничена сверху. Поэтому рассмотрим периодические НЛА, у которых длина УП, а следовательно, и теста, заведомо конечна.

В том же разделе 2.3 было установлено, что для периодического НЛА длина УП не превосходит λn где λ - период главной характеристической матрицы A НЛА \tilde{A} , n - ее размерность. Что касается второй части теста, то ее можно рассматривать как ДП и

поэтому длина ее не превосходит n . Таким образом, общая длина теста, построенного таким методом, не превышает величины $\lambda n + n = (\lambda + 1)n$.

Сравним эту оценку с соответствующей верхней оценкой длины теста для произвольного автомата, не являющегося НЛА. Как показано в [18], верхняя оценка длины теста в этом случае такова:

$$\nu \leq (2q - 1)^2 = O(q^2).$$

Здесь q - число состояний автомата. Поскольку для НЛА $q = p^n$, в терминах параметра q оценка длины теста для НЛА не превышает $(\lambda + 1) \log_p q$, что значительно ниже $O(q^2)$.

Вопросы и упражнения

1. Дайте определение теста, обнаруживающего заданную неисправность в ЛА.
2. Сформулируйте определение конечной памяти глубины μ линейного автомата и μ -определенного автомата.
3. Приведите постановку задачи построения теста, обнаруживающего заданную неисправность.
4. К какой математической задаче сводится задача построения теста?
5. Опишите метод построения теста для μ -определенного и синхронизируемого линейного автомата.
6. Опишите метод построения теста для произвольных стационарных линейных автоматов.
7. Чем от предыдущего метода отличается метод построения теста для нестационарных линейных автоматов?
8. Как длина минимального теста для ЛА связана с глубиной его памяти?
9. Для примера ЛА из раздела 2.1 лекции 2 постройте тест для обнаружения неисправности, отличающейся от заданной в разделе 2.1 тем, что вторая (единичная) строка в главной матрице неисправного ЛА заменена нулевой строкой.

Линейные автоматы существенно без потери информации

Описывается класс автоматов без потери информации и доказывается эквивалентность двух разных его определений, данных Гиллом и Хаффменом. Вводится новая разновидность автоматов БПИ, названных автоматами существенно БПИ конечного порядка, и приводится критерий принадлежности автомата этому классу. Описывается метод синтеза комбинационной схемы, восстанавливающей первый символ неизвестного слова, поданного на вход автомата БПИК.

В работах Хаффмена [76] и Ивена [74] были впервые введены автоматы без потери информации (БПИ) и автоматы БПИ конечного порядка (БПИК). Эти автоматы позволяют восстановить неизвестную входную последовательность по наблюдаемой реакции и известному начальному состоянию. В работах [43], [46] были введены и исследованы так называемые автоматы существенно без потери информации (СБПИ) и автоматы СБПИ конечного порядка (СБПИК), являющиеся неинициальными, т. е. такими, у которых начальное состояние неизвестно. В упомянутых работах в терминах графов переходов автомата были получены критерии принадлежности конкретных автоматов перечисленным классам. Заметим, что проверка этих критериев является достаточно трудоемким делом. Ниже будет показано, что для линейных автоматов аналогичные критерии могут быть сформулированы в терминах характеристических матриц и проверяются они достаточно просто.

Отметим, что дискретные устройства, описываемые моделями ЛА СБПИ, обладают одной важной особенностью - для них могут быть построены сравнительно простые и эффективные схемы встроенного контроля (СВК). Такие СВК основаны на сравнении сигналов, поступающих на входы устройств, с сигналами, восстанавливаемыми по выходам. Рассогласование этих сигналов свидетельствует о неправильном функционировании автомата и дает возможность зафиксировать этот факт в момент первого его проявления, не допуская распространения появившейся ошибки.

Использование СВК в той роли, которая только что описана, позволяет осуществлять так называемый функциональный контроль устройства.

Этот вид контроля ведется непрерывно в процессе функционирования устройства и параллельно с его работой, направленной на выполнение функций, для выполнения которых он создан. В отличие от тестового контроля, требующего прерывания работы устройства для подачи на вход устройства тестов, функциональный контроль необходим в тех случаях, когда прерывание невозможно по условиям эксплуатации. Безусловно, и тестовый, и функциональный контроль имеют как достоинства, так и недостатки. Так, тестовый контроль дает возможность более точно локализовать неисправности в устройстве и не требует дополнительных аппаратурных затрат. Вместе с тем необходимость прерывания штатной работы устройства, несомненно, является существенным недостатком. Недостатками функционального контроля являются требуемые для его реализации дополнительные аппаратурные затраты (в виде СВК) и невысокая точность локализации неисправностей.

Для преодоления указанных выше недостатков естественно стремление разумным образом сочетать для дискретных устройств как средства тестового, так и функционального контроля. Для устройств, описываемых моделями ЛА, это можно осуществить, выделив из них такие максимальные по размерам подсхемы, которые представляют собой либо автоматы БПИ, либо некоторую их модификацию, с тем, чтобы организовать для них функциональный контроль с применением СВК. Оставшуюся часть исходного устройства можно контролировать с помощью тестов, затраты на построение которых уменьшаются по сравнению с затратами на построение тестов для устройства в целом. В качестве одной из модификаций ЛА БПИ являются обобщенные автоматы БПИ, исследованные в лекции 7.

Напомним определение автомата БПИ, данное Хаффменом [76] в связи с рассмотрением следующей задачи. На автомат, находящийся в известном начальном состоянии \bar{s}_0 , подается неизвестная входная последовательность (слово) \mathcal{P} и наблюдается реакция w автомата на это слово. Требуется построить эксперимент, проводимый с автоматом после подачи слова \mathcal{P} , позволяющий распознать его. Автоматы, для которых сформулированная задача может быть решена независимо от входной последовательности и начального состояния, называются автоматами БПИ.

В монографии А. Гилла [19] применительно к линейным автоматам также введено иное определение автомата БПИ: линейный автомат называется автоматом БПИ, если по его начальному состоянию и наблюдаемой выходной реакции можно однозначно восстановить неизвестное входное слово. Таким образом, в этом определении, в отличие от определения Хаффмена, проведения дополнительного эксперимента не требуется.

Покажем, что эти два определения применительно к ЛА являются эквивалентными. Докажем утверждение, которое понадобится для этого.

Лемма 15.1. Если ранг матрицы D ЛА A равен l , то из любой вершины графа переходов этого автомата не может выходить более одной дуги, помеченной одним и тем же выходным сигналом.

Доказательство. Проведем его методом от противного. Пусть из некоторого состояния $\bar{s} \in S_n$ выходят две дуги, соответствующие двум различным входным сигналам \bar{u}_1 и \bar{u}_2 , помеченные одним и тем же выходным символом \bar{y} . Это означает, что $\bar{u} = C\bar{s} + D\bar{u}_1$ и $\bar{y} = C\bar{s} + D\bar{u}_2$. Вычитая одно равенство из другого, получим $D(\bar{u}_1 - \bar{u}_2) = [0]$, или, что то же самое, $D\bar{u} = [0]$. Последнее равенство можно рассматривать как линейную однородную систему уравнений относительно неизвестных $u = (u_1, u_2, \dots, u_l)'$, записанную в матричной форме и имеющую ненулевое решение. Известно, что нетривиальное решение однородная система имеет тогда и только тогда, когда ранг матрицы D меньше l , что противоречит условиям леммы.

Вернемся теперь к вопросу об эквивалентности двух приведенных выше определений автомата БПИ.

Пусть автомат Мили A , находящийся в состоянии \bar{s}_0 , после подачи на него неизвестного входного слова P выдал слово w и перешел в состояние \bar{s} , неизвестное экспериментатору. Подадим на вход установочное слово P_1 (оно всегда существует, если автомат Мили минимален) и по реакции w_1 определим его конечное состояние \bar{s}_1 . В [18] доказано, что по состоянию \bar{s}_1 однозначно устанавливается

состояние \bar{s} . Наконец, знание состояний \bar{s}_0 и \bar{s}_1 и реакции w дает возможность распознать входное слово P . Таким образом, установочный эксперимент играет роль лишь промежуточного звена для определения состояния \bar{s} .

В случае линейного автомата БПИ, учитывая лемму 15.1, справедлив следующий факт: при наблюдаемой реакции w ЛА, стартующего в состоянии \bar{s}_0 , на неизвестное входное слово P в графе ЛА существует единственный путь из \bar{s}_0 в \bar{s} , но не несколько путей, как это может быть в случае автомата Мили БПИ. Отсюда следует, что для ЛА дополнительный установочный эксперимент будет давать избыточную информацию для распознавания слова P , поскольку отказ от такого эксперимента не повлияет на возможность восстановления упомянутого слова. Итак, знания начального состояния и выходной реакции ЛА достаточно для распознавания неизвестного входного слова и потому применительно к ЛА оба приведенных определения автомата БПИ эквивалентны.

Приведем критерий принадлежности автомата классу автоматов БПИ, доказанный в [19].

Теорема 15.1. Для того чтобы ЛА A являлся ЛА БПИ, необходимо и достаточно, чтобы ранг матрицы D равнялся l .

Доказательство. Из уравнения переходов ЛА (10.1) следует, что

$$\bar{y}(0) - C\bar{s}(0) = D(0)$$

Последнее равенство можно интерпретировать как систему линейных алгебраических уравнений относительно неизвестных $u_1(0), u_2(0), \dots, u_l(0)$, являющихся координатами входного вектора $\bar{u}(0)$. Из алгебры известно, что эта система имеет решение тогда и только тогда, когда ранг D равен l . Зная вектор $\bar{u}(0)$, по формуле (1.1) вычисляем состояние $\bar{s}(1)$, а затем и вектор $\bar{u}(1)$ из системы

$$\bar{y}(1) - C\bar{s}(1) = Du(1)$$

Ясно, что справедливость теоремы легко теперь доказать методом

математической индукции.

Введем еще одну разновидность автоматов БПИ.

Пусть неизвестная входная последовательность P подается на ЛА A , начальное состояние которого неизвестно и может быть любым состоянием из пространства S_n , и наблюдается реакция w на эту последовательность. Если для заданного ЛА A существует такое натуральное число $N(A)$, что знания начального отрезка длины $N(A)$ слова w достаточно для однозначного определения первого символа слова P независимо от входной последовательности P и начального состояния ЛА, то A называют ЛА существенно БПИ конечного порядка (СБПИК).

Для однозначного определения числа $N(A)$ будем считать, что оно является минимальным из всех возможных чисел такого рода и называть порядком ЛА A . Условимся далее ЛА СБПИК порядка k обозначать как ЛА СБПИК- k .

Перейдем теперь к нахождению критерия принадлежности ЛА классу автоматов СБПИК- k .

Введем в рассмотрение следующие матрицы:

$$\begin{aligned}
 G(0) &= [D, C], \\
 G(1) &= \begin{bmatrix} D & 0 & C \\ CB & D & CA \end{bmatrix}, \\
 G(k) &= \begin{bmatrix} D & [0] & \dots & [0] & [0] & C \\ CB & D & \dots & [0] & [0] & CA \\ \dots & \dots & \dots & \dots & \dots & \dots \\ CA^{k-1}B & CA^{k-2}B & \dots & CB & D & CA^k \end{bmatrix} \quad (15).
 \end{aligned}$$

Принципы построения этих матриц легко получить из формулы (1.4).

Условимся через $\bar{a}(k)$ и $\bar{y}(k)$ обозначать векторы

$$\bar{a}(k) = (u_1(0), \dots, u_l(0), \dots, u_1(k-1), \dots, u_l(k-1), s_1(0), \dots, s_k(0))',$$

$$\bar{y}(k) = (y_1(0), \dots, y_m(0), \dots, y_1(k-1), \dots, y_m(k-1))'$$

Используя эти обозначения, запишем в матричной форме следующую СЛАУ:

$$G(k)\bar{a}(k) = \bar{y}(k) \quad (15.2)$$

где неизвестными являются координаты вектора $\bar{s}(k)$.

Используя метод Гаусса решения СЛАУ [33], матрицу $G(k)$, к которой добавлен столбец $\bar{y}(k)$, всегда можно привести к виду

$$\bar{G}(k) = \begin{bmatrix} 1 & 0 & \dots & g_{1,r+1} & \dots & g_{1,kl} & b_1 & \dots & \dots \\ 0 & 1 & \dots & g_{2,r+1} & \dots & g_{2,kl} & b_2 & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & g_{r,r+1} & \dots & g_{r,kl} & b_r & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & g_{d,r+1} & \dots & g_{d,kl} & b_d & \dots \end{bmatrix}$$

Здесь $d \leq m(k-1)$, $\bar{b} = (b_1, \dots, b_d)'$ - вектор, компоненты которого получены из компонент вектора $\bar{y}(k)$ в процессе преобразования по методу Гаусса матрицы $G(k)$ с добавленным к ней столбцом $\bar{y}(k)$, в матрицу $\bar{G}(k)$. Выделим из $\bar{G}(k)$ подматрицу, состоящую из первых l строк $\bar{G}(k)$ с отброшенным последним столбцом \bar{b} , а также столбцами с 1-го по r -й:

$$\bar{G}_1(k) = \begin{bmatrix} g_{1,r+1} & \dots & g_{1,kl} \\ g_{2,r+1} & \dots & g_{2,kl} \\ \dots & \dots & \dots \\ g_{l,r+1} & \dots & g_{l,kl} \end{bmatrix}$$

Теорема 15.2. Для того чтобы ЛА был автоматом СБПИК- k , необходимо и достаточно, чтобы $\bar{G}_1(k)$ была нулевой матрицей.

Доказательство. Восстановление

первого

символа

$$\bar{u}(0) = (u_1(0), \dots, u_l(0))'$$

неизвестного

слова

$\bar{u}(k) = \bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ по известной реакции $\bar{y}(k)$, как нетрудно понять, эквивалентно нахождению решения СЛАУ

$$[E_l \bar{G}_1(k)] \bar{u}(k) = \bar{b}_1 \quad (15.3)$$

где E_l - единичная матрица размерности $l \times l$, \bar{b}_1 - вектор, состоящий из первых l координат вектора \bar{b} . Если $\bar{G}_1(k) = [0]$, то система (15.3) в координатной форме записи примет вид

$$u_1(0) = b_1, u_2(0) = b_2, \dots, u_l(0) = b_l$$

Таким образом, $u_1(0), \dots, u_l(0)$, составляющие первый символ неизвестного входного слова ЛА, находятся однозначно. Отсюда следует достаточность условий теоремы.

Понятно, что если $\bar{G}_1(k) \neq [0]$, то система (15.3) имеет не единственное решение и, следовательно, однозначное восстановление неизвестного первого входного символа невозможно. Таким образом, условие теоремы является и необходимым.

Пользуясь теоремой 15.2, порядок СБПИК можно определить, проверяя равенство $\bar{G}_1(k) = [0]$ последовательно для $k = 0, 1, \dots$. Наименьшее значение k , для которого это равенство будет выполняться, и будет давать искомый порядок.

Проиллюстрируем изложенное на примере ЛА, заданного над полем $GF(2)$ следующими характеристическими матрицами ($n = m = 4, l = 2$):

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 1 & 1 \\ 0 & 1 \end{bmatrix},$$

$$C = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$D = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Пусть на этот ЛА подан неизвестный входной символ $\bar{u}(0) = (u_1(0), u_2(0))'$ и на его выходе наблюдается символ $\bar{y}(0) = (1, 0, 0, 0)'$.

Построим матрицу

$$G(0) = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{matrix} y_1(0) \\ y_2(0) \\ y_3(0) \\ y_4(0) \end{matrix}$$

Преобразование ее по методу Гаусса (с отбрасыванием нулевых строк) дает матрицу

$$G(0) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} y_1(0) \\ y_4(0) \\ y_2 \end{matrix}$$

Поскольку в $G(0)$ отсутствует строка вида $(0\ 1\ 0\ \dots)$, то символ $u_2(0)$ заведомо не восстанавливается по одному наблюдаемому выходному символу. Отсюда следует, что порядок рассматриваемого ЛА больше единицы.

Пусть на исследуемый ЛА подан еще один неизвестный входной символ и на него получена реакция $\bar{y}(1) = (0, 0, 0, 0)'$. Сформируем матрицу

$$G(1) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \begin{matrix} y_1(0) \\ y_2(0) \\ y_3(0) \\ y_4(0) \\ y_1(1) \\ y_2(1) \\ y_3(1) \\ y_4(1) \end{matrix}$$

Преобразование ее по методу Гаусса с отбрасыванием нулевых строк дает матрицу

$$\bar{G}(1) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{matrix} y_1(0) + y_4(1) \\ y_2(0) + y_2(1) \\ y_1(1) \\ y_4(1) \\ y_2(0) \end{matrix}$$

В последней матрице подматрица $\bar{G}_1(1) = [0]$. Из первых двух строк матрицы $\bar{G}(1)$ получаем искомый результат:

$$\begin{aligned} u_1(0) &= y_1(0) + y_4(1) = 1 + 0 = 1, \\ u_2(0) &= y_2(0) + y_2(1) = 0 + 0 = 0. \end{aligned}$$

Таким образом, порядок исследуемого ЛА равен 2, а при заданных реакциях на входное слово длины 2 первый символ этого слова есть $(1, 0)'$.

Остановимся теперь на вопросе о том, какой схемой можно восстановить неизвестный входной символ входного слова. Общий вид схемы, решающей эту задачу для ЛА СБПИК- k , изображен на [рис.15.1](#), где D - элемент единичной задержки. Она содержит m сдвиговых регистров, каждый из которых состоит не более чем из $k - 1$ элементов задержки. Выходы элементов задержки представляют собой значения $y_1(0), \dots, y_m(0), \dots, y_1(k - 1), \dots, y_m(k - 1)$, которые подаются на комбинационную схему с l выходами. Эта схема выполняет над перечисленными сигналами операции в соответствии с выражениями, полученными в первых l позициях последнего столбца матрицы $\bar{G}(k - 1)$.

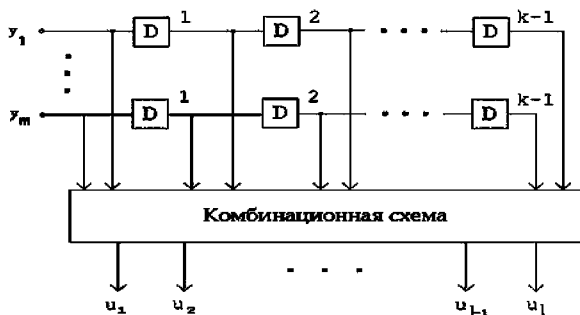


Рис. 15.1.

Понятно, что для ЛА над полем $GF(2)$ соответствующая комбинационная схема содержит только сумматоры по модулю 2, а ее выходами являются компоненты первого символа неизвестного входного слова.

Для рассмотренного примера описанная схема восстановления сигналов $u_1(0)$ и $u_2(0)$ представлена на [рис.15.2](#).

Подводя итоги изложенному, отметим, что по сравнению с автоматами Мили и Мура, являющимися в общем случае нелинейными, ЛА за счет их специфики позволяют ослабить требования, обеспечивающие восстановление неизвестного входного слова по наблюдаемой реакции. Отметим также, что сама процедура упомянутого восстановления для ЛА значительно менее трудоемка, чем для автоматов Мили. В самом деле, для ЛА она сводится к решению СЛАУ, для чего разработан целый ряд эффективных и простых в реализации методов, и не требует проведения никаких дополнительных экспериментов, необходимых для автоматов Мили.

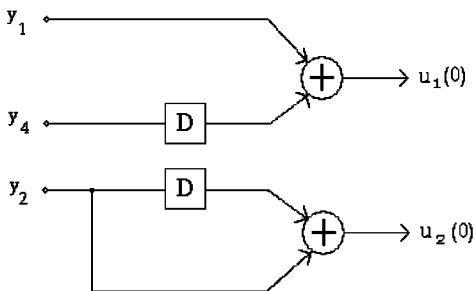


Рис. 15.2.

Кроме того, если для автоматов Мили критерии принадлежности их различным разновидностям автоматов БПИ сформулированы в [19] в терминах довольно громоздких конструкций так называемых проверочных графов, то для ЛА соответствующие критерии удалось сформулировать в терминах их характеристических матриц, довольно просто проверяемых.

Вопросы и упражнения

1. Приведите определение автомата без потери информации.
2. Сформулируйте критерий принадлежности автомата этому классу.

3. Определите класс линейных автоматов существенно без потери информации конечного порядка.
4. Сформулируйте критерий принадлежности автомата этому классу.
5. Опишите метод синтеза комбинационной схемы для восстановления первого символа неизвестного слова, поданного на вход автомата БПИК.

Обобщенные линейные автоматы без потери информации

Вводится класс обобщенных линейных автоматов без потери информации и даются условия принадлежности автомата этому классу. Определяется понятие его оптимального подавтомата, описывается процедура выделения такого подавтомата из ЛА и доказывается его единственность.

Рассмотрим автоматы, у которых входы и выходы являются многоканальными, т. е. их входные и выходные алфавиты структурированы. Каждый из таких автоматов может не быть автоматом БПИ в классическом смысле, но вместе с тем он может допускать возможность восстановления некоторого подмножества компонент входного символа при наблюдении реакций на подмножестве его выходных каналов. Такого рода автоматы были исследованы в работах [48], [49] и названы обобщенными автоматами БПИ (ОБПИ).

В этой лекции будут исследованы обобщенные линейные автоматы БПИ. В частности, в терминах характеристических матриц ЛА будет получен критерий принадлежности автомата упомянутому классу.

Введем некоторые понятия, которые понадобятся далее.

Выходной канал y_i ЛА A назовем избыточным, если в любой момент времени t значение сигнала $y_i(t)$ на нем есть линейная комбинация сигналов на остальных выходных каналах этого же автомата.

Линейный автомат назовем неизбыточным по выходам, если в нем отсутствуют избыточные выходные каналы.

Из сказанного выше следует, что множество $\tilde{Y} = \{y_{j1}, \dots, y_{jk}\}$ неизбыточных каналов представляет собой такое минимальное по мощности подмножество множества $Y = \{y_1, \dots, y_m\}$ всех выходных каналов ЛА, значений сигналов на которых достаточно для определения значений сигналов на каналах из множества $\frac{Y}{\tilde{Y}}$.

Теорема 16.1. Для того чтобы ЛА \tilde{A} был избыточным по выходам, необходимо и достаточно, чтобы число его выходных каналов было равно рангу матрицы $[C, D]$ системы (1.2) уравнений выходов этого автомата.

Доказательство. Предположим, что ранг матрицы системы линейных уравнений (10.2) относительно неизвестных

$s_1(t), \dots, s_n(t), u_1(t), \dots, u_l(t)$ равен ν , где $\nu < m$. Известно [19], что это условие является необходимым и достаточным для того, чтобы $m - \nu$ уравнений упомянутой системы представляли собой линейные комбинации остальных уравнений. Применительно к линейному автомату этот факт означает, что $m - \nu$ его выходных каналов являются избыточными. Отсюда и вытекает справедливость теоремы.

Опишем общий метод, позволяющий определить избыточные выходные каналы в ЛА. Этот метод основан на применении известного метода Гаусса [33] для решения систем линейных алгебраических уравнений.

Пусть Q - матрица системы (1.2). Методом Гаусса приведем матрицу Q к ступенчатой форме. В процессе такого приведения некоторые строки могут вырождаться в нулевые, что свидетельствует о линейной зависимости соответствующих уравнений системы (10.2) от других уравнений. Очевидно, что выходные каналы ЛА, которым соответствуют нулевые строки в ступенчатой матрице, и являются избыточными.

Проиллюстрируем изложенное на примере ЛА над полем $GF(2)$ с характеристическими матрицами

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Преобразование методом Гаусса матрицы $[C, D]$ в данном случае осуществляется за два шага, и полученные при этом матрицы имеют вид

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix} \begin{matrix} y_1 \\ y_2 \\ y_2 \oplus y_3 \\ y_4 \end{matrix} \quad \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{matrix} y_1 \\ y_2 \\ y_2 \oplus y_3 \\ y_2 \oplus y_3 \oplus y_4 \end{matrix}$$

Равенство нулю четвертой строки последней матрицы означает, что $y_2 \oplus y_3 \oplus y_4 = [0]$. Последнее равенство равносильно трем следующим: $y_2 = y_3 \oplus y_4$, $y_3 = y_2 \oplus y_4$, $y_4 = y_2 \oplus y_3$. Эти равенства говорят о том, что любой, но только один канал из множества $\{y_2, y_3, y_4\}$ в действительности можно считать избыточным. В самом деле, значение сигнала на одном из упомянутых каналов есть линейная комбинация сигналов на двух других каналах из этой тройки.

Теорема 16.2. Для того чтобы ЛА БПИ являлся неизбыточным по выходам, необходимо и достаточно, чтобы число его выходных каналов m было не меньше числа l его входных каналов.

Доказательство.

Необходимость. Пусть ЛА A есть БПИ и не имеет избыточных выходов. Покажем, что тогда $m \geq l$. Поскольку ЛА есть автомат БПИ, то по теореме 3.1 $D = l$. На основании теоремы 16.1 неизбыточность ЛА по выходам равносильна условию $[C, D] = m$. Из определения ранга матрицы следует, что $[C, D] \geq D$, т. е. $m \geq l$.

Достаточность. Пусть ЛА есть автомат БПИ и $m \geq l$. Докажем, что ЛА будет неизбыточен по выходам. Доказательство проведем от противного. Пусть $m < l$. Нетрудно показать, что

$$[C, D] \geq \max[C, D] \quad (16.1)$$

Поскольку ЛА есть БПИ, то $D = l$. С учетом неравенств $m < l$ и $C \leq m$, получаем

$$\max[C, D] = l$$

Таким образом, из (16.1) следует, что $[C, D] \geq l$. Поскольку $l \neq m$, то $[C, D] \neq m$ и по теореме 16.1 рассматриваемый ЛА является избыточным по выходам. Полученное противоречие завершает доказательство теоремы.

Введем теперь понятие обобщенного линейного автомата БПИ (ОБПИ).

Пусть $\bar{g} = (g_1, g_2, \dots, g_h)'$ - вектор-столбец, представляющий собой входной символ ЛА; i_1, i_2, \dots, i_ν - некоторые натуральные числа, где $i_a \neq i_b$, если $a \neq b$ и $1 \leq I \leq h$. Назовем проекцией вектора \bar{g} по входным каналам i_1, \dots, i_ν вектор-столбец $(g_{i_1}, \dots, g_{i_\nu})'$ и обозначим ее как $pr_{i_1, \dots, i_\nu} = barg$. Пусть

$\hat{g} = (\bar{g}(1), \bar{g}(2), \dots, \bar{g}(k))$ - упорядоченная последовательность вектор-столбцов, которую будем называть входным словом. Тогда проекцией слова \hat{g} по каналам i_1, \dots, i_ν назовем упорядоченную последовательность вектор-столбцов

$$(pr_{i_1, \dots, i_\nu} \bar{g}(1), \dots, pr_{i_1, \dots, i_\nu} \bar{g}(k)).$$

Рассмотрим следующую задачу. На ЛА A , находящийся в известном начальном состоянии $\bar{s}(0)$, подается неизвестное входное слово $\hat{a} = \bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ и по выходным каналам с номерами $j_1 < j_2 < \dots < j_\mu$, где $1 \leq j_i \leq m$ наблюдается выходная реакция. Требуется распознать проекцию слова \hat{a} по каналам с номерами i_1, i_2, \dots, i_ν , где $1 \leq i_j \leq l$

Условимся далее через $A(I, J)$, где $I = \{i_1, \dots, i_\nu\}$, $J = \{j_1, \dots, j_\mu\}$, обозначать такой подавтомат ЛА A , у которого:

- множество состояний, входной и выходной алфавиты и функция переходов те же, что у ЛА A ;
- выходной символ подавтомата в момент времени t есть $pr_{j_1, \dots, j_\mu} \bar{y}(t)$, где $\bar{y}(t)$ - выходной символ ЛА A в тот же

момент времени.

Понятно, что решение сформулированной выше задачи сводится к выяснению того, существует ли для заданного ЛА A такой подавтомат $A(I, J)$, для которого восстановление $pr_{i_1, \dots, i_v} \bar{u}(t)$ по наблюдаемой реакции $pr_{j_1, \dots, j_r} \bar{y}(t)$ возможно независимо от входного слова $\bar{u}(t)$ и действительного начального состояния ЛА A .

Далее такой подавтомат $A(I, J)$ будем называть обобщенным без потери информации (ОБПИ).

Существование у ЛА A подавтомата ОБПИ $A(I, J)$ содержательно означает, что исходный ЛА, не являясь автоматом БПИ в классическом смысле, тем не менее позволяет восстановить фрагменты его неизвестных входных слов по наблюдаемым фрагментам выходных слов и известному начальному состоянию.

Перейдем теперь к описанию процедуры, с помощью которой для заданного ЛА и заданного подмножества его выходных каналов можно установить, существует ли в ЛА такое подмножество I его входных каналов, что подавтомат $A(I, J)$ является ОБПИ.

Условимся далее выходные каналы с номерами из подмножества J называть наблюдаемыми, а каналы с номерами из подмножества $\bar{J} = \{1, 2, \dots, m\}$ - ненаблюдаемыми.

Пусть $\bar{J} = \{z_1, \dots, z_k\}$, т. е. выходные каналы с номерами z_1, z_2, \dots, z_k - ненаблюдаемые. Условимся, что строки всех характеристических матриц ЛА пронумерованы числами $1, 2, \dots$ сверху вниз, а столбцы - числами $1, 2, \dots$ слева направо. Напомним, что строкам $1, 2, \dots, h$ матриц A и B и столбцам $1, 2, \dots, h$ матриц A и C соответствуют компоненты s_1, \dots, s_n вектор-состояния, строкам $1, 2, \dots, m$ матриц C и столбцам $1, 2, \dots, l$ матриц B и D - компоненты u_1, \dots, u_l входного вектора \bar{u} .

Перейдем теперь к описанию упомянутой выше процедуры.

1. В матрице D вычеркиваются строки, соответствующие ненаблюдаемым выходам y_{z_1}, \dots, y_{z_k} . Оставшуюся часть матрицы размерности $(m - k) \times l$ обозначим через D_j . Находим в ней подматрицу с максимальным рангом, который обозначим через q . Пусть $\tilde{u}_1, \dots, \tilde{u}_q$ - переменные, соответствующие столбцам подматрицы с максимальным рангом q , которые будем называть потенциально восстанавливаемыми. Вычеркнем из D_j столбцы, соответствующие переменным $\tilde{u}_1, \dots, \tilde{u}_q$, и полученную матрицу обозначим $[D_j]$. Если $[D_j] = [0]$, то перейдем к пункту 2, иначе процедура завершается - при удалении выходов y_{z_1}, \dots, y_{z_k} не существует такого подмножества выходов I , что подавтомат $A(I, J)$ является ОБПИ.
2. В матрице C вычеркиваются строки, соответствующие ненаблюдаемым выходам y_{z_1}, \dots, y_{z_k} . Оставшуюся часть матрицы обозначим через C_j . Пусть $\tilde{s}_1, \dots, \tilde{s}_h$ - переменные, которым соответствуют столбцы в матрице C_j , содержащие нулевые элементы. Очевидно, что $\tilde{s}_1, \dots, \tilde{s}_h$ - компоненты вектор-состояния, которые необходимо знать для вычисления по известной реакции неизвестных компонент $\tilde{u}_1, \dots, \tilde{u}_q$ выходного вектора. Переходим далее к пункту 3 процедуры.
3. В матрице A выделим строки, соответствующие компонентам $\tilde{s}_1, \dots, \tilde{s}_h$ вектор-состояния, и полученную матрицу обозначим через A_j . Вычеркнем из нее столбцы, соответствующие переменным $\tilde{s}_1, \dots, \tilde{s}_h$, в результате чего получим матрицу $[A_j]$. Если $[A_j] = [0]$, то перейдем к пункту 4, иначе процедура завершается - при выходах y_{z_1}, \dots, y_{z_k} не существует такого подмножества выходов I , что подавтомат $A(I, J)$ является ОБПИ.
4. В матрице B выделим строки, соответствующие компонентам $\tilde{s}_1, \dots, \tilde{s}_h$, и полученную матрицу обозначим через B_j . Вычеркнем из нее столбцы, соответствующие переменным $\tilde{u}_1, \dots, \tilde{u}_q$, в результате чего получим матрицу $[B_j]$. Если

$[BJ] = [0]$, то подавтомат $A(I, J)$, где $I = \{\tilde{u}_1, \dots, \tilde{u}_q\}$, $J = \{z_1, \dots, z_k\}$ является ОБПИ, иначе в ЛА при заданном J не существует подавтомата ОБПИ. На этом процедура свою работу завершает.

Описанная процедура позволяет не только определить, существует ли такое множество I , что подавтомат $A(I, J)$ является ОБПИ, но и установить состав множества I . Построенный подавтомат $A(I, J)$ имеет в качестве входного вектор $[\tilde{u}_1, \dots, \tilde{u}_q]'$, в качестве выходного вектор $[y_{j_1}, \dots, y_{j_\mu}]'$, в качестве вектор-состояния $[\tilde{s}_1, \dots, \tilde{s}_h]'$, а его характеристическими матрицами являются матрицы A, B, C, D , из которых удалены все строки и столбцы, за исключением тех, которые соответствуют множествам переменных $\{\tilde{u}_1, \dots, \tilde{u}_q\}, \{\tilde{y}_1, \dots, \tilde{y}_\mu\}, \{\tilde{s}_1, \dots, \tilde{s}_h\}$.

Прокомментируем описанную процедуру. Условие $D_j = [0]$ в пункте 1 процедуры означает, что при вычислении реакций подавтомата $A(I, J)$ используются только потенциально восстанавливаемые переменные $\tilde{u}_1, \dots, \tilde{u}_q$ и никакие другие. Условия $[A_j] = [0]$ и $[B_j] = [0]$ в пунктах 3, 4 процедуры означают, что при вычислении компонент $\tilde{s}_1, \dots, \tilde{s}_h$ следующего состояния подавтомата $A(I, J)$ по формуле (1.1) в составляющей $A\bar{s}(t)$ задействуются только эти же компоненты, а в составляющей $B\bar{u}(t)$ - только потенциально восстанавливаемые компоненты $\tilde{u}_1, \dots, \tilde{u}_q$ входного вектора.

Учитывая эти комментарии, несложно провести строгое обоснование описанной процедуры.

Из описанной процедуры, с учетом использованных в ней обозначений, вытекает справедливость следующего утверждения.

Теорема 16.3. Для того чтобы у ЛА A существовал подавтомат ОБПИ $A(I, J)$, где I и J - непустые собственные подмножества множеств

входных и выходных каналов ЛА соответственно, необходимо и достаточно, чтобы матрица D содержала подматрицу ранга $|I|$, а матрицы $[D_j]$, $[B_j]$, $[A_j]$ были ненулевыми.

Выше уже отмечалось, что у заданного ЛА A в общем случае может существовать несколько различных подавтоматов ОБПИ $A(I, J)$. Наибольший интерес среди них представляет такой подавтомат, у которого подмножество I максимально, а подмножество J минимально по мощности. Далее такой подавтомат $A(I, J)$ будем называть оптимальным ОБПИ подавтоматом ЛА A .

Содержательно оптимальный подавтомат ОБПИ $A(I, J)$ дает возможность восстанавливать информацию максимального объема, поступающую на вход ЛА A , при минимальных затратах, определяемых числом наблюдаемых выходных каналов.

Восстановление фрагмента неизвестного входного слова $\bar{u}(0), \bar{u}(1), \dots$ оптимального подавтомата $A(I, J)$ осуществляется путем последовательного решения систем линейных алгебраических уравнений

$$D\bar{u}(t) = C\bar{s}(t) - \bar{y}(t)$$

для $t = 0, 1, \dots$. В силу теоремы 6.1 необходимым и достаточным условием разрешимости систем является условие $\text{rank} D = l$. Если $m > l$, то $m - l$ уравнений этих систем являются линейными комбинациями l остальных уравнений. Последнее означает, что для упомянутого восстановления в действительности требуется наблюдать не все m выходных каналов, а только l из них. По существу оставшиеся $-l$ каналов являются при этом избыточными.

Из этих рассуждений вытекает справедливость следующего утверждения.

Теорема 16.4. Если $A(I, J)$ является оптимальным ОБПИ подавтоматом ЛА A , то $|I| = |J|$.

Опишем теперь способ, позволяющий для заданного ЛА найти его оптимальный подавтомат ОБПИ.

Предположим, что исходный ЛА является БПИ, причем $m > l$. Из теоремы 16.4 следует, что в нем $m - l$ каналов при восстановлении неизвестного входного слова являются избыточными. Приведа методом Гаусса матрицу D ЛА к ступенчатой форме, найдем избыточные каналы и исключим их из дальнейшего рассмотрения. Очевидно, что полученный при этом подавтомат является оптимальным ОБПИ подавтоматом.

Если исходный ЛА не является автоматом БПИ, то оптимальный ОБПИ подавтомат, если таковой существует, можно найти методом перебора начиная с подавтомата $A(H, H)$, где $|H| = l$. Перебор будем осуществлять, например, следующим образом.

Начнем с попытки удаления у ЛА одного выходного канала $y_i, i = \overline{1, m}$. Если удаление одного очередного канала y_i дает ОБПИ подавтомат, то он и является искомым. Если же ни один из получаемых при этом подавтоматов не является ОБПИ, необходимо перейти к удалению из ЛА всевозможных пар выходов $\{y_i, y_j\}$, где $i \neq j$ и $1 \leq i, j \leq m$. Среди полученных на втором этапе подавтоматов может найтись ОБПИ, который и будет являться искомым оптимальным подавтоматом. Продолжим этот процесс далее аналогичным образом до тех пор, пока на очередном этапе либо не будет найден подавтомат ОБПИ, который и является искомым оптимальным ОБПИ, либо при исключении $m - 1$ выходов (рассматриваются все возможные сочетания из $m - 1$ выходов) ни один подавтомат не является ОБПИ. Последнее означает отсутствие у исходного ЛА подавтоматов ОБПИ.

При реализации описанного способа придется обращаться к описанной ранее в этом разделе процедуре отыскания по заданному множеству J выходных каналов такого множества I входных каналов, что подавтомат $A(I, J)$ является ОБПИ. В худшем случае таких обращений будет $\sum_{i=1}^{m-1} C_m^i$.

Проиллюстрируем описанный способ на примере ЛА над полем $CF(2)$, заданного следующими характеристическими матрицами:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Поскольку для рассматриваемого ЛА $l = 3$, поиск оптимального подавтомата ОБПИ начнем с анализа подавтоматов с тремя выходами. В связи с этим среди четырех выходных каналов этого ЛА необходимо исключить один.

Заметим: в начале раздела для приведенных матриц C и D было установлено, что в качестве избыточного выхода может быть принят как Y_4 , так и Y_3 и Y_2 . Для определенности остановимся на первом варианте. Тогда удаление выхода Y_4 приведет к удалению из матриц C и D четвертой строки.

Поскольку $|D| = 0$, то $D < 3$, следовательно, подавтомат с тремя выходами не является БПИ. Поэтому для поиска оптимального подавтомата ОБПИ воспользуемся перебором, организация которого была описана выше.

Положим $i = 1$; тогда удаление выхода Y_1 (или, что все равно, первой строки из матрицы D в соответствии с процедурой установления существования оптимального ОБПИ $A(I, J)$ при заданном множестве J) приводит к матрице

$$D_{Y_2 \cdot Y_3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Поскольку любые ее подматрицы размерности 2×2 имеют ранг, меньший 2, то условия теоремы 16.3 не выполняются и поэтому исключение выхода Y_1 не приведет к выделению подавтомата ОБПИ с двумя входными (выходными) каналами.

Положим $i = 2$; тогда удаление выхода Y_2 приводит к матрице

$$D_{y_1, y_2} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

По той же причине, что и в предыдущем случае, исключение выхода Y_2 также не приведет к выделению подавтомата ОБПИ.

Положим $i = 3$; тогда удаление выхода Y_3 приводит к матрице

$$D_{y_1, y_2} = \begin{array}{c} \begin{array}{ccc} u_1 & u_2 & u_3 \end{array} \\ \begin{bmatrix} \dots & \dots & \\ \vdots 0 & 1 & \vdots 0 \\ \vdots 1 & 0 & \vdots 0 \\ \dots & \dots & \end{bmatrix} \end{array}$$

Ее подматрица, выделенная пунктиром, имеет ранг 2, следовательно, потенциально восстанавливаемыми являются компоненты u_1 и u_2 входного вектора. Поскольку удаление упомянутой подматрицы из D_{y_1, y_2} приводит к нулевой матрице, перейдем к пункту 2 упомянутой выше процедуры.

Удалив из матрицы C третью строку, соответствующую ненаблюдаемому выходу Y_3 , получим матрицу

$$C_{y_1, y_2} = \begin{array}{c} \begin{array}{ccc} s_1 & s_2 & s_3 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$$

в которой столбцы s_1 и s_2 содержат ненулевые элементы. Следовательно, s_1 и s_2 - компоненты, которые необходимы для вычисления неизвестных u_1 и u_2 .

В соответствии с пунктом 3 процедуры построим матрицу $[A_{y_1, y_2}]$, удалив из A строки и столбцы с номерами 1, 2, которым соответствуют переменные s_1 и s_2

$$[A_{y_1, y_2}] = [00]'$$

Поскольку эта матрица нулевая, то в соответствии с процедурой переходим к пункту 4. Выделим в матрице B две первые строки, соответствующие компонентам s_1 и s_2 , и из полученной матрицы удалим два первых столбца, соответствующих переменным u_1 и u_2 . Оставшаяся матрица

$$[B_{y_1, y_2}] = [00]'$$

является нулевой. Построенные в процессе выполнения процедуры матрицы $[D_{y_1, y_2}]$, $[B_{y_1, y_2}]$, $[A_{y_1, y_2}]$, а также матрица D заданного ЛА удовлетворяют условиям теоремы 16.3. Таким образом, подавтомат, полученный за счет удаления выхода u_3 , является ОБПИ.

В соответствии с процедурой найденный оптимальный ОБПИ имеет следующие характеристические матрицы:

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, D = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Пусть, например, начальным состоянием рассматриваемого автомата является $s(0) = [1, 0, 1]'$, а на выходе наблюдается вектор $[y_1(0), y_2(0)]' = [0, 1]'$. Тогда по формуле (1.2) получаем

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \oplus \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix}$$

что в координатной форме дает систему

$$0 = 1 \oplus u_2(0), 1 = 0 \oplus u_1(0)$$

Отсюда получаем $u_1(0) = 1, u_2(0) = 1$. Вычислим теперь следующее состояние подавтомата по формуле (1.1):

$$\begin{bmatrix} s_1(1) \\ s_2(1) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} s_1(0) \\ s_2(0) \end{bmatrix} \oplus \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix}$$

Тогда в координатной форме получаем

$$s_1(1) = s_1(0) + u_1(0) = 1 \oplus 1 = 0, s_2(1) + u_2(0) = 1$$

Далее по аналогии с изложенным выше для $t = 1$ получаем систему

$$y_1(1) = s_1(1) \oplus u_2(1), y_2(1) = s_2(1) \oplus u_1(1)$$

которая при наблюдаемом, например, векторе $[0,1]'$ становится такой:

$$1 = 0 \oplus u_2(1), 1 = 1 \oplus u_1(1)$$

Отсюда $u_1(1) = 0, u_2(1) = 1$. Процесс восстановления последующих входных векторов может быть продолжен и далее аналогичным образом.

Заметим, что если у ЛА, для которого ищется оптимальный подавтомат, ОБПИ таков, что $m > l$, то этот подавтомат не единственен. В самом деле, пусть у ЛА имеются дублирующие друг друга выходы, т. е. по крайней мере один его выход является избыточным. В этом случае при исключении избыточных выходов множество неизбыточных выходов может быть разным по составу в зависимости от того, какие из дублирующих выходов оставлены и какие исключены. Отсюда следует, что такому ЛА соответствуют различные неизбыточные по выходам автоматы, отличающиеся друг от друга множествами выходных каналов. Очевидно, что построенные для них оптимальные подавтоматы ОБПИ также будут отличаться друг от друга составом выходных каналов.

Если заданный ЛА является неизбыточным по выходам, имеет место следующее утверждение.

Теорема 16.5. Если у неизбыточных по выходам ЛА A оптимальный подавтомат ОБПИ существует, то он единственен.

Доказательство. Проведем его методом от противного. Пусть для ЛА A , у которого $m = l$, существует два различных оптимальных подавтомата ОБПИ $A_1(I_1, J_1)$ и $A_2(I_2, J_2)$. Рассмотрим все возможные соотношения между подмножествами I_1, I_2, J_1, J_2 :

1. $I_1 \neq I_2, J_1 \cap J_2 = \emptyset$ или $J_1 \cap J_2 \neq \emptyset$.

В этом случае в силу оптимальности подавтоматов $A_1(I_1, J_1)$ и $A_2(I_2, J_2)$ соответствующие им системы уравнений $D\bar{u}(t) = C\bar{s}(t) - \bar{y}(t)$, из которых определяются неизвестные, сопоставляемые входным каналам подмножеств I_1 и I_2 , имеют единственное решение. Объединим эти системы в одну и будем рассматривать ее как систему относительно неизвестных, сопоставляемых входным каналам множества $I_1 \cup I_2$. Понятно, что последняя система также имеет единственное решение. Это означает, что у исходного ЛА существует такой подавтомат ОБПИ, который при наблюдении сигналов на каналах из множества $J_1 \cap J_2$ позволяет восстановить сигналы на входных каналах из множества $I_1 \cup I_2$. Поскольку $I_1 \neq I_2$, то $|I_1 \cup I_2| > |I_1|$ и $|I_1 \cup I_2| > |I_2|$, следовательно, подавтоматы A_1 и A_2 не являются оптимальными ОБПИ подавтоматами, что противоречит исходному предположению.

2. $I_1 = I_2$ и $J_1 \neq J_2$.

Как и в предыдущем случае, объединим две упоминавшиеся системы в одну и будем рассматривать ее как систему относительно неизвестных, соответствующих входным каналам множества $I_1(I_2)$. Поскольку $A_1(I_1, J_1)$ и $A_2(I_2, J_2)$ - оптимальные подавтоматы ОБПИ, ранг объединенной системы равен $|I_1|$. Из теоремы 16.4 следует, что для $A_1(I_1, J_1)$ и $A_2(I_2, J_2)$ выполняются равенства $|I_1| = |J_1|$ и $|I_2| = |J_2|$. Тогда из неравенства $J_1 \neq J_2$ вытекает, что $|J_1 \cup J_2| > |I_1|$ (или, что все равно, $|I_2|$).

Ввиду того что ранг матрицы D исходного ЛА равен $|I_1| = |I_2|$, а $|J_1 \cup J_2 - 2| > |I_1|$, в упомянутой объединенной системе имеется $|J_1 \cup J_2| - |I_1|$ уравнений, являющихся линейными комбинациями

других уравнений той же системы. Последнее означает, что исходный ЛА A , у которого существует два различных оптимальных ОБПИ подавтомата, имеет избыточные выходы, что противоречит условию теоремы.

Из изложенного следует, что в действительности возможен лишь случай $I_1 = I_2$ и $J_1 = J_2$, что и доказывает теорему.

Вопросы и упражнения

1. Дайте определение избыточного выходного канала линейного автомата.
2. Сформулируйте ранговый критерий отсутствия у линейного автомата избыточных выходных каналов.
3. Сформулируйте тот же критерий в терминах отношений между числом входных и выходных каналов линейного автомата БПИ.
4. Опишите общий метод нахождения избыточных выходных каналов ЛА.
5. Сформулируйте определение обобщенного линейного автомата БПИ.
6. Дайте определение обобщенного линейного подавтомата $A(I, J)$ линейного автомата A , где I и J - заданные подмножества входных и выходных каналов автомата A .
7. Опишите процедуру построения обобщенного линейного подавтомата $A(I, J)$.
8. Сформулируйте необходимые и достаточные условия существования обобщенного линейного подавтомата $A(I, J)$.
9. Дайте определение оптимального обобщенного линейного подавтомата $A(I, J)$ линейного автомата A .
10. В случае существования у ЛА оптимального обобщенного линейного подавтомата $A(I, J)$ является ли такой подавтомат единственным?

Минимизация времени восстановления неизвестных входных сигналов в сети из автоматов без потери информации

Введено понятие корректной сети из линейных автоматов без потери информации, и для нее исследуется задача минимизации времени восстановления неизвестных входных сигналов по наблюдаемым выходам. Предложен метод решения этой задачи, основанный на сведении ее к классической задаче о потоке минимальной стоимости из теории потоков.

Задача, упомянутая в заголовке лекции, возникает, в частности, при организации функционального контроля ЦУ. Как было сказано ранее, для этой цели могут быть использованы СВК, позволяющие восстанавливать неизвестные входные сигналы ЦУ по его выходной реакции, с последующим сравнением их с реальными сигналами, поступившими на вход ЦУ. Рассогласование этих сигналов свидетельствует о неправильном функционировании ЦУ и может быть зафиксировано в момент первого его проявления. Последнее обстоятельство очень важно, поскольку позволяет не допускать распространения появившейся ошибки.

Восстановление неизвестной входной последовательности по известному начальному состоянию автомата и наблюдаемой реакции в случае ЛА сводится, как известно [19], к решению системы линейных уравнений. Сложность решения такой системы, определяемая, например, необходимым для этого временем, зависит от размерности, а также от сложности вычисления выражений, определяющих значения правых частей системы, которая, в свою очередь, зависит от структуры схемы ЛА. Сеть, построенную из d штук ЛА БПИ над полем $GF(p)$, можно описать моделью одного ЛА БПИ, размерность которого равна

$q = \sum_{i=1}^d q_i$, где q_i - размерности входящих в состав сети автоматов-компонентов. Тогда восстановление неизвестных входных сигналов сети связано с решением системы линейных уравнений с w неизвестными, где w - число входов сети. Вместе с тем упомянутое восстановление, поскольку структура сети известна, может быть сведено к решению d систем линейных уравнений меньшей

размерности, каждая из которых соответствует отдельному компоненту сети. Понятно, что последний способ восстановления может оказаться более эффективным, т. е. требовать существенно меньшего времени на реализацию. В случае, когда время является критическим параметром, СВК, построенная вторым способом, будет предпочтительнее СВК, построенной первым способом, поскольку время решения систем уравнений определяет быстрдействие СВК.

Отметим, что СВК, синтезируемая вторым способом, определяется неоднозначно, поскольку восстановление одного и того же компонента входного сигнала сети может быть осуществлено по различным путям рассматриваемой сети и потому время восстановления этого компонента может быть различным в зависимости от выбранного пути. В связи с изложенным возникает задача минимизации общего времени восстановления входных сигналов сети из ЛА БПИ, которая рассматривается в этой лекции.

Пусть задано некоторое конечное множество ЛА $\tilde{A} = \{A_1, \dots, A_k\}$, которое далее именуется базисом, а каждый элемент этого множества - базисным. Предполагается, что сеть содержит в качестве компонентов ЛА $A_i \in \tilde{A}$. Помимо элементов сеть содержит входные и выходные полюсы. Из базисных элементов сети будут строиться по определенным правилам. Опишем эти правила, используя индукцию:

- совокупность первичных входных полюсов, соответствующих различным переменным сети, есть сеть, и все эти полюсы являются ее вершинами;
- результат присоединения к вершинам сети входов (необязательно всех) некоторого базисного элемента есть сеть; вершинами новой сети являются все вершины исходной сети, незадействованные (если таковые имеются) входы, а также выходы присоединенного элемента;
- результат присоединения к вершинам сети входов (кроме внешних полюсов) и/или выходов некоторого базисного элемента есть снова сеть, вершинами которой являются все вершины сети, полученной на предыдущем этапе, незадействованные (если таковые имеются) входы, а также все выходы присоединенного элемента.

В сети выделяется некоторое подмножество вершин, именуемых внешними выходными полюсами.

Сеть, в которой отсутствуют незадействованные входы используемых в ее составе базисных элементов, назовем корректной. Далее рассматриваются только корректные сети.

Из описанных правил построения сети следует, что с содержательной точки зрения корректной сети соответствует "правильная" электрическая схема, содержащая в общем случае параллельно и/или последовательно соединенные компоненты, а также линии обратной связи.

Далее предполагается, что у каждого компонента сети отсутствуют соединения его выходов с его же входами, т. е. отсутствуют петли обратной связи. Это предположение не является ограничительным, поскольку в случае существования у компонента такого вида соединения его можно отнести к его внутренней структуре, а сам компонент рассматривать как "черный ящик", содержащий только те входы и выходы, которые связывают его с другими базисными элементами сети или внешними входами (выходами).

Выходной канал y_i ЛА назовем избыточным, если в любой момент времени t значение сигнала $y_i(t)$ на нем есть линейная комбинация сигналов на других выходных каналах этого же ЛА.

Назовем ЛА неизбыточным по выходам, если в нем отсутствуют избыточные выходные каналы.

Напомним, что любой ЛА над полем $GF(p)$ задается системами уравнений состояний и выходов соответственно:

$$\bar{s}(t+1) = A\bar{s}(t) + B\bar{u}(t) \quad (17.1)$$

$$\bar{y}(t) = C\bar{s}(t) + D\bar{u}(t) \quad (17.2)$$

Выражение (17.2) можно интерпретировать как систему линейных уравнений относительно $n+l$ неизвестных $s_1(t), \dots, s_n(t), u_1(t), \dots, u_l(t)$, где n - размерность ЛА, l - число его входных каналов.

Лемма 17.1. Для того чтобы ЛА был неизбыточным по выходам, необходимо и достаточно, чтобы число его выходных каналов было равно рангу матрицы системы (17.2).

Доказательство. Пусть ранг системы (17.2) равен ν , где $\nu < m$ (где m - число выходных каналов ЛА). Известно из [33], что это условие является необходимым и достаточным для того, чтобы $m - \nu$ уравнений упомянутой системы были линейными комбинациями остальных уравнений. Применительно к ЛА этот факт означает, что $m - \nu$ его выходных каналов являются избыточными. Отсюда и вытекает справедливость леммы.

Лемма 17.2. Для того чтобы ЛА был БПИ и неизбыточным по выходам, необходимо и достаточно, чтобы выполнялись следующие условия:

1. число входных каналов равно числу выходных ($l = m$);
2. ранг матрицы D равен l .

Доказательств. В [19] доказано, что ЛА есть ЛА БПИ тогда и только тогда, когда ранг матрицы D равен l , т. е. второе условие леммы выполняется. Для ЛА БПИ начальное состояние $\bar{s}(0)$ считается известным. Зная $\bar{s}(0)$ и реакцию $\bar{y}(0)$, решив систему (17.2) относительно переменных $u_1(0), \dots, u_l(0)$, получим неизвестный входной символ $\bar{u}(0)$. Зная $\bar{u}(0)$, с помощью формулы (17.2) вычислим состояние $\bar{s}(1)$. Аналогичным образом можно определить состояние $\bar{s}(t)$ для $t = 2, 3, \dots$. Тогда система (17.2) для ЛА БПИ вырождается в систему

$$D\bar{u}(t) = \bar{y}(t) - C\bar{s}(t) \quad (17.3)$$

относительно неизвестных $\bar{u}_1(t), \dots, \bar{u}_l(t)$ с матрицей D . Отсюда на основании леммы 17.1 можно утверждать, что неизбыточность по выходам ЛА БПИ имеет место тогда и только тогда, когда число m его выходных каналов будет равно рангу матрицы D , т. е. числу l .

Далее рассматриваются сети, построенные только из ЛА БПИ,

неизбыточных по выходам. Очевидно, что любая сеть из ЛА БПИ такого типа представляет собой ЛА БПИ, но необязательно избыточный по выходам.

Перейдем теперь к постановке задачи. Рассмотрим сеть из ЛА БПИ, избыточных по выходам, в которой вершинами являются автоматы-компоненты и все ее внешние входы и внешние выходы. Каждому каналу (соединению) сети, связывающему две его вершины, поставим в соответствие целое неотрицательное число, называемое весовым коэффициентом. Значение весового коэффициента определяется типом канала. Если канал связывает выход некоторого автомата-компонента с входом другого автомата-компонента или некоторый вход сети с входом автомата-компонента, то соответствующий этому каналу весовой коэффициент положим равным числу тактов времени, необходимому для вычисления значения сигнала на упомянутом входе путем решения соответствующей системы линейных уравнений вида (17.3). В случае, когда канал связывает выход некоторого автомата-компонента с одним из внешних выходов сети, соответствующий ему весовой коэффициент положим равным нулю. Последнее объясняется тем, что значение сигнала на выходе сети может наблюдаться непосредственно, т. е. не требует никаких вычислений.

Если в сети существует путь от внешнего входа a к внешнему выходу b , то будем считать временем восстановления значения сигнала на входе a вдоль упомянутого пути сумму весовых коэффициентов каналов, входящих в путь. Множество путей в сети, связывающее каждый внешний вход сети с одним из ее внешних выходов, назовем покрывающим множеством. Очевидно, что знание покрывающего множества путей сети позволяет восстановить неизвестный входной вектор сети по наблюдаемой реакции сети и ее известному начальному состоянию. Назовем суммарные временные затраты на восстановление всех составляющих входного вектора сети вдоль всех путей покрывающего множества временем восстановления входного вектора.

Рассматриваемая здесь задача заключается в построении для заданной сети из ЛА БПИ покрывающего множества путей с минимальным временем восстановления входного вектора.

Перейдем к описанию метода решения сформулированной задачи. Идея

предлагаемого метода состоит в редукции ее к классической задаче о потоке минимальной стоимости в сети [3] порожденной исходной сетью. Обозначим исходную сеть, построенную из ЛА БПИ, избыточных по выходам, через \tilde{A} , а порожденную ею сеть - через A^* .

Опишем правила, по которым из исходной сети строится порожденная ею сеть. Условимся, что все входные и выходные каналы каждого автомата-компонента сети пронумерованы числами от 1 до \mathcal{G} , если он имеет \mathcal{G} входов

Каждому автомату-компоненту A_i сети \tilde{A} , имеющему \mathcal{G} входов (выходов), в порожденной сети \mathcal{G}^* соответствует \mathcal{G} вершин $A_i^1, A_i^2, \dots, A_i^{\mathcal{G}}$. Если в исходной сети ν -й выход автомата A_i соединен каналом связи с μ -м входом автомата A_j той же сети, то в порожденной сети соединим вершину A_j^μ с вершиной A_i^ν дугой, заходящей в A_i^ν (направление этой дуги противоположно направлению распространения сигнала в реальной сети по этому каналу). Далее, если в исходной сети ее внешний вход (внешний выход) a соединен с ν -м входом (выходом) автомата-компонента A_i , то в порожденной сети соединим вершину a с вершиной A_i^ν дугой, заходящей в вершину a (в вершину A_i^ν). Направление упомянутых дуг также противоположно направлению распространения сигнала по соответствующим каналам в реальной сети. К множеству вершин порожденной сети, описанному выше, добавим еще две вершины u и v . Вершину u соединим дугами со всеми вершинами порожденной сети, соответствующими внешним выходам исходной сети, заходящими в них. Аналогично, все вершины порожденной сети, соответствующие внешним входам исходной сети, соединим дугами с вершиной v , заходящими в нее.

Опишем теперь правила назначения пропускных способностей дугам порожденной сети:

1. пропускную способность каждой дуги вида (x, ν) порожденной сети положим равной 1;
2. пропускные способности всех остальных дуг порожденной сети

положим равными бесконечности.

Теперь остается описать правила назначения каждой дуге порожденной сети соответствующего весового коэффициента, который играет роль стоимости прохождения единицы потока в классической задаче о потоке минимальной стоимости.

Каждой дуге, исходящей из вершины u порожденной сети, назначим весовой коэффициент, равный 0.

Пусть автомат-компонент A_i сети имеет g входов (выходов). Как уже упоминалось выше, восстановление значений сигналов на его входах по наблюдаемой реакции при известном начальном состоянии требует решения соответствующей системы линейных уравнений. Пусть время решения такой системы равно a_i единицам времени, если в качестве единицы принять временной такт. Каждому автомату-компоненту A_i исходной сети с g входами соответствует g вершин $A_i^1, A_i^2, \dots, A_i^g$ порожденной сети. Положим значения весовых коэффициентов дуг, исходящих из упомянутых вершин, равными величине a_i .

Таким образом, по исходной сети \tilde{A} нами построена порожденная сеть A_* с одним истоком u и одним стоком v , где каждой дуге соответствует пара чисел, первое из которых есть весовой коэффициент, а второе - пропускная способность.

Отметим, что используемые ниже, но не определяемые здесь понятия и терминология из теории потоков в сетях трактуются так, как они определены в [34].

Установим два свойства порожденной сети, сформулированные в виде лемм.

Лемма 17.3. Величина максимального потока в порожденной сети A_* равна числу d внешних входов исходной сети \tilde{A} .

Доказательство. Рассмотрим разрез (X, v) сети A_* , где X - множество всех вершин порожденной сети, исключая вершину v . Поскольку пропускная способность каждой дуги этого разреза равна 1, а

из правил построения порожденной сети следует, что число дуг, входящих в разрез (X, ν) , равно числу d первичных входов исходной сети, то пропускная способность этого разреза равна d . Из правил построения порожденной сети следует, что никакой другой разрез порожденной сети не может содержать дуг меньше, чем d . Отсюда на основании теоремы о максимальном потоке и минимальном разрезе [34] вытекает справедливость леммы.

Лемма 17.4. Множеству всех увеличивающих цепей в порожденной сети A_* , обеспечивающих максимальное значение потока минимальной стоимости, соответствует покрывающее множество путей в исходной сети \tilde{A} с минимальным временем восстановления входного вектора.

Доказательство. Пусть в порожденной сети построен максимальный поток d , имеющий минимальную стоимость, который проходит по множеству увеличивающих цепей $\Pi = \{\Pi_1, \Pi_2, \dots, \Pi_j\}$. Из леммы 17.3 и того факта, что все дуги вида (x, ν) в порожденной сети имеют пропускную способность 1, следует, что число γ увеличивающих цепей в Π должно равняться d . Действительно, если на i -м шаге процесса решения задачи о потоке минимальной стоимости построена увеличивающая цепь, последняя дуга которой есть (A_b^a, ν) , то все другие увеличивающие цепи, появляющиеся в процессе решения на последующих шагах работы алгоритма, не могут содержать дуги (A_b^a, ν) . Это объясняется тем, что после i -го шага дуга (A_b^a, ν) является насыщенной и потому в дальнейшем она блокирует появление увеличивающих цепей, проходящих через нее.

Рассмотрим произвольную цепь $(u, A_{i_1}^{j_1}), (A_{i_1}^{j_1}, A_{i_2}^{j_2}), \dots, (A_{i_{k-1}}^{j_{k-1}}, A_{i_k}^{j_k}), (A_{i_k}^{j_k}, \nu)$ из множества Π . Поставим ей в соответствие путь в исходной сети $A_{i_k}(j_k), A_{i_{k-1}}(j_{k-1}), \dots, A_{i_1}(j_1)$, где $A_{i_l}(j_l), l = \overline{1, k}$, означает входной (или, что все равно, выходной) полюс J_l автомата-компонента A_{i_l} исходной сети. Из способа построения порожденной сети следует, что $A_{i_1}(j_1)(A_{i_k}(j_k))$ является внешним входом (выходом) сети. Таким

образом, приведенный путь указывает перечень входов автоматов-компонентов исходной сети и тот порядок, в котором следует восстанавливать сигналы на них с тем, чтобы в конечном итоге восстановить сигнал на внешнем входе сети. Поскольку среди конечных вершин множества упомянутых путей обязательно имеются все внешние входы исходной сети, то по определению оно является покрывающим. Очевидно, что в силу минимальной стоимости потока общее время восстановления входного вектора, осуществляемое вдоль путей построенного покрывающего множества, будет минимальным.

Леммы 17.3 и 17.4 служат обоснованием метода решения рассматриваемой задачи, состоящего из трех этапов. На первом этапе по заданной сети из ЛА БПИ строится порожденная ею сеть. Второй этап состоит в решении известным методом, описанным, например, в [34], задачи о потоке величины a минимальной стоимости, где a - число внешних входов исходной сети. На третьем этапе по найденному в процессе решения на втором этапе множеству увеличивающих цепей строится искомое множество покрывающих путей.

Проиллюстрируем предложенный метод на примере сети из ЛА БПИ, неизбыточных по выходам, которая изображена на [рис.17.1](#)

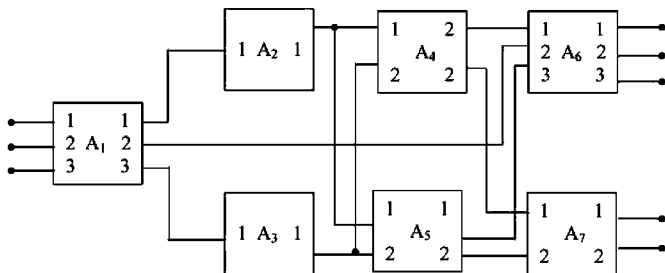


Рис. 17.1.

Построим по описанным выше правилам порожденную сеть, изображенную на [рис.17.2](#).

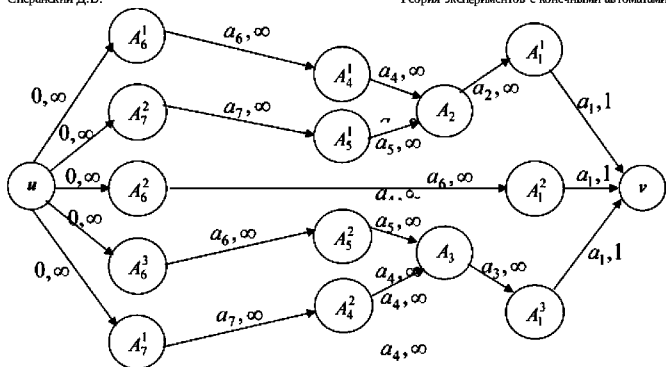


Рис. 17.2.

Каждой дуге этой сети поставим в соответствие упорядоченную пару чисел, из которых первое является весовым коэффициентом, а второе - пропускной способностью. Правила назначения дугам упомянутых значений описаны выше. Весовые коэффициенты на дугах рис.17.2 обозначены числами a_1, a_2, \dots, a_7 . Пусть для рассматриваемого примера эти значения таковы:

$a_1 = 3, a_2 = 1, a_3 = 1, a_4 = 2, a_5 = 2, a_6 = 3, a_7 = 2$. Теперь

приступим к решению задачи о максимальном потоке (для заданной сети его величина равна 3, поскольку сеть имеет 3 внешних входа) минимальной стоимости, воспользовавшись методом, описанным в [34]. Процесс решения отображен в приведенной таблице 17.1, которую кратко прокомментируем.

Столбцы этой таблицы со 2-го по 17-й соответствуют вершинным числам $p(u), p(A_6^1), \dots, p(A_1^3), p(v)$, которым на нулевом шаге присваивается значение 0. Перед началом работы окрашенным считается только исток сети, т. е. вершина u . Поскольку на всех дугах, исходящих из вершины u , стоимость прохождения единицы потока равна 0, имеют место равенства $p(A_6^1) - p(u) = \dots = p(A_7^2) - p(v) = 0$. Поэтому окрашиваются дуги, перечень которых приведен в столбце 18 таблицы нулевого шага, а

также концевые вершины этих дуг (столбец 19). Теперь перейдем к следующему шагу. Для всех неокрашенных вершин увеличиваем вершинные числа на 1, что отражено во 2-й строке таблицы (шаг 1). Однако на этом шаге перечень окрашенных дуг и вершин остается прежним, так как теперь ни для каких других пар вершин разность их вершинных чисел не совпадает со "стоимостью" на соответствующей дуге.

Таблица 17.1.

№ стр	n	$P(u)$	$P(A_6^1)$	$P(A_6^2)$	$P(A_3^6)$	$P(A_7^1)$	$P(A_7^2)$	$P(A_4^1)$	$P(A_4^2)$	$P(A_5^1)$
1	1	2	3	4	5	6	7	8	9	10
2	0	0	0	0	0	0	0	0	0	0
3	1	0	0	0	0	0	0	1	1	1
4	2	0	0	0	0	0	0	2	2	2
5	3	0	0	0	0	0	0	3	2	2
6	4	0	0	0	0	0	0	3	2	2
7	5	0	0	0	0	0	0	3	2	2
8	6	0	0	0	0	0	0	3	2	2
9	-	0	0	0	0	0	0	3	2	2
10	7	0	0	0	0	0	0	3	2	2
11	8	0	0	0	0	0	0	3	2	2
12	-	0	0	0	0	0	0	3	2	2
13	-	0	0	0	0	0	0	3	2	2

Продолжение таблицы

№ стр	$P(A_1^1)$	$P(A_1^2)$	$P(A_1^3)$	$P(\nu_6)$	Окрашенные дуги
1	14	15	16	17	18
2	0	0	0	0	$(u, A_6^1), (u, A_6^2), (u, A_6^3), (u, A_7^1), (u, A_7^2)$
3	1	1	1	1	те же
4	2	2	2	2	те же + $(A_7^1, A_4^2), (A_7^2, A_5^1)$

5	3	3	3	3	те же $+(A_6^1, A_4^1), (A_6^3, A_5^2), (A_6^2, A_1^2)$
6	5	3	5	5	те же $+(A_5^1, A_2), (A_4^2, A_3)$
7	5	3	5	5	те же $+(A_2, A_1^1), (A_3, A_3^1)$
8	5	3	5	5	те же $+(A_1^2, \nu)$
9	5	3	5	6	нет
10	5	3	5	7	все, исключая $(A_1^1, \nu), (A_1^2, \nu), (A_1^3, \nu)$
11	5	3	5	8	те же $+(A_1^1, \nu)$
12	5	3	5	8	нет
13	5	3	5	8	все, исключая $(A_1^1, \nu), (A_1^2, \nu)$

Далее процесс продолжается аналогичным образом до 6-го шага включительно. На этом шаге из окрашенных дуг может быть построена увеличивающая цепь $(u, A_6^2), (A_6^2, A_1^2), (A_1^2, \nu)$, по которой поток может быть увеличен на 1. Напомним, что в соответствии с используемым методом из [34] перед началом его работы значения потоков на всех дугах сети полагаются равными 0. Поскольку на 6-м шаге найдена увеличивающая цепь, перед выполнением следующего шага окраска снимается со всех дуг и со всех вершин, кроме истока u . На 8-м шаге из окрашенных дуг вновь может быть построена увеличивающая цепь $(u, A_7^2), (A_7^2, A_5^1), (A_5^1, A_2), (A_2, A_1^1), (A_1^1, \nu)$, по которой поток может быть увеличен на 1. Наконец, на 9-м шаге из окрашенных дуг строится увеличивающая цепь $(u, A_7^1), (A_7^1, A_4^2), (A_4^2, A_3), (A_3, A_1^3), (A_1^3, \nu)$, по которой поток также может быть увеличен на 1.

На этом алгоритм решения задачи о потоке минимальной стоимости заканчивает работу, поскольку в сети получен максимальный поток величины 3. Полученным выше трем увеличивающим цепям соответствуют три следующих пути в исходной сети, по которым должно производиться восстановление сигнала на первичных входах сети:

1. $A_6(2), A_1(2)$;

2. $A_7(2), A_5(1), A_2, A_1(1)$;
3. $A_7(1), A_4(2), A_3, A_1(3)$.

Время восстановления сигналов по этим путям на первичных входах сети соответственно равно 6, 8 и 8 тактам времени. Следовательно, общее минимальное время восстановления входного вектора равно 22 тактам. Целью рассмотрения этого простого примера является иллюстрация методики построения порожденной сети по исходной, а также возможности существования в сети альтернативных путей для восстановления сигнала на внешнем входе. Так, для входа A_1^1 сети существует еще один путь восстановления сигнала на нем - $A_6(1), A_4(1), A_2, A_1(1)$, кроме приведенного выше пути. Однако время восстановления сигнала по этому пути равно 9 тактам. Аналогичная ситуация имеет место и для внешнего входа A_1^3 : существует еще один путь - $A_6(3), A_5(2), A_3(2), A_3, A_1(3)$ - с временем восстановления 9 тактов.

В заключение заметим, что предложенный метод решения рассмотренной задачи может быть распространен и для сетей из ЛА БПИ, которые являются избыточными по входам. Для этого необходимо идентифицировать избыточные выходы в каждом автомате-компоненте и исключить их из дальнейшего рассмотрения. Такую идентификацию можно осуществить путем преобразования матрицы системы уравнений (17.2) по методу Гаусса решения систем линейных уравнений. Если в процессе преобразования некоторые строки матрицы вырождаются в нулевые, то это и является признаком линейной зависимости соответствующего выхода от других, т. е. его избыточности.

Вопросы и упражнения

1. Сформулируйте правила построения корректной сети из базисных ЛА.
2. Дайте определение покрывающего множества путей сети.
3. Приведите постановку задачи минимизации времени восстановления неизвестных входных сигналов в сети из

линейных автоматов без потери информации.

4. Какова основная идея метода решения задачи, сформулированной в пункте 3 ?
5. Опишите идею способа сведения исследуемой задачи к задаче о потоке минимальной стоимости из теории потоков.
6. Что такое порожденная сеть и каковы правила ее построения?
7. Опишите три этапа алгоритма решения задачи о минимизации времени восстановления неизвестных входных сигналов по наблюдаемым выходам.

Оптимальные эксперименты с линейными автоматами

Вводится понятие линейного автомата с взвешенным входным алфавитом в пространстве обобщенных состояний. Исследуется задача построения обобщенных синхронизирующих последовательностей с минимальным весом, переводящих ЛА из любого обобщенного состояния в заданное. Показано, что задача построения синхронизирующей последовательности минимального веса сводится к задаче целочисленного линейного программирования с линейными ограничениями. Исследуется задача построения обобщенной синхронизирующей последовательности с минимальным числом перепадов и описывается метод ее решения, который также основан на ее редукции к упомянутой задаче линейного программирования.

При исследовании различных видов экспериментов с автоматами в классической теории чаще всего рассматриваются задачи построения минимальных по длине входных последовательностей. Других критериев, кроме длины, для сравнения различных экспериментов в этой теории не используется. Это обстоятельство объясняется тем, что по умолчанию предполагаются равными затраты на подачу различных входных символов, из которых составлена соответствующая входная последовательность, подаваемая в процессе проведения эксперимента.

Вместе с тем упомянутые затраты в действительности могут быть различны, и при подаче двух последовательностей одинаковой длины, отличных по составу входных символов, суммарные затраты могут существенно отличаться. Это, в частности, происходит в том случае, когда моделью автомата описывается некоторый технологический процесс, включающий в качестве входных воздействий разного рода механические действия. Последние могут потребовать выполнения значительной работы, измеряемой, к примеру, в эргах, или затрат мощности, измеряемой в ваттах, и т. п.

В силу сказанного возникает проблема оптимального управления экспериментом по критериям минимизации общих затрат. Заметим, что некоторые эксперименты с автоматами, в частности, эксперименты по распознаванию состояний, можно рассматривать как процессы управления. Например, синхронизацию автомата можно трактовать как процесс управления автоматом путем подачи входной

последовательности, устанавливающей автомат в известное заключительное состояние независимо от того, в каком начальном состоянии он находился.

В теории оптимального управления как непрерывными [38], [40], [41], так и дискретными системами [14] [42], входные воздействия, обеспечивающие достижения цели управления, оцениваются по различным содержательным критериям.

Вообще говоря, подобного рода критерии могут быть использованы и при построении теории экспериментов с автоматами.

Хотя линейный автомат и является дискретной системой, однако его задание над конечным полем $GF(p)$ исключает возможность использования результатов упомянутой теории оптимального управления, поскольку последняя развита для случая вещественного и комплексного полей.

Оптимальные синхронизирующие эксперименты

В этом разделе рассматриваются ЛА, фазовые пространства которых состоят из обобщенных состояний (ОС), определенных в лекции 3. Условимся считать, что в вектор-столбце, представляющем ОС, неопределенными являются последние μ его координат.

Напомним, что ЛА, для которого существует обобщенная синхронизирующая последовательность (ОСП), называется обобщенно синхронизируемым. Состояние, в котором рассматриваемый ЛА оказывается после подачи ОСП, называется обобщенным синхросостоянием.

Поскольку перевод ЛА в одно и то же обобщенное синхросостояние может осуществляться несколькими ОСП, то введем критерий, по которому будем сравнивать различные ОСП.

Каждому входному символу \bar{u} ЛА поставим в соответствие действительное число $W(\bar{u})$, называемое весом символа. Весом

входной последовательности $u = \bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ назовем величину

$$W(\hat{u}) = \sum_{t=0}^k W(\bar{u}(t))$$

Содержательно вес входной последовательности u можно интерпретировать как суммарные затраты на ее подачу.

Рассмотрим следующую задачу. Пусть задан обобщенно синхронизируемый ЛА и некоторое обобщенное синхросостояние \bar{s} . Требуется найти ОСП минимального веса, переводящую ЛА из произвольного начального состояния в ОС \bar{s} . При этом предполагается, что множество допустимых начальных состояний ЛА совпадает со всем множеством его состояний.

Прежде чем перейти к решению этой задачи, исследуем некоторые свойства множества всех синхросостояний ЛА и его ОСП.

Условимся ОСП наименьшей длины для ЛА называть далее минимальной ОСП и обозначать ее длину через k_{min} .

Теорема 18.1. Если ЛА является обобщенно синхронизируемым, то для любого $k \geq k_{min}$ множество синхросостояний, порождаемых всеми ОСП длины k , совпадает с множеством синхросостояний, порождаемых всеми ОСП длины k_{min} .

Доказательство. Предположим, что ОСП $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ длины k переводит ЛА в обобщенное синхросостояние \bar{s} . Это означает, что

$$[A_{\mu}^k \bar{s}(0) + [A_{\mu}^{k-1}]_{\mu} \bar{u}(0) + \dots + [B]_{\mu} \bar{u}(k-1) = [\bar{s}]_{\mu} \quad (18.1)$$

Поскольку $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$ есть ОСП, то по теореме 1.18

$$[A^k]_{\mu} = [0] \text{ и тогда (18.1) примет вид}$$

$$[A^{k-1} B]_{\mu} \bar{u}(0) + \dots + [B]_{\mu} \bar{u}(k-1) = [\bar{s}]_{\mu} \quad (18.2)$$

ИЛИ

$$Q(k)u = [\bar{s}]_{\mu} \quad (18.2)$$

где

$$Q(k) = [A^{k-1}B, \dots, B]_{\mu}, u = [\bar{u}(0), \dots, \bar{u}(k-1)]'$$

Далее (18.2) будем интерпретировать как СЛАУ относительно неизвестных, являющихся координатами вектора \hat{u} . Как известно из алгебры [33], необходимым и достаточным условием разрешимости СЛАУ является представление столбца свободных членов $[\bar{s}]_{\mu}$ в виде линейной комбинации линейно независимых столбцов матрицы $Q(k)$ системы (18.2).

Поскольку для любого $k \leq k_{min}$

$$[A_{\mu}^k = [A^{k_{min}} A^{k-k_{min}}]_{\mu} = [A^{k_{min}}]_{\mu} A^{k-k_{min}} = [0]$$

то

$$Q(k) = [[0], \dots, [0], A^{k_{min}-1}B, \dots, B]_{\mu} \quad (18.3)$$

Таким образом, вектор $[\bar{s}]_{\mu}$ является линейной комбинацией линейно независимых столбцов матрицы (18.3) или, что то же самое, линейно независимых столбцов матрицы $Q(k_{min})$.

Следствие 1. Мощность множества всех различных синхросостояний ЛА, заданного над полем $GF(p)$, равна величине $p^{Q(k_{min})}$.

Заметим, что нулевое обобщенное синхросостояние всегда входит во множество всех синхросостояний ЛА, поскольку при подаче нулевой входной последовательности длины k_{min} обобщенно синхронизируемый ЛА переходит в ОС $[0]_{\mu}$.

Следствие 2. Если $[B]_{\mu} \neq [0]$, то обобщенно синхронизируемый ЛА

имеет ненулевое обобщенное синхросостояние .

Это вытекает из того, что если $[B]_{\mu} \neq [0]$, то $[A^{k_{min}-1}B, \dots, B]_{\mu} \geq 1$, но тогда $p^{rank Q(k_{min})}$ для любого p .

Что касается определения множества всех обобщенных синхросостояний, то, как это следует из теоремы 18.1, оно сводится к нахождению линейного подпространства, порожденного базисом матрицы $Q(k)$.

Теорема 18.2. Пусть u_{min} - минимальная ОСП, а u - произвольная ОСП длины $k \geq k_{min}$, переводящая ЛА в одно и то же синхросостояние, и пусть $W(\bar{u}) \geq 0$ для любого входного символа этого ЛА. Тогда $W(u_{min}) \leq W(\bar{u})$.

Доказательство. Предположим, что ОСП $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k-1)$, где $k \geq k_{min}$ переводит ЛА в обобщенное синхросостояние \bar{s} . Это означает, что

$$[A^{k-1}B]_{\mu}\bar{u}(0) + \dots + [B]_{\mu}\bar{u}(k-1) = [\bar{s}]_{\mu}$$

Учитывая, что в силу обобщенной синхронизируемости для всех $k \geq k_{min}$ справедливо равенство $[A^k]_{\mu} = [0]$, получаем

$$[0] + \dots + [0] + [A^{k_{min}-1}B]_{\mu}\bar{u}(k-k_{min}) + \dots + [B]_{\mu}\bar{u}(k-1) = [\bar{s}]_{\mu}$$

где слагаемые, содержащие A_i при $I \geq k_{min} - 1$, равны $[0]$. Из полученного равенства следует, что ОСП $\bar{u}(k-k_{min}), \dots, \bar{u}(k-1)$ длины k_{min} переводит ЛА в обобщенное синхросостояние \bar{s} . В силу неотрицательности весовой функции $W(\bar{u})$ отсюда следует утверждение теоремы.

Из этой теоремы вытекает следующий вывод: если весовая функция $W(\bar{u})$ является неотрицательной, то минимальную по весу ОСП рассматриваемого обобщенно синхронизируемого ЛА следует искать

среди ОСП минимальной длины.

В случае, когда весовая функция $W(\bar{u}) < 0$ по крайней мере для одного входного символа ЛА, сформулированная в начале этого раздела задача имеет решение только тогда, когда длина искомой ОСП предполагается заранее заданной. Легко показать, что если на длину ОСП ограничений не накладывать, то для этого ЛА можно построить ОСП, вес которой будет меньше любого наперед заданного отрицательного числа.

Далее предполагается, что весовая функция $W(\bar{u})$ является неотрицательной и матрица $Q(k)$ в СЛАУ (18.2), обозначаемая далее как Q , соответствует $k = k_{min}$.

Вернемся теперь к рассматриваемой задаче. Множество всех ОСП, переводящих ЛА в обобщенное состояние $[\bar{s}]_{min}$, как это следует из сказанного выше, должно удовлетворять равенству

$$D\hat{u} = [s]_{\mu}$$

Здесь

вектор-столбец

$\hat{u} = [u_1(0), \dots, u_l(0), \dots, u_1(k_{min}), \dots, u_l(k_{min})]'$. В соответствии с формулировкой задачи искомое решение должно доставлять минимум весовой функции $W(\hat{u})$. Учитывая, что переменные $u_i(j)$, $i = \overline{1, l}$, $j = \overline{0, k_{min}}$ по смыслу задачи являются целыми неотрицательными числами, не превосходящими характеристику p поля $GF(p)$, рассматриваемую задачу можно описать в следующем виде:

$$W(\hat{u}) \rightarrow \min \quad (18.4)$$

$$\begin{aligned} D\hat{u} &= [\bar{s}]_{\mu}, \\ 0 \leq u_i(j) &\leq p - 1, 1 \leq i \leq l, 0 \leq j \leq k_{min} \end{aligned} \quad (18.5)$$

Условимся для упрощения обозначений координаты вектора \hat{u} считать перенумерованными сверху вниз натуральными числами от 1 до lk_{min} . Тогда последнее неравенство примет вид

$$0 \leq u_i \leq p - 1, 1 \leq i \leq lk_{min} \quad (18.6)$$

Сформулированная задача относится к классу задач математического программирования. Скажем более точно: она представляет собой задачу целочисленного программирования [35] с линейными ограничениями (18.5).

Перепишем систему ограничений (18.5) в виде системы сравнений

$$Q\hat{u} = [\bar{s}]_{\mu} \text{ mod } p \quad (18.7)$$

Как известно, сравнение $a \equiv b \text{ mod } p$ эквивалентно равенству $a - b = pd$ для некоторого целого d . Поэтому (18.7) эквивалентна СЛАУ в целых числах:

$$Q\hat{u} = [\bar{s}]_{\mu} + p\bar{d}$$

где $\bar{d} = (d_1, \dots, d_{\mu})'$ - вектор-столбец, число координат которого равно числу уравнений в системе (18.7). Таким образом, поставленная задача эквивалентна следующей задаче целочисленного программирования:

$$W(\hat{u}) \rightarrow \min \quad (18.8)$$

$$Q\hat{u} = [\bar{s}]_{\mu} + p\bar{d} \quad (18.9)$$

$$0 \leq u_i \leq p - 1, 1 \leq i \leq lk_{min} \quad (18.10)$$

Заметим, что из (18.9) нетрудно установить диапазон изменения координат вектора \bar{d} :

$$0 \leq d_i \leq lk_{min}(p - 1), 1 \leq i \leq \mu$$

Подведя итоги изложенного, сформулируем следующее утверждение.

Теорема 18.3. Задача построения ОСП минимального веса, переводящей обобщенно синхронизируемый ЛА в заданное обобщенное синхросостояние, всегда может быть сведена к задаче целочисленного программирования с линейными ограничениями.

В настоящее время известен ряд методов решения задач целочисленного программирования [35], в том числе метод сечения, метод разветвленного поиска и оценки и др. Остановимся кратко на том, как может быть адаптирован последний из названных методов для исследуемой нами задачи, поскольку ее специфика позволяет такую адаптацию осуществлять эффективно.

Напомним кратко идею метода разветвленного поиска посредством разделения и оценки, относящегося к группе методов ветвей и границ, широко применяемых в целочисленном программировании

Предположим, что поставлена следующая задача:

$$W(\hat{u}) \rightarrow \min_{\text{при } \hat{u} \in U}$$

где U - некоторое конечное множество целочисленных векторов. Используя верхние оценки для координат вектора \hat{u} , можно с помощью известных методов [28] свести задачу к двоичной. Поэтому далее будем считать, что U есть множество двоичных векторов.

Используя дискретность множества U , представим его в форме некоторого разветвления. Вершине нулевого уровня, т. е. корню ветвления, соответствует все множество U . Для построения вершины уровня 1 выберем некоторую переменную u_i . Этот уровень содержит две вершины, которые соответствуют следующим подмножествам U : подмножество векторов U_{i_0} , для которых $u_i = 0$, и подмножество U_{i_1} , для которых $u_i = 1$. Понятно, что эти два подмножества образуют разбиение множества U . Будем говорить, что множество U разделено относительно переменных u_i . Аналогично, для построения уровня 2 выберем вторую переменную u_j и разделим каждое из полученных на предыдущем этапе подмножеств относительно переменной u_j . Таким образом, на уровне 2 получим четыре подмножества: подмножество векторов $U_{i_0j_0}$, для которых $u_i = 0, u_j = 0$, подмножество векторов $U_{i_1j_0}$, для которых $u_i = 1, u_j = 0$, подмножество векторов $U_{i_0j_1}$, для которых $u_i = 0, u_j = 1$, и подмножество векторов $U_{i_1j_1}$, для которых $u_i = 1, u_j = 1$. Аналогичным образом строятся уровни 3, ..., N , где N - число переменных задач. Каждой вершине

построенного дерева соответствует конкретный двоичный вектор из U .

Перейдем теперь к понятию оценки. Предположим, что для каждой вершины U' разветвления удалось получить оценку снизу для лучшего решения из множества U' :

$$f(U') \leq \min_{\hat{u} \in U'} W(\hat{u})$$

Функцию f будем называть далее функцией оценки. Построим шаг за шагом, начиная с уровня 0, разветвления множества U . Пусть для некоторого допустимого вектора u_0 известно значение $W(u_0)$. Предположим, что для некоторой вершины U' построенного разветвления имеет оценку

$$f(U') > W(u_0)$$

Следовательно, по определению f множество U' не содержит оптимального решения задачи. Это позволяет избежать исследования всех вершин разветвления, следующего за u' . Используя данный принцип, можно значительно уменьшить перебор элементов допустимого множества, что весьма существенно при большом числе переменных.

Заметим, что выбор множества, которое необходимо разделить на очередном этапе, и выбор переменной для разделения в общем случае произволен. Вместе с тем осуществлять такой выбор в каждом конкретном случае необходимо с учетом специфики задачи. Для получения оценок весовой функции часто решается соответствующая непрерывная задача линейного программирования, что является достаточно трудоемким этапом. Тот факт, что при $p = 2$ все коэффициенты, кроме свободных членов в ограничениях для u_j , равны 0 или 1, существенно облегчает решение соответствующих задач. Этот случай часто встречается на практике, поскольку ЛА над полем $GF(2)$ является адекватной моделью различных широко распространенных на практике цифровых устройств. Ниже будет показано, что специфика конкретной задачи позволяет иногда избежать этого этапа, вычисляя функции оценки из других соображений.

Синтез обобщенной синхронизирующей последовательности с минимальным числом перепадов

Ниже рассматривается задача, аналогичная задаче предыдущего раздела, где вместо минимизации весовой функции $W(u)$ будет осуществляться минимизацией иной функции.

Предварительно дадим краткое содержательное пояснение. Известно, что для цифровых устройств, реализованных с применением так называемой К-МОП- технологии, из двух равных по длине входных последовательностей более предпочтительной является та из них, у которой меньше число перепадов. Под перепадом сигналов в последовательности понимается наличие ситуации, когда два соседних символа в этой последовательности инверсны друг другу. К примеру, последовательность 00101011 имеет 5 перепадов.

Введем в рассмотрение следующую функцию:

$$G(a, b) = \begin{cases} 0, & \text{если } a = b, \\ 1, & \text{если } a \neq b \end{cases}$$

Если a и b - два символа некоторой входной последовательности, то по значению $G(a, b)$ можно судить о наличии или отсутствии перепада сигналов. Тогда задача построения ОСП $\hat{u} = u_1, u_2, \dots, u_{lk_{min}-1}$ с минимальным числом перепадов сигналов, переводящей обобщенно синхронизируемый автомат в заданное ОС $[\bar{s}]_\mu$, формулируется следующим образом:

$$\sum_{i=1}^{lk_{min}-1} G(u_i, u_{i+1}) \rightarrow \min \quad (18.11)$$

$$Q\hat{u} = [\bar{s}]_\mu + p\bar{d} \quad (18.12)$$

$$0 \leq u_i \leq p - 1, 1 \leq i \leq lk_{min} \quad (18.13)$$

Заметим, что количество переменных в этой задаче может быть уменьшено. Действительно, некоторые элементы матрицы Q могут

быть нулевыми и потому в левой части системы (18.12) соответствующие переменные будут отсутствовать. Далее эти переменные будем именовать независимыми, а остальные переменные - зависимыми. Предположим, что в оптимальном плане \hat{u} задач (18.11)-(18.13) u_i и u_j есть значения зависимых координат, а координаты с номерами $i+1, i+2, \dots, j-1$ являются независимыми. Тогда если $u_i = u_j$, то на участке $(u_i, u_{i+1}, \dots, u_{j-1}, u_j)$ входной последовательности \hat{u} перепады сигналов отсутствуют, а если $u_i \neq u_j$, то имеется ровно один перепад. Понятно, что если на упомянутом участке входной последовательности \hat{u} исключить переменные $u_i, u_{i+1}, \dots, u_{j-1}$, то значение целевой функции (18.11) на укороченной таким образом входной последовательности не изменится. Поэтому целевую функцию (18.11) можно рассматривать на множестве векторов, составленных только из зависимых координат вектора \hat{u} . Видоизмененную таким образом целевую функцию будем называть приведенной, так же будем называть и соответствующую задачу.

Ясно, что минимальное значение целевой функции в приведенной задаче равно ее минимальному значению в исходной задаче (18.11)-(18.13). Оптимальный план исходной задачи можно получить из оптимального плана приведенной задачи, если зависимые переменные положить равными соответствующим переменным из решения приведенной задачи, а значение каждой независимой переменной u_j положить равным значению зависимой переменной u_i с максимальным номером i , таким, что $i < j$. Если такой номер отсутствует, т. е. все переменные u_i, u_{i+1}, \dots, u_j являются независимыми, то положим их равными значению зависимой переменной u_j с таким минимальным номером i , что $i < j$.

Из вида целевой функции (18.11) следует, что она не является линейной. Вместе с тем покажем, что исходная задача (18.11)-(18.13) может быть сведена к линейной задаче за счет введения дополнительных переменных.

Рассмотрим задачу

$$\sum_{i=1}^{lk_{min}-1} g_i \rightarrow \min \quad (18.14)$$

со следующими ограничениями:

$$Q\hat{u} = [\bar{s}]_{\mu} + p\bar{d} \quad (18.15)$$

$$- \leq u_i \leq p - 1, 1 \leq i \leq lk_{min} \quad (18.16)$$

$$-g_i(p-1) \leq u_i - u_{i+1} \leq g_i(p-1), g_i \in \{0, 1\}, 1 \leq i \leq lk_{min} - 1$$

Введем следующее обозначение: $\bar{g} = (g_1, g_2, \dots, g_{lk_{min}-1})$.

Теорема 18.4. Если $(\hat{u}, \bar{d}, \bar{g})$ - оптимальный план задачи (18.14)-(18.17), то (\hat{u}, \bar{d}) является оптимальным планом задачи (18.11)-(18.13) и значения (18.14) и (18.11) на соответствующих оптимальных планах совпадают.

Доказательство. Пусть (\hat{u}, \bar{d}) есть оптимальный план задачи (18.11)-(18.13). Положим $g_i = G(u_i, u_{i+1})$ при $1 \leq i \leq lk_{min} - 1$, тогда для построенного таким образом множества величин g_i выполняются неравенства (18.17). В самом деле, если $u_i = u_{i+1}$, то $g_i = 0$ и неравенство (18.17) выполнено; если $u_i \neq u_{i+1}$, то $g_i = 1$ и неравенство (18.17) принимает вид

$$-p + 1 \leq u_i - u_{i+1} \leq p - 1$$

При ограничении (18.16) на значения u_i последнее неравенство также выполняется. Таким образом, поскольку ограничения на вектор \hat{u} в задачах (18.11)-(18.13) и (18.14)-(18.17) совпадают, то для последней задачи $(\hat{u}, \bar{d}, \bar{g})$ является допустимым планом, причем значения целевых функций соответствующих задач на двух приведенных планах совпадают.

Итак, всякому допустимому плану (\hat{u}, \bar{d}) задачи (18.11)-(18.13) соответствует допустимый план $(\hat{u}, \bar{d}, \bar{g})$ задачи (18.14)-(18.17) с одинаковыми значениями целевых функций, а на всяком оптимальном

плане $(\hat{u}, \bar{d}, \bar{g})$ задачи (18.14)-(18.17) значение целевой функции (18.14) совпадает со значением целевой функции (18.11) на плане (\hat{u}, \bar{d}) задачи (18.11)-(18.13). Отсюда следует, что если $(\hat{u}, \bar{d}, \bar{g})$ есть оптимальный план задачи (18.14)-(18.17), то (\hat{u}, \bar{d}) является оптимальным планом задачи (18.11)-(18.13). В самом деле, в противном случае существует план (\hat{u}_1, \bar{d}_1) , для которого $W(\hat{u}_1) < W(\hat{u})$, но тогда план $(\hat{u}_1, \bar{d}_1, \bar{g}_1)$ лучше плана $(\hat{u}, \bar{d}, \bar{g})$, что противоречит выбору.

Из теоремы 18.4 следует, что рассматриваемая нами задача (18.11)-(18.13) может быть сведена к линейной задаче и по найденному оптимальному плану последней легко построить оптимальный план исходной задачи.

Проиллюстрируем изложенное на примере ЛА, заданного над полем $GF(2)$ следующими характеристическими матрицами:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Система уравнений перехода в координатной форме для этого ЛА имеет следующий вид:

$$\begin{aligned} s_1(t+1) &= u_1(t), \\ s_2(t+1) &= s_1(t) + u_2(t), \\ s_3(t+1) &= s_1(t) + s_2(t) + u_3(t), \\ s_4(t+1) &= s_1(t) + s_2(t) + s_3(t) + u_4(t). \end{aligned}$$

Пусть $\mu = 3$ и требуется установить этот ЛА в обобщенное состояние $\bar{s} = (1, 1, 1, x)'$. Начнем с проверки того, является ли заданный ЛА обобщенно синхронизируемым. Вычислим с этой целью матрицы

$$A^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Отсюда видно, что необходимое условие обобщенной синхронизируемости $[A]_\mu$ выполняется при $k = 3$ и не выполняется при $k < 3$. Следовательно, минимальная ОСП для данного ЛА имеет длину 3.

Найдем матрицу линейных ограничений на допустимый план

$$Q = [A^2B, AB, B]_3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & & & & & & & & \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & & & & & & & & \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & & & & & & & & \end{bmatrix}$$

Система сравнений $Q\hat{u} = [\bar{s}]_\mu \text{mod} 2$, представляя собой линейные ограничения на переменные задачи, принимает вид

$$Q\hat{u} = (1, 1, 1) \text{mod} 2$$

где

$$\hat{u} = (u_1(0), \dots, u_4(0), u_1(1), \dots, u_4(1), u_1(2), \dots, u_4(2))' = (u_1, u_1, \dots, u_{12})'$$

Перепишем эту систему в координатной форме:

$$u_1(2) \equiv 1 \text{mod} 2$$

$$u_1(1) + u_2(2) \equiv 1 \text{mod} 2$$

$$u_1(0) + u_1(1) + u_2(1) + u_3(2) \equiv 1 \text{mod} 2$$

В соответствии с изложенным выше эквивалентная система линейных алгебраических уравнений (18.15) примет вид (с учетом перенумерации переменных от 1 до 12)

$$\begin{aligned}u_9 - 2d_1 &= 1, \\u_5 + u_{10} - 2d_2 &= 1, \\u_1 + u_5 + u_6 + u_{11} - 2d_3 &= 1,\end{aligned}$$

где $u_j \in \{0, 1\}$, d_i - целые неотрицательные числа, $2 \leq i \leq 3, 1 \leq j \leq 12$.

Из полученных уравнений вытекает, что $d_1 = d_2 = 0, d_3 \leq 1$. Тогда исходная задача построения ОСП с минимальным числом перепадов сигналов для заданного ЛА эквивалентна задаче целочисленного линейного программирования:

$$\begin{aligned}\sum_{i=0}^{11} g_i &\rightarrow \min, \\u_9 &= 1, \\u_5 + u_{10} &= 1, \\u_1 + u_5 + u_6 + u_{11} - 2d_3 &= 1, \\i \leq u_i - u_{i+1} \leq g_i, \dots, g_i, u_i, d_3 &\in \{0, 1\}, 1 \leq i \leq 12\end{aligned}$$

Заметим, что в данном случае линейность задачи не столь существенна, поскольку на этапе оценки решать непрерывные задачи большой размерности с помощью симплекс-метода не потребуется. Поэтому применим к исходной задаче метод, изложенный в предыдущем разделе. Эта задача выглядит так:

$$\begin{aligned}\sum_{i=0}^{11} G(u_i, u_{i+1}) &\rightarrow \min, \\u_9 - 2d_1 &= 1, \\u_5 + u_{10} - 2d_2 &= 1, \\u_1 + u_5 + u_6 + u_{11} - 2d_3 &= 1\end{aligned}$$

где $d_3, u_j \in \{0, 1\}$, d_i - целые неотрицательные числа, $1 \leq i \leq 12$. Формально эта задача имеет 13 переменных.

Способом, описанным выше, сократим количество переменных до 6, оставляя для рассмотрения лишь зависимые переменные

$$u_1, u_5, u_6, u_9, u_{10}, u_{11}$$

Таким образом, приведенная задача примет вид

$$\begin{aligned} W(\hat{u}) &= G(u_1, u_5) + G(u_5, u_6) + G(u_6, u_9) + G(u_9, u_{10}) + G(u_{10}, u_{11}) \rightarrow \min, \\ u_9 - 2d_1 &= 1, \\ u_5 + u_{10} - 2d_2 &= 1, \\ u_1 + u_5 + u_6 + u_{11} - 2d_3 &= 1, \end{aligned}$$

где

$$d_3, u_i \in \{0, 1\}, 1 \leq i \leq 12$$

Из условия $u_5 + u_{10} = 1$ следует, что $u_5 \neq u_{10}$, поэтому среди пар (u_5, u_6) , (u_6, u_9) , (u_9, u_{10}) соседних координат вектора приведенной задачи имеется не менее одного перепада.

Таким образом, получаем оценку $W(\hat{u}^*) \geq 1$ для оптимального плана \hat{u}^* . Если существует план u , для которого $W(u) = 1$, то он может быть выбран в качестве оптимального. Попытаемся найти такой план.

Разделим множество двоичных векторов задачи относительно переменной d_3 . При $d_3 = 1$ получаем $u_1 + u_5 + u_6 + u_{11} = 3$, из чего следует, что ровно одно из чисел равно 0, остальные равны 1.

Если $u_1 \neq u_5$, то среди пар (u_1, u_5) , (u_5, u_6) , (u_6, u_9) , (u_9, u_{10}) имеется не менее двух перепадов сигналов, поэтому рассмотрим случай $u_1 = u_5$. Тогда получаем $u_1 = u_5 = 1, u_{10} = 0, u_6 \neq u_{11}$.

Если $u_{10} \neq u_{11}$, то снова получается по крайней мере два перепада сигналов среди пар (u_1, u_5) , (u_5, u_6) , (u_6, u_9) , (u_9, u_{10}) , (u_{10}, u_{11}) , в случае же $u_{10} = u_{11}$ получаем $u_{11} = 0, u_6 = 1$ и среди пар соседних координат вектора $(1, 1, 1, 1, 0, 0)$ имеется один перепад сигналов. Это значение соответствует минимальной оценке функции W , поэтому последний план является оптимальным планом приведенной задачи.

Для получения решения исходной задачи дополним полученный вектор

значениями независимых координат так, как это было описано выше:

$$\hat{u}^* = (1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 0)'$$

Значения зависимых координат, по которым достраивался оптимальный план \hat{u}^* , выделены жирным шрифтом.

Из полученного оптимального плана \hat{u}^* сформируем теперь для заданного ЛА ОСП с минимальным числом перепадов сигналов:

$$\bar{u}(0) = (1, 1, 1, 1)', \dots, \bar{u}(1) = (1, 1, 1, 1)', \dots, \bar{u}(2) = (1, 0, 0, 0)'$$

Заметим в заключение, что специфика области применения предложенных методов в большинстве случаев позволяет понизить размерность задачи оптимизации путем сокращения количества ее переменных.

Что касается этапа оценки при решении задачи в используемом методе ветвей и границ, то в общем случае он сводится к решению соответствующих непрерывных задач.

Однако при $p = 2$ их решение значительно упрощается, поскольку коэффициенты в ограничениях этих задач равны 0 или 1. Таким образом, специфика рассматриваемой задачи значительно сокращает трудность ее решения по сравнению со многими другими задачами, решаемыми с помощью метода ветвей и границ.

Вопросы и упражнения

1. Поясните понятие взвешенного алфавита и его содержательный смысл и понятие оптимального эксперимента.
2. Сформулируйте математическую постановку задачи построения синхронизирующей последовательности минимального веса, переводящей ЛА из любого состояния в заданное синхросостояние.
3. Опишите метод сведения задачи поиска обобщенной синхронизирующей последовательности минимального веса для ЛА к задаче целочисленного линейного программирования с линейными ограничениями.

4. Дайте содержательные пояснения понятию перепада сигналов во входной последовательности.
5. Опишите идею метода решения задачи построения обобщенной синхронизирующей последовательности для ЛА с минимальным числом перепадов.

Интервальная арифметика над конечным полем и ее приложения к теории экспериментов с автоматами

Описывается интервальная арифметика над полем $GF(p)$. Вводятся понятия интервала, обобщенного интервала, правильного и неправильного интервалов, бинарные арифметические операции над интервалами и исследуются свойства этих операций. Рассматриваются вопросы решения интервальных уравнений. Вводятся понятия алгебраического, объединенного, допустимого и управляемого множества решений интервальных уравнений.

Об интервальном анализе

В интервальном анализе основным объектом исследования является интервал, представляющий собой замкнутый числовой промежуток.

Так, интервал $[a, b]$ содержит все вещественные числа, заключенные между границами интервала, включая и сами границы a и b . Когда говорят об интервальной неопределенности, подразумевают неполное знание о некоторой величине, т. е. невозможность указать ее точное значение, но возможность обозначить только границы ее изменения. Понятно, что ширина интервала есть мера неопределенности интересующей нас величины. Математическая теория, изучающая задачи с интервальными неопределенностями, получила название интервального анализа.

Ясно, что арифметические операции над величинами, имеющими интервальную неопределенность, будут давать результаты, также содержащие интервальную неопределенность.

Методы интервального анализа [4] [26], развитые к настоящему времени, базируются на использовании арифметических операций с вещественными и комплексными числами. Эти методы находят применение в различных сферах. Так, они оказались полезными при решении многих практических задач, использующих в качестве переменных величины с интервальной неопределенностью.

Заметим, что источниками неопределенностей различных величин могут быть разнообразные факторы. В частности, таким источниками

могут быть ограниченность разрядной сетки компьютера, ошибки различных преобразований, к примеру, преобразований чисел из одной системы счисления в другую, неточности измерений значений величин из-за естественного несовершенства измерительных приборов и т. п.

Необходимо отметить, что интервальный анализ - это не единственная теория, которая оперирует с величинами, содержащими неопределенности. Так, с подобного рода величинами имеют дело теории, использующие их размытые (нечеткие) или вероятностные описания. В связи с этим возникает естественный вопрос: имеет ли интервальный анализ какие-либо достоинства, которые отсутствуют, например, у теории вероятностей или теории нечетких множеств, нечетких графов и т. д.? Надо сказать, что такие достоинства у интервального анализа имеются. В частности, одно из них состоит в более высоком уровне развития математического аппарата для исследований. Так, ни в теории нечетких множеств, ни в теории вероятностей пока не достигнут тот качественный уровень методов решения систем уравнений с неопределенностями, который имеет место для интервальных систем уравнений.

Отметим также еще одно достоинство интервального анализа. Ранее все модели неопределенности, используемые при оценке параметров и идентификации, имели главным образом стохастический или вероятностный характер, базируясь на известных распределениях рассматриваемых величин и т. п. К сожалению, на практике часто бывает совершенно недостаточно информации, чтобы считать неопределенные факторы подчиняющимися какой-либо вероятностной модели, либо эти факторы не удовлетворяют тем или иным условиям, которые на них налагает вероятностная модель неопределенности. К примеру, таковыми могут быть условия независимости величин или специальный вид их распределения. С этой точки зрения интервальные неопределенности являются наименее ограничительными и более адекватны многим практическим задачам.

Укажем еще одну сферу применения интервального анализа - вычисления с приближенными числами, обеспечивающие достаточно точный учет ошибок округлений. Хорошо известно, что все современные компьютеры оперируют с числами, представленными в форме с плавающей запятой. Однако, к сожалению, они не вполне

адекватны не только реальному физическому миру, но и его математическим моделям. Недостатками чисел с плавающей запятой является отсутствие информации о точности тех величин, которые они представляют. Недостатком их является невозможность представить многие используемые вещественные величины в виде чисел с конечной длиной мантиссы. Наконец, еще одним недостатком является отсутствие адекватности между свойствами арифметических операций над числами с плавающей запятой и свойствами идеальных арифметических операций над вещественными числами из-за ошибок округления.

Из сказанного следует, что вычисления с плавающей запятой не позволяют отслеживать вычислительных ошибок и потому не дают возможности произвести анализ точности результатов. Последнее чревато серьезными негативными последствиями при принятии критических решений на основе полученных приближенных результатов.

Обратимся теперь к рассматриваемым нами в курсе лекций линейным автоматам, задаваемым над конечным полем $GF(p)$. С физической точки зрения значения напряжений для входных и выходных сигналов, а также напряжений, которые связаны с представлениями сигналов в элементах памяти электронных устройств, описываемых моделями линейных автоматов, можно измерять в квантованных (калибровочных) единицах. Поскольку по техническим условиям значения напряжений имеют ограничения сверху, предельное значение этого напряжения, выраженное в квантованных единицах, можно принять в качестве характеристики p поля $GF(p)$. Такое постулирование приводит к необходимости оперировать с квантованными значениями напряжений, т. е. к вычислениям в арифметике по модулю p .

Поскольку измерение значений входных и выходных сигналов из-за ограниченной точности измерительных приборов приводит к появлению интервалов, к которым принадлежат эти значения, возникает необходимость разработки фрагмента специфического интервального анализа. Основным объектом такого анализа является интервал с конечным числом элементов, границами которого являются элементы поля $GF(p)$. Заметим, что некоторые результаты такого

анализа не имеют аналогов в интервальном анализе над вещественным полем \mathbb{R} .

Отметим, наконец, что на первый взгляд конечность поля $GF(p)$ должна была бы привести к упрощению соответствующей арифметики, однако на самом деле произошло ее усложнение, что будет видно из результатов, изложенных в следующем разделе этой лекции.

Рассмотренные в предыдущих лекциях синхронизирующие, установочные и диагностические эксперименты с линейными автоматами предполагали возможность получения в процессе их проведения точных значений реакций автоматов. Но, как уже упоминалось ранее, из-за несовершенства измерительных приборов более адекватным является предположение, что реакции автоматов представлены в интервальной форме. Понятно, что методы теории экспериментов, изложенные в предыдущих лекциях, не могут быть непосредственно применены для решения соответствующих задач в интервальной постановке. Однако их решение становится возможным с применением интервальной арифметики над полем $GF(p)$ и специально разработанных для этих целей методов. В некоторых случаях такие методы являются аналогами "точных" методов, разработанных ранее, в иных же случаях "интервальные" методы таких аналогов не имеют. Таким образом, интервальная арифметика над конечным полем, о которой речь пойдет в следующем разделе этой лекции, нашла приложения в теории экспериментов с линейными автоматами в интервальной постановке.

Интервальная арифметика над полем $GF(p)$

Начнем с обозначений и основных определений. Строчные греческие буквы $\alpha, \beta, \gamma, \dots$, а также латинские с чертой снизу или сверху $\underline{a}, \bar{a}, \underline{b}, \bar{b}, \dots$ будут далее обозначать элементы поля $GF(p) = \{0, 1, \dots, p-1\}$, где p - простое число. Элемент этого поля представляет собой класс вычетов. Условимся далее считать, что каждый такой класс заменяется его представителем - наименьшим целым положительным числом, входящим в этот класс. На множестве

классов вычетов определим отношения порядка $(<, \leq, \geq, >)$, которые естественным образом индуцируются соответствующими отношениями порядка на множестве целых чисел.

Подмножество $\subseteq GF(p)$, такое, что $a = [\underline{a}, \bar{a}] = \{\alpha \underline{a} \leq \alpha \leq \bar{a}; \underline{a}, \bar{a} \in GF(p)\}$ будем называть правильным замкнутым интервалом, где \underline{a} и \bar{a} - соответственно его нижняя и верхняя границы.

Запись вида $b = [\underline{b}, \bar{b}]$, где $\underline{b} > \bar{b}$, будем интерпретировать как множество $GF(p) [\bar{b} + 1, \underline{b} - 1]$ и называть это множество неправильным интервалом.

Интервал вида $[\underline{a}, \bar{a}]$, где $\underline{a} = \bar{a}$, будем называть вырожденным и интерпретировать его как элемент поля $GF(p)$.

Множество всех интервалов над $GF(p)$ обозначим через $IGF(p)$, а латинские буквы зарезервируем для обозначения интервалов.

Исходя из интерпретации неправильного интервала $[\underline{b}, \bar{a}]$ как множества "внешних" элементов $GF(p)$ по отношению к правильному интервалу $[\bar{b} + 1, \underline{b} - 1]$, любой неправильный интервал может быть представлен в следующем виде:

$$[\underline{b}, \bar{b}] = [0, \bar{b}] \cup [\underline{b}, p - 1] \quad (19.1)$$

Определение 19.1. Два интервала $a = [\underline{a}, \bar{a}]$ и $b = [\underline{b}, \bar{b}]$ называются равными (и записывается это как $a = b$), если они равны в смысле теории множеств.

Из этого определения следует, что $\underline{a} = \bar{a} = \bar{b} \rightarrow a = b$, но обратная импликация, в отличие от случая вещественных интервалов, места не имеет. Например, при $p = 5$ интервалы $[2, 1]$ и $[3, 2]$ равны как множества, однако их границы не совпадают. Легко видеть, что для

любого $\alpha \in GF(p)$ справедливо равенство $[\alpha, \alpha - 1] = [0, p - 1]$

Очевидно, что отношение равенства в $IGF(p)$ рефлексивно, симметрично и транзитивно.

Введем операции над элементами $IGF(p)$. Далее знаками $+$, $-$, \cdot , $/$ будем обозначать соответствующие арифметические операции над вещественными числами, а знаками $\oplus, \ominus, \odot, \oslash$ операции над элементами поля $GF(p)$.

Определение 19.2. Пусть $\otimes \in \{\oplus, \ominus, \odot, \oslash\}$ - бинарная арифметическая операция над элементами поля $GF(p)$. Если $a, b \in IGF(p)$, то

$$a \otimes b = \{\xi = \alpha \otimes \beta / \alpha \in \beta, \beta \in b\} \quad (19.2)$$

определяет бинарную арифметическую операцию над $IGF(p)$. (В случае деления предполагаем, что $0 \notin b$.)

Очевидно, что операции сложения и умножения элементов $IGF(p)$ коммутативны.

Если интерпретировать поле $GF(p)$ как точки числовой оси, то каждый правильный интервал $[\underline{a}, \bar{a}]$ - это множество точек, расположенных вплотную одна за другой между \underline{a} и \bar{a} .

Отметим, что результатом операции над интервалом может оказаться множество точек, не являющихся одним интервалом, а представляющее собой объединение нескольких интервалов, разбросанных по числовой оси. Например, для $p = 7$, $[1, 2] * [2, 3] = [2, 4][6, 6]$.

Введем еще одно понятие. Подмножество $A \subseteq GF(p)$, такое, что

$$a = \bigcup_{i \in I} a_i \quad (19.3)$$

где $a_i \in IGF(p)$, I - конечное множество индексов и для $I \neq j, a_i \cup a_j = \emptyset$, назовем обобщенным интервалом поля $GF(p)$.

Обобщенный интервал будем обозначать строчной буквой. Понятно, что интервал из $IGF(p)$ есть частный случай обобщенного интервала. За множеством всех обобщенных интервалов сохраним обозначение $IGF(p)$, поскольку из контекста каждый раз ясно, какой именно интервал имеется в виду в конкретном случае.

Введем бинарные операции над обобщенными интервалами.

Определение 19.3. Пусть $*$ $\in \{+, -, \cdot, \cdot\}$ - бинарная операция. Если

$$A = \bigcup_{i \in I} a_i, B = \bigcup_{j \in J} b_j,$$

где a_i, b_j - обычные интервалы поля $GF(p)$, то

$$A * B = (\bigcup_{i \in I} a_i) * (\bigcup_{j \in J} b_j) = \bigcup_{i \in I, j \in J} a_i * b_j$$

Заметим, что множество $IGF(p)$ обобщенных интервалов является замкнутым относительно введенных бинарных операций.

Как следует из формулы (19.4), любая бинарная операция над обобщенными интервалами сводится к соответствующей операции над обычными интервалами поля $GF(p)$.

Введем унарную операцию над обычным интервалом $x = [\underline{x}, \bar{x}] \in IGF(p)$:

$$-x = [-\underline{x}, -\bar{x}]$$

где " ξ " - это элемент поля $GF(p)$, обратный к элементу ξ по сложению, т. е. такой, что $\xi \oplus (-\xi) = 0$. Аналогичную операцию для обобщенного интервала $X = \bigcup_{i \in I} x_i$ определим так:

$$-X = \bigcup_{i \in I} (-x_i)$$

С использованием этой операции бинарная операция вычитания выражается через операцию сложения:

$$A - B = A + (-B) \quad (19.5)$$

Например, при $p = 5, a = [0, 1], b = [4, 1]$, тогда $[0, 1] - [4, 1] = [0, 1] + (-[4, 1]) = [0, 1] + [4, 1] = [4, 2]$.

Введем операцию умножения элемента $\alpha \in GF(p)$ на обобщенный интервал X :

$$\alpha \cdot X = Y_{\xi \in X} [\alpha \odot \xi, \alpha \odot \xi] \quad (19.6)$$

Используя эту операцию, бинарную операцию умножения обобщенных интервалов можно записать в виде

$$AB = Y_{\alpha \in A} \alpha B \quad (19.7)$$

Например, если $p = 5, a = [2, 4], b = [1, 2]$ то имеем

$$a \cdot b = 2[1, 2] \cup 3[1, 2] \cup 4[1, 2] = [2, 2] \cup [4, 4] \cup [3, 3] \cup [1, 1] \cup [3, 4] = [1, 4]$$

И наконец, операцию деления обобщенных интервалов можно записать в виде

$$A/B = A(Y_{\beta \in B} 1/\beta) \quad (19.8)$$

где $1/\beta$ - это элемент $GF(p)$, обратный элементу β по умножению, т. е. такой, что $\beta \cdot (1/\beta) = 1$.

Например, если $p = 5, a = [1, 3], b = [3, 4]$, то

$$[1, 3]/[3, 4] = [1, 3](1/3 \cup 1/4) = [1, 3] * (2 \cup 4) = 2 * [1, 3] \cup 4 * [1, 3] = [1, 2] \cup [4, 4] \cup [2, 4] = [1, 4]$$

Остановимся на свойствах арифметических операций в $IGF(p)$.

Напомним, что в классической интервальной арифметике [2] функция

$$\zeta = f(\xi, \eta)$$

где $f(\xi, \eta) = \xi * \eta, * \in \{+, -, \cdot, /\}, \xi \in x, \eta \in y, x, y \in I(R)$ (при делении предполагается, что $0 \notin y$) - непрерывные функции на компактном множестве, и потому $f(\xi, \eta)$ принимает как наименьшее, так и наибольшее значения. Таким образом, $x * y$ есть также замкнутый вещественный интервал, который полностью заполнен вещественными числами, являющимися результатами упомянутых бинарных операций над соответствующими числами. При этом границы результата выражаются через операции с границами операндов. Некоторая, но не полная, аналогия справедлива и в интервальной арифметике над полем $GF(p)$. Исходя из указанной аналогии, найдем формулы, позволяющие выражать результаты бинарных операций над обычными интервалами через их границы.

Введем понятие ширины обобщенного интервала $X = Y_{i \in I} x_i$ и обычного интервала $x_i = [\underline{x}_i, \bar{x}_i]$, обозначаемых далее как $w(X)$ и $w(x_i)$ соответственно:

$$w(X) = \sum_{i \in I} w(x_i),$$

$$w(x_i) = \begin{cases} \bar{x}_i - \underline{x}_i + 1, & \text{если } x_i \text{ правильный интервал,} \\ \bar{x}_i - x_i + p + 1, & \text{если } x_i \text{ неправильный интервал} \end{cases}$$

Понятно, что $w(x_i)$ и $w(X)$ - это число элементов в соответствующих интервалах. Очевидно также, что $w(x_i) = \bar{x}_i \ominus \underline{x}_i + 1$ в любом случае.

Пусть $a = [\underline{a}, \bar{a}], b = [\underline{b}, \bar{b}]$ - обычные интервалы поля $GF(p)$, тогда для сложения интервалов справедлива следующая формула

$$a+b = \{matrix[\underline{a} \oplus \underline{b}, \bar{a} \oplus \bar{b}], \quad \text{если } w(a) + w(b) \leq p, \quad \text{в противном случае}$$

Доказательство этой формулы приведено в [31].

Аналогично, для операции вычитания обычных интервалов справедлива следующая формула

$$a - b = \begin{cases} [\underline{a} \ominus \bar{b}, \bar{a} \ominus \underline{b}], & \text{если } w(a) + w(b) \leq p, \\ [0, p - 1] & \text{в противном случае} \end{cases}$$

Произведение и частное интервалов представить через операции с границами сомножителей в общем случае не удастся. Ситуация по сравнению с классической интервальной арифметикой осложняется в связи с тем, что, к примеру, для произведения интервалов не выполняется даже включение: $ab \subseteq [\underline{a} \odot \underline{b}, \bar{a} \odot \bar{b}]$.

Следующее утверждение, доказательство которого приведено в [31], касается свойств введенных интервальных операций и самих интервалов.

Теорема 19.1. Пусть $A, B, C \in IGF(p)$, тогда:

1. $A + B = B + A, A * B = B * A$ (коммутативность);
2. $(A + B) + C = A + (B + A), (AB)C = A(BC)$ (ассоциативность);
3. $(\forall A \in IGF(p)) A = A + [0, 0] = [0, 0] + A, A * [1, 1] = [1, 1] * A$, то есть $[0, 0]$ и $[1, 1]$ являются единственными нейтральными элементами для сложения и умножения соответственно;
4. $IGF(p)$ не имеет делителей нуля;
5. Произвольный невырожденный интервал из $IGF(p)$ не имеет обратного ни по сложению, ни по умножению, и $0 \in A - A, 1 \in A$;
6. $A(B + C) \in AB + AC$ (субдистрибутивность);
7. $\lambda(A + B) = \lambda A + \lambda B$, где $\lambda \in GF(p)$ (дистрибутивность умножения на число);
8. $\forall \alpha \in GF(p) [\alpha, \alpha - 1] = [0, p - 1]$;
9. $w(A * B) \leq w(A) * w(B)$, где $*$ $\in \{\oplus, \ominus, \odot, \oslash\}$;
10. $w(a + b) = w(a - b) = w(a) + w(b) - 1$ если a, b -

обычные интервалы и $w(a) + w(b) < p$;

$$11. w(\lambda A) = w(A), \lambda \in GF(p) \text{ и } \lambda \neq 0.$$

Заметим, что $IGF(p)$ с введенными операциями сложения и умножения на число не является линейным пространством, и даже квазилинейным, так как не выполняется аксиома линейности

$$(\lambda + \mu)A = \lambda A + \mu A,$$

где

$$\lambda, \mu \in GF(p), A \in IGF(p)$$

Как и в классической интервальной арифметике, в интервальной арифметике над конечным полем $GF(p)$ возникают задачи поиска решений интервальных уравнений и систем. Кратко остановимся на этом.

Рассмотрим уравнение относительно X

$$f(A, X) = B \quad (19.9)$$

зависящее от некоторых параметров $(A_1, \dots, A_s)^T$, где $A_i, B \in IGF(p)$, $i = \overline{1, s}$, $f(A, X)$ - рациональное выражение, которое состоит из интервалов A, X соединенных знаками арифметических операций.

Определение 19.4. Назовем алгебраическим решением уравнения (19.9) такой обобщенный интервал $X \in IGF(p)$, что при его подстановке в (19.9) получается точное равенство.

Пусть, например, $p = 5$ и рассматривается уравнение $[1, 2] + X = [4, 2]$. Можно убедиться, что интервалы $X_1 = [0, 0] \cup [3, 3]$ и $X_2 = [3, 0] ([3, 0] = [0, 0] \cup [3, 3] \cup [4, 4])$, подставленные в заданное уравнение, дают точное равенство, т. е. оба

этих интервала являются алгебраическими решениями. Из этого примера следует, что алгебраическое решение уравнения $A + X = B$ в общем случае не единственно.

Следующие определения множеств решений были введены в [51] для множеств решений интервальных систем линейных алгебраических уравнений, здесь эти множества решений определяются для уравнений (19.9).

Определение 19.5. Назовем объединенным множеством решений уравнения (19.9) следующее множество элементов:

$$X_{\exists\exists} = \{\xi \in GF(p) \mid (\exists\alpha \in A)(\exists\beta \in B)f(\alpha, \xi) = \beta\}$$

Так, для уравнения $[1, 2] + X = [4, 2]$ объединенное множество решений есть $X_{\exists\exists} = [0, 4]$.

Определение 19.6. Назовем допустимым множеством решений уравнения (5.9) следующее множество:

$$X_{\forall\exists} = \{\xi \in GF(p) \mid (\forall\alpha \in A)(\exists\beta \in B)f(\alpha, \xi) = \beta\}$$

Для рассмотренного выше примера $X_{\forall\exists} = [3, 0]$, т. е. для этого уравнения имеет место включение $X_{\forall\exists} \subset X_{\exists\exists}$. Легко показать, что такое включение справедливо и в общем случае. Кроме того, для этого же примера $X_{\forall\exists} = X_2$, т. е. допустимое множество решений совпадает с одним из алгебраических решений.

Определение 19.7. Назовем управляемым множеством решений уравнения (19.9) следующее множество:

$$X_{\forall\forall} = \{\xi \in GF(p) \mid (\forall\beta \in B)(\exists\alpha \in A)f(\alpha, \xi) = \beta\}$$

Так, для $p = 5$ уравнение $[1, 3] + X = [0, 1]$ управляемое множество решений есть $X_{\forall\forall} = [3, 4]$.

Теорема 19.2. Линейное уравнение $a + X = b$, где a, b - обычные

интервалы над полем $GF(p)$, имеет алгебраическое решение X в виде обобщенного интервала тогда и только тогда, когда $w(a) \leq w(b)$.

Доказательство.

Необходимость. Докажем ее от противного. Пусть $w(a) > w(b)$. Покажем, что в этом случае рассматриваемое уравнение решений не имеет. Пусть $a = \{\alpha_1, \dots, \alpha_\mu\}$. Выберем произвольный элемент $\xi \in GF(p)$ и построим множество $a_\xi = \{\alpha_1 \oplus \xi, \dots, \alpha_\mu \oplus \xi\}$. Очевидно, что в a_ξ все элементы попарно различны т. е. $w(a_\xi) = \mu$. Элемент $\xi \notin X$ поскольку среди элементов A_ξ существует по крайней мере один такой элемент $a_i \oplus \xi \notin b$ ибо $w(b) < \mu$. Из произвольности элемента ξ следует, что в X нет ни одного элемента $GF(p)$, т. е. алгебраическое решение пусто.

Достаточность. Пусть $w(a) \leq w(b)$ тогда алгебраическое решение есть $X = [\underline{b} \ominus \underline{a}, \bar{b} \ominus \bar{a}]$, $a + X = [a, a] + [\underline{b} \ominus \underline{a}, \bar{b} \ominus \bar{a}] = \underline{a} \oplus \underline{b} \ominus \underline{a}, \bar{a} \oplus \bar{b} \ominus \bar{a} = [\underline{b}, \bar{b}] = b$. При этом по свойству II в теореме 19.1 $w(b) = w(a + X) = w(a) + w(X) - 1$. Отсюда следует, что $w(b) - w(a) = w(X) - 1$ Тогда, учитывая, что $w(a) \leq w(b)$, получим $w(X) \geq 1$, т. е. множество X не пусто.

Заметим, что теорема 19.2 для случая, когда в уравнении $A + X = B$ вместо обычных интервалов использованы обобщенные интервалы, неверна. В самом деле, пусть $p = 7$, $A + [1, 1] \cup [3, 3] \cup [5, 5]$, $B + [6, 2]$. Перебором можно проверить, что уравнение в этом случае алгебраического решения не имеет.

Для существования алгебраического решения уравнения $AX = B$ условие $w(A) \leq w(B)$ является необходимым, но не является достаточным даже для случая, когда A и B - обычные интервалы. Например, при $p = 7$ уравнение $[2, 4] * X = [2, 5]$ алгебраического

решения не имеет, хотя $w(A) = 3 < w(B) = 4$. В то же время это уравнение имеет объединенное множество решений $X_{\exists\exists} = [1, 6]$ и допустимое множество решений $X_{\forall\exists} = [1, 1] \cup [6, 6]$.

Для уравнения

$$A + X = B$$

введенные множества решений определяются следующим образом:

$$X_{\exists\exists} = Y_{\alpha \in A} Y_{\beta \in B} (\beta \ominus \alpha) = B_A \quad (19.10)$$

$$X_{\forall\exists} = I_{\alpha \in A} Y_{\beta \in B} (\beta \ominus \alpha) \quad (19.11)$$

$$X_{\forall\forall} = I_{\beta \in B} Y_{\alpha \in A} (\beta \ominus \alpha) \quad (19.12)$$

Для уравнения

$$AX = B$$

где

$$0 \notin A$$

множество решений определяются аналогично:

$$X_{\exists\exists} = Y_{\alpha \in A} Y_{\beta \in B} (\beta \oslash \alpha) = B/A \quad (19.13)$$

$$X_{\forall\exists} = I_{\alpha \in A} Y_{\beta \in B} (\beta \oslash \alpha) \quad (19.14)$$

$$X_{\forall\forall} = I_{\beta \in B} Y_{\alpha \in A} (\beta \oslash \alpha) \quad (19.15)$$

Равенства (19.10)-(19.12), (19.13)-(19.15) следуют из определения множеств $X_{\exists\exists}$, $X_{\forall\exists}$, $X_{\forall\forall}$ и определения теоретико-множественных операций объединения и пересечения.

Рассмотрим пример. Пусть $p = 7$ и уравнение имеет вид

$$[2, 4] * X = [2, 5]$$

Для наглядности построим таблицу деления $\beta \oslash \alpha$, где $\beta \in [2, 5]$, $\alpha \in [2, 4]$.

$\alpha \setminus \beta$	2	3	4	5
2	1	5	2	6
3	3	1	6	4
4	4	6	1	3

Объединение всех элементов строк (столбцов) таблицы составляет объединенное множество решений $X_{\exists\exists} = [1, 6]$. Пересечение элементов строк таблицы составляет допустимое множество решений $X_{\forall\exists} = [1, 1][6, 6]$. Пересечение элементов столбцов таблицы пусто, поэтому $X_{\exists\forall} = \emptyset$ и управляемых решений уравнение не имеет.

Теорема 19.3. Все множество решений уравнений $A + X = B$ и $AX = B$ и алгебраическое решение, если оно существует, включаются в объединенное множество решений, т. е. в $B - A$ для уравнения $A + X = B$ и в B для уравнения $AX = B$

Доказательство этой теоремы следует из равенств (19.10)-(19.12), (19.13)-(19.15) для множеств решений. Для алгебраического решения очевидно, что если оно существует, то все его элементы являются допустимыми решениями и, следовательно, принадлежат также и объединенному множеству решений.

Вопросы и упражнения

1. Дайте определение равенства двух интервалов.
2. Приведите определения различных типов интервалов: обычных, обобщенных, правильных, неправильных, вырожденных.
3. Дайте определение бинарных операций над различными типами интервалов.
4. Перечислите свойства интервалов и операций над ними.
5. Приведите определения различных типов решений интервального уравнения $f(A, X) = B$: алгебраического,

объединенного, допустимого, управляемого.

6. Сформулируйте критерий существования алгебраического решения линейного интервального уравнения $a + X = b$.

Диагностическая задача в интервальной постановке

Исследуется диагностическая задача для линейных автоматов, когда наблюдаемая на выходах реакция автоматов представлена в виде интервалов в поле $GF(p)$. Предложены методы решения диагностической задачи с использованием модифицированной конструкции дерева преемников и путем сведения ее к решению систем линейных алгебраических уравнений с интервальной правой частью. Описан генетический алгоритм для решения упомянутой системы, значительно сокращающий время поиска решений.

Постановка задачи и алгебраический метод решения

В классической диагностической задаче для линейных автоматов предполагается, что каждая выходная реакция в момент времени t - это вектор $\bar{y}(t) = [y_1(t), \dots, y_m(t)]'$, координаты $y_i(t)$ которого представляют собой точные значения. Поскольку практически значения координат наблюдаемой реакции автомата получаются в результате измерений, они неизбежно содержат погрешности, размеры которых зависят от точности измерительных приборов. Поэтому более реальной по сравнению с классической задачей является диагностическая задача, в которой реакция автомата представлена в виде интервалов в поле $GF(p)$:

$$\bar{y}(t) = [[y_1(t), \bar{y}_1(t)], \dots, [y_m(t), \bar{y}_m(t)]]' \quad (20.1)$$

Решением интервальной диагностической задачи, заключающейся в определении начального состояния автомата по наблюдаемой его реакции на подаваемую диагностическую последовательность, будем считать такое состояние автомата из множества допустимых начальных состояний, если, стартуя из него, автомат в каждый момент времени t порождает выходной вектор, координаты которого принадлежат интервальному вектору (20.1). Заметим, что, как и в случае классической задачи, мы будем считать интервальную диагностическую задачу разрешимой, если искомое начальное состояние определяется однозначно.

Понятно, что чем больше ширина интервалов в выходных векторах, тем сложнее найти начальное состояние автомата. Если же интервалы в выходных векторах полностью покрывают поле $GF(p)$, то нахождение начальных состояний становится невозможным. Такая ситуация вполне может иметь место при малых значениях p , поэтому далее предполагается, что $p \geq 5$.

Для решения диагностической задачи необходимо иметь ДП, общий метод построения которой в классическом случае описан в [18]. Остановимся кратко на методе построения ДП, ориентированном на решение интервальной диагностической задачи. Построение ДП в этом случае может быть осуществлено с использованием конструкции классического диагностического дерева [18] с внесенными в нее некоторыми незначительными изменениями. Эти изменения касаются способа формирования преемника A -группы.

В классическом случае каждое состояние автомата любого сигма-множества A -группы имеет единственного преемника в одном и только одном сигма-множестве преемника упомянутой A -группы. Отсюда следует, что в классическом случае каждая A -группа, связанная с любой ветвью диагностического дерева, содержит в своих сигма-множествах одно и то же суммарное число состояний, равное мощности множества допустимых начальных состояний автомата.

Построение преемника сигма-множества A -группы по некоторому входному сигналу в интервальном диагностическом дереве будем осуществлять следующим образом. Сначала вычисляем все состояния-преемники этого сигма-множества по упомянутому входному сигналу и соответствующие "интервальные" реакции автомата. Каждая "интервальная" реакция ЛА заменяется множеством точных реакций, которое представляет собой все возможные комбинации из точных значений, принадлежащих соответствующему интервалу. Таким образом, каждому состоянию-преемнику будет соответствовать некоторое конечное множество точных реакций. Состояния-преемники будут помещаться в одно и то же сигма-множество, если в соответствующих им множествах точных реакций есть совпадающие между собой. Это правило формирования сигма-множеств A -группы следующего уровня может привести к тому, что одно сигма-множество

предшествующего уровня в следующем уровне дерева попадает в несколько сигма-множеств. Таким образом, в двух новых сигма-множествах может находиться одно и то же состояние-преемник. Отсюда вытекает, что общее число состояний в преемнике A -группы в интервальном диагностическом дереве может оказаться больше мощности множества допустимых начальных состояний автомата.

Что касается правил, по которым некоторая ветвь интервального диагностического дерева становится оконечной, то они остаются теми же, что и в классическом диагностическом дереве. Проиллюстрируем построение интервального диагностического дерева на примере.

Рассмотрим ЛА над полем $GF(7)$, заданный с помощью следующих характеристических матриц:

$$A = \begin{bmatrix} 2 & 1 & 4 \\ 5 & 2 & 6 \\ 3 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix}, C = \begin{bmatrix} 5 & 0 & 3 \\ 1 & 4 & 6 \end{bmatrix}, D = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

Пусть множество допустимых начальных состояний этого ЛА содержит два следующих состояния: $[456]'$ $[105]'$. Классическое диагностическое дерево для рассматриваемого автомата представлено на рис.20.1. Над каждой A -группой на этом рисунке в прямоугольнике помещена информация о состояниях-преемниках (\bar{s}) и выходных реакциях (\bar{y}). Ниже помещены сигма-множества (в фигурных скобках), составляющие A -группу. Из этого рисунка следует, что рассматриваемый автомат имеет 7 диагностических последовательностей длины 1: $u = 0, u = 1, \dots, u = 6$.

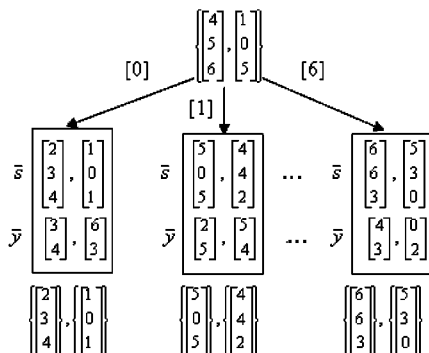


Рис. 20.1.

Теперь построим фрагменты интервального диагностического дерева (одну его ветвь для входного сигнала $u = 0$) в предположении, что при измерении реакции ЛА по каждой координате выходного вектора может произойти ее искажение на одну единицу как в меньшую, так и в большую сторону. Этот фрагмент представлен на рис.20.2.

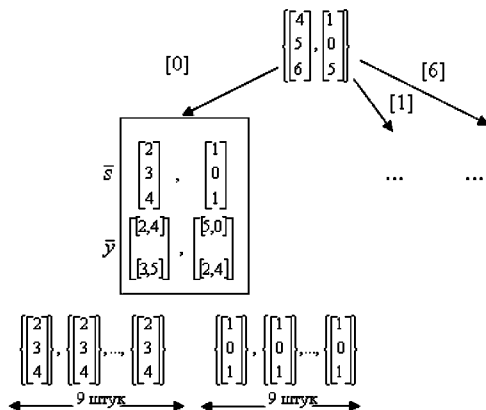


Рис. 20.2.

При переходах ЛА из состояний $[456]'$, $[105]'$ в состояния $[234]'$, $[101]'$ соответственно вычисляемые точные реакции есть $[34]'$, $[63]'$. При наличии оговоренных выше погрешностей измерения эти точные реакции превращаются в интервальные реакции $[[2, 4]'$, $[3, 5]']$, $[[5, 0]$, $[2, 4]']$. Каждую из этих реакций заменяем соответственно множеством всех возможных комбинаций величин из приведенных интервалов:

$$\left\{ \begin{array}{l} \left[\begin{array}{c} 2 \\ 3 \\ 5 \end{array} \right], \left[\begin{array}{c} 2 \\ 4 \\ 5 \end{array} \right], \left[\begin{array}{c} 2 \\ 5 \\ 6 \end{array} \right], \left[\begin{array}{c} 3 \\ 3 \\ 6 \end{array} \right], \left[\begin{array}{c} 3 \\ 4 \\ 6 \end{array} \right], \left[\begin{array}{c} 3 \\ 5 \\ 6 \end{array} \right], \left[\begin{array}{c} 4 \\ 3 \\ 2 \end{array} \right], \left[\begin{array}{c} 4 \\ 4 \\ 3 \end{array} \right], \left[\begin{array}{c} 4 \\ 5 \\ 4 \end{array} \right] \\ \left[\begin{array}{c} 5 \\ 5 \\ 2 \end{array} \right], \left[\begin{array}{c} 5 \\ 6 \\ 3 \end{array} \right], \left[\begin{array}{c} 5 \\ 6 \\ 4 \end{array} \right], \left[\begin{array}{c} 6 \\ 2 \\ 2 \end{array} \right], \left[\begin{array}{c} 6 \\ 3 \\ 3 \end{array} \right], \left[\begin{array}{c} 6 \\ 4 \\ 2 \end{array} \right], \left[\begin{array}{c} 0 \\ 0 \\ 2 \end{array} \right], \left[\begin{array}{c} 0 \\ 0 \\ 3 \end{array} \right], \left[\begin{array}{c} 0 \\ 0 \\ 4 \end{array} \right] \end{array} \right\},$$

С использованием этих множеств построим теперь сигма-множества A -группы. Так, поскольку реакция $[23]'$ из первого множества не совпадает ни с одной реакцией из второго, то соответствующее ей

состояние-преемник $[234]'$ образует простое, т. е. одноэлементное, сигма-множество. Аналогичная ситуация имеет место для всех остальных элементов обоих множеств. Следовательно, преемником множества допустимых начальных состояний является A -группа, содержащая 18 одноэлементных сигма-множеств, среди которых по 9 штук содержит одно и то же состояние.

Вернемся теперь к интервальной диагностической задаче. Для нахождения неизвестного начального состояния $\bar{s}(0) = [s_1(0), \dots, s_n(0)]'$ по известной ДП $\bar{u}(0), \bar{u}(1), \dots, \bar{u}(k)$ и наблюдаемой в процессе проведения эксперимента с ЛА выходной последовательности $\bar{y}(0), \bar{y}(1), \dots, \bar{y}(k)$, сформируем следующую СЛАУ:

$$\begin{cases} CA(0) = \bar{y}(0) - D\bar{u}(0), \\ CA\bar{s}(0) = \bar{y}(1) - CB\bar{u}(0) - D\bar{u}(1), \\ \dots \\ CA^k\bar{s}(0) = \bar{y}(k) - CA^{k-1}B\bar{u}(0) - \dots - CB\bar{u}(k-1) - D\bar{u}(k) \end{cases}$$

Эта система получена на основе формулы полной реакции ЛА для $t = \overline{0, k}$.

Поскольку выходные реакции $\bar{y}(0), \bar{y}(1), \dots, \bar{y}(k)$ представляют собой интервальные вектора, то правые части системы (20.2) также являются интервальными векторами, тогда как матрица системы (20.2) является точной.

В общем случае число уравнений в системе (20.2), которое обозначим через ν , может не совпадать с числом неизвестных, равным размерности n ЛА. Если $\nu < n$, то, как известно из алгебры, общее решение такой системы представляется с использованием свободных переменных и матрица системы приводится к квадратной. Если $\nu > n$, то из факта существования у ЛА диагностической последовательности вытекает существование решения системы (20.2). Последнее означает, что ее ранг равен n и поэтому матрица системы и в этом случае может быть приведена к квадратной.

Исходя из сказанного, решение рассматриваемой интервальной диагностической задачи в математическом плане сводится к решению СЛАУ с точной квадратной матрицей и интервальной правой частью.

Представим такую систему в общем виде:

$$Ax = B \quad (20.3)$$

где $A = |a_{ij}|_{n \times n}$ - матрица, элементами которой являются элементы поля $GF(p)$, $B = [B_1, \dots, B_n]'$, а каждое B_i является интервалом или мультиинтервалом (объединением нескольких интервалов).

Тривиальный способ нахождения всех решений системы (20.3) состоит в замене ее интервальной правой части всевозможными конкретными значениями из соответствующих интервалов и поиска решений получающихся при этом обычных систем линейных уравнений одним из известных методов. Однако такой способ является трудоемким, поскольку ведет к необходимости решения z шпук систем, где

$$z = w(B_1) \dots w(B_n) = \prod_{i=1}^n w(B_i)$$

Доказываемая ниже теорема позволяет значительно уменьшить трудоемкость нахождения всех решений системы (20.3).

Предварительно введем следующие обозначения:

$$\{e_k\} = \{0, \text{если } i \neq k, 1, \text{если } i = k \quad , k = \overline{1, 0}; i = \overline{1, n}; \\ \underline{B} = [\underline{B}_1, \underline{B}_2, \dots, \underline{B}_n]'$$

где

$$\underline{B}_i = \alpha_i, \alpha_i \in B_i, \forall \beta_i \in B_i (\alpha \leq \beta), i = \overline{1, n}$$

Теорема 20.1. Пусть $(\xi_1, \dots, \xi_n)'$ есть решение системы

$$A\xi = \underline{B} \quad (20.4)$$

а $(\delta_1^{(k)}, \dots, \delta_n^{(k)})'$ есть решения n штук систем

$$Ax = e_k, k = \overline{1, n} \quad (20.5)$$

Тогда решение $x = (x_1, \dots, x'_n)$ системы (20.3) имеет вид

$$x_j = \xi_j + \sum_{k=1}^n m_k \delta_j^{(k)}, j = \overline{1, n} \quad (20.6)$$

где $\underline{B}_k + m_k \in B_k$ т. е. m_k пробегает независимо все значения от 0 до $p - 1$, при этом из них точно $w(B_k)$

Доказательство. Рассмотрим i -ю строку системы (20.3):

$$\begin{aligned} (Ax)_i &= \sum_{j=1}^n a_{ij} x_j = \sum_{j=1}^n a_{ij} \left(\xi_j + \sum_{k=1}^n m_k \delta_j^{(k)} \right) = \sum_{j=1}^n \left(a_{ij} \xi_j + \sum_{k=1}^n m_k \delta_j^{(k)} \right) \\ &= \sum_{j=1}^n a_{ij} \xi_j + \sum_{j=1}^n \sum_{k=1}^n m_k \delta_j^{(k)} = \underline{B}_i + \sum_{k=1}^n m_k \sum_{j=1}^n a_{ij} \delta_j^{(k)} = \underline{B}_i + m_i, \end{aligned}$$

поскольку

$$\sum_{j=1}^n a_{ij} \delta_j^{(k)} = \begin{cases} 0, & \text{если } i \neq k, \\ 1, & \text{если } i = k \end{cases}$$

Таким образом, $(Ax)_i = \underline{B}_i + m_i$, а $\underline{B}_i + m_i \in B_i$, следовательно, $(Ax)_i = B_i$, откуда и вытекает справедливость теоремы.

Из этой теоремы следует, что в действительности для построения всех решений системы (20.3) необходимо решить только $(n + 1)$ систему линейных уравнений, а не $w(B_1) \dots w(B_n)$ систем, как это было в тривиальном способе.

В качестве примера рассмотрим применение этого метода для решения СЛАУ с точной матрицей и интервальной правой частью.

Пусть $p = 37$ и необходимо найти все решения системы

$$\begin{cases} 3x_1 + 5x_2 = [30, 36], \\ 5x_1 + 2x_2 = [20, 30] \end{cases}$$

Вначале рассмотрим систему вида (20.4):

$$\begin{cases} 3x_1 + 5x_2 = 30, \\ 5x_1 + 2x_2 = 20 \end{cases}$$

Решение этой системы есть пара $(\xi_1, \xi_2) = (6, 32)$.

Теперь рассмотрим две системы вида (20.5):

$$\begin{cases} 3\delta_1^{(1)} + 5\delta_1^{(2)} = 1, & \begin{cases} 3\delta_1^{(2)} + 5\delta_2^{(2)} = 0, \\ 5\delta_1^{(2)} + 2\delta_2^{(2)} = 1 \end{cases} \end{cases}$$

Их решениями являются

$$\delta^{(1)} = (\delta_1^{(1)}, \delta_2^{(1)}) = (33, 10), \delta^{(2)} = (\delta_1^{(2)}, \delta_2^{(2)}) = (10, 31).$$

Тогда все множество решений системы таково:

$$x_1 = \xi_1 + i\delta_1^{(1)} + j\delta_1^{(2)}, x_2 = \xi_2 + i\delta_2^{(1)} + j\delta_2^{(2)},$$

где

Таким образом, получаем:

$$x_1 = 6 + 33i + 10j, \text{ где } 0 \leq i \leq 6, 0 \leq j \leq 10$$

Все эти решения представлены в следующей таблице:

b_1	2	20	21	22	23	24	25	26	27	28	29	30
30	(6, 32)	(16, 26)	(26, 20)	(36, 14)	(9, 8)	(19, 2)	(29, 33)	(2, 27)	(12, 21)	(22, 15)	(32, 9)	
31	(2, 5)	(12, 36)	(22, 30)	(32, 24)	(5, 18)	(15, 12)	(25, 6)	(35, 0)	(8, 31)	(18, 25)	(28, 19)	

32	(35, 15)	(8, 9)	(18, 3)	(28, 34)	(1, 28)	(11, 22)	(21, 16)	(31, 10)	(4, 4)	(14, 35)	(24, 29)
33	(31, 25)	(4, 19)	(14, 13)	(24, 7)	(34, 1)	(7, 32)	(17, 26)	(27, 20)	(0, 14)	(10, 8)	(20, 2)
34	(27, 35)	(0, 29)	(10, 23)	(20, 17)	(30, 11)	(3, 5)	(13, 36)	(23, 30)	(33, 24)	(6, 18)	(16, 12)
35	(23, 8)	(33, 2)	(6, 33)	(16, 27)	(26, 21)	(36, 15)	(9, 9)	(19, 3)	(29, 34)	(2, 28)	(12, 22)
36	(19, 18)	(29, 12)	(2, 6)	(12, 0)	(22, 31)	(32, 25)	(5, 19)	(15, 13)	(25, 7)	(35, 1)	(8, 32)

Таким образом, в рассмотренном примере предлагаемый метод отыскания всех решений СЛАУ потребовал решения 3 систем линейных уравнений вместо 77 при использовании тривиального метода.

Теперь проиллюстрируем применение метода решения СЛАУ на конкретной диагностической задаче.

Рассмотрим ЛА над полем $GF(7)$, заданный с помощью характеристических матриц, которые приведены выше в этом разделе. Пусть множество допустимых начальных состояний содержит 4 состояния:

$$\{[203]', [001]', [456]', [105]'\}$$

Диагностическая последовательность для заданного множества допустимых начальных состояний, построенная с использованием описанного выше диагностического дерева, такова:

$u(0) = [0]$, $u(1) = [0]$, $u(2) = [0]$. Предположим, что измерения выходной реакции ЛА на эту диагностическую последовательность дали следующие результаты:

$$y(0) = \begin{bmatrix} [5, 6] \\ [2, 4] \end{bmatrix}, y(1) = \begin{bmatrix} [0, 2] \\ [0, 1] \end{bmatrix}, y(2) = \begin{bmatrix} [0, 1] \\ [3, 5] \end{bmatrix}$$

Заметим, что эти результаты получены в несколько иных предположениях об искажениях при измерении, чем были сделаны при

построении фрагмента интервального диагностического дерева на рис.20.2.

Для решения интервальной диагностической задачи необходимо найти все решения системы уравнений вида (20.2).

$$\begin{bmatrix} 5 & 0 & 3 \\ 1 & 4 & 6 \\ 5 & 5 & 2 \\ 5 & 2 & 6 \\ 6 & 1 & 3 \\ 3 & 2 & 3 \end{bmatrix} \begin{bmatrix} s_1(0) \\ s_2(0) \\ s_3(0) \end{bmatrix} = \begin{bmatrix} [5, 6] \\ [2, 4] \\ [0, 2] \\ [0, 1] \\ [0, 1] \\ [3, 5] \end{bmatrix} \quad (20.7)$$

Заметим, что для нахождения решения этой интервальной системы уравнений тривиальным способом необходимо решить 216 систем.

Для приведенной системы количество неизвестных равно 3, а число уравнений равно 6. Приводя матрицу к квадратной, получаем следующую систему:

$$\begin{bmatrix} 5 & 0 & 3 \\ 1 & 4 & 6 \\ 5 & 5 & 2 \end{bmatrix} \begin{bmatrix} s_1(0) \\ s_2(0) \\ s_3(0) \end{bmatrix} = \begin{bmatrix} [5, 6] \\ [2, 4] \\ [0, 2] \end{bmatrix} \quad (20.8)$$

Обозначим матрицу полученной системы (20.8) через F' . В соответствии с описанным выше методом необходимо решить следующие 4 системы линейных уравнений:

$$\begin{aligned} F'[\xi_1 \xi_2 \xi_3]' &= [000]', \\ f[\delta_1^{(1)} \delta_2^{(1)} \delta_3^{(1)}]' &= [100]', \\ F'[\delta_1^{(2)} \delta_2^{(2)} \delta_3^{(2)}]' &= [010]', \\ F'[\delta_1^{(3)} \delta_2^{(3)} \delta_3^{(3)}]' &= [001]' \end{aligned}$$

Решая их, получаем следующие результаты:

$$(\xi_1, \xi_2, \xi_3) = (0, 0, 0)$$

$$(\delta_1^{(1)}, \delta_2^{(1)}, \delta_3^{(1)}) = (1, 0, 1)$$

$$(\delta_1^{(2)}, \delta_2^{(2)}, \delta_3^{(2)}) = (6, 5, 4),$$

$$(\delta_1^{(3)}, \delta_2^{(3)}, \delta_3^{(3)}) = (6, 5, 1)$$

Теперь в соответствии с формулой (20.6) выпишем решение системы (20.7) в параметрическом виде:

$$s_1(0) = \xi_1 + i_1 \delta_1^{(1)} + i_2 \delta_1^{(2)} + i_3 \delta_1^{(3)},$$

$$s_2(0) = \xi_2 + i_1 \delta_2^{(1)} + i_2 \delta_2^{(2)} + i_3 \delta_2^{(3)},$$

$$s_3(0) = \xi_3 + i_1 \delta_3^{(1)} + i_2 \delta_3^{(2)} + i_3 \delta_3^{(3)}$$

где $5 \leq i_1 \leq 6, 2 \leq i_2 \leq 4, 0 \leq i_3 \leq 2$

Таким образом, получаем 18 решений системы уравнений (20.8), представленных в формате $s(0) = [s_1(0), s_2(0), s_3(0)]'$:

$$[336]', [120]', [611]', [213]', [004]', [565]', [160]', [651]', [442]', [430]', [221]', [012]', [314]', [105]', [666]', [261]', [052]', [543].$$

Полученные решения интервальной системы уравнений (20.8) являются "кандидатами" в решение интервальной системы уравнений (20.7). Подставляя их в систему (20.7), можно убедиться, что только одно из этих 18 решений, $s(0) = [105]'$, удовлетворяет всем уравнениям (20.7). Поскольку это решение принадлежит множеству допустимых начальных состояний рассматриваемого ЛА, то оно и есть искомое решение интервальной диагностической задачи.

Генетический алгоритм для решения интервальной диагностической задачи

В предыдущем разделе было указано, что мощность множества всех решений интервальной системы уравнений с квадратной матрицей равна величине $z = \prod_{i=1}^n w(B_i)$. Из формулы видно, что эта мощность быстро растет с ростом ширины интервалов в правой части системы

(20.3). Даже при сравнительно небольших значениях упомянутых параметров величина z может быть столь велика, что множество всех решений системы становится практически необозримым. В [24] отмечается, что это свойство является атрибутом труднорешаемой задачи, причем не менее важным, чем экспоненциальное время ее решения. По этой причине, если мощность множества всех "кандидатов" в решение системы очень велика, возникает проблема нахождения среди них единственного решения интервальной диагностической задачи. В данном разделе рассматривается еще один подход к решению поставленной интервальной диагностической задачи, который можно рассматривать как "быстрый" метод ее решения.

Простейшим способом поиска какого-либо решения рассматриваемой задачи является перебор всех возможных допустимых начальных состояний ЛА и проверка того, удовлетворяют ли эти состояния системе линейных уравнений (20.3). Поскольку для ЛА, заданного над полем $GF(p)$, число таких всевозможных начальных состояний p^n , где n - размерность ЛА, то уже при $n = 30$ и $p = 2$ число состояний, подлежащих проверке, превышает 10^9 . Учитывая, что проверка допустимости состояния ЛА является нетривиальной (вычисление матричных произведений), можно констатировать, что проверка всех возможных состояний неосуществима за приемлемое время.

Накопленный опыт практического решения задач переборного типа показал достаточно высокую эффективность применения генетических алгоритмов (ГА). Ниже будет показано, каким образом можно построить генетический алгоритм для рассматриваемой в этом разделе задачи, и оценена его эффективность.

Заметим, что используемые далее, но не определенные здесь понятия, связанные с генетическими алгоритмами, заимствованы из [5], [32].

Принцип работы генетических алгоритмов основан на моделировании некоторых механизмов популяционной генетики: манипулировании хромосомным набором при формировании генотипа новой биологической особи путем наследования участков хромосомных наборов родителей (кроссинговер), случайном изменении генотипа (мутация). Другим важным механизмом, заимствованным у природы, является процедура естественного отбора, направленная на улучшение

от поколения к поколению приспособленности членов популяции к "выживанию" особей, обладающих определенными признаками.

Применение генетических алгоритмов требует подходящего выбора способа кодирования решения, представленного в виде хромосомы, поскольку их эффективность во многом зависит от этого. Кроме того, целевая функция, применяемая для оценки приспособленности хромосомы, должна быть выбрана соответствующим образом, чтобы адекватно отразить конечную цель поиска. Важными моментами при использовании генетических алгоритмов являются также реализация операторов отбора, скрещивания и мутации и выбор вероятности их применения.

В данной работе используется простой генетический алгоритм (ПГА), который был впервые описан Гольдбергом на основе работ Холланда. Изложение сути этого алгоритма приведено в [32]. ПГА использует операторы отбора (селекции), скрещивания (кроссинговера) и мутации. Механизм ПГА несложен, и ниже кратко описаны его этапы.

1. Генерация случайной популяции, состоящей из N хромосом, длина каждой из которых равна l (каждая хромосома - кандидат на роль решения задачи). Эта начальная популяция образует поколение с номером 0.
2. Вычисление функции приспособленности для каждой хромосомы в популяции.
3. Выполнение следующих шагов, пока не будет получено N потомков:
 - Выбор пары родительских хромосом из текущей популяции, при этом вероятность выбора больше у тех хромосом, чья приспособленность выше. Выбор делается с возвращением, т. е. одна и та же хромосома может быть выбрана в качестве родителя более одного раза.
 - С вероятностью P_c (вероятность кроссинговера) производится кроссинговер пары в произвольно выбранной точке (одноточечный кроссинговер), которая выбирается с равномерно распределенной вероятностью, и формируются два потомка. Если кроссинговер не производится, то формируются два потомка, которые являются точными копиями родителей.

- Производится мутация двух потомков в каждом гене с вероятностью P_m (вероятность мутации), и получившиеся хромосомы помещаются в новую популяцию.
 - Если N нечетное и количество потомков равно $N + 1$, то один произвольно выбранный потомок удаляется из новой популяции.
4. Текущая популяция замещается новой популяцией с увеличением номера поколения.
 5. Если номер текущего поколения меньше заданного максимального числа поколений, то переход к шагу 2, иначе окончание алгоритма.

Остановимся подробнее на отмеченных нами узловых точках построения генетического алгоритма. Обратимся вначале к способу кодирования решения в рассматриваемой интервальной диагностической задаче.

Большая часть существующей теории генетических алгоритмов (включая ранние работы Холланда) базируется на предположении, что хромосомы являются битовыми строками. Таким образом, двоичное кодирование решения является основным способом кодирования решений при использовании генетических алгоритмов.

Каждому допустимому решению s сопоставим битовую строку, представляющую собой хромосому. Сопоставление осуществим следующим образом: вектору $s = [s_1, s_2, \dots, s_n]'$ поставим в соответствие число в двоичной системе, которое равно числу s_1, s_2, \dots, s_n в системе счисления с основанием p . Такая схема сопоставления является взаимно-однозначной, что позволяет проводить как кодирование, так и декодирование хромосомы.

Заметим, что при $p = 2$ используются системы счисления по основанию 2 и перевод производить не требуется. Если $p > 2$, то число бит, необходимых для представления решения, отличается от n в сторону увеличения.

При таком способе кодирования длина хромосомы l определяется следующим образом:

$$l = \begin{cases} n * k, & \text{если } p = 2^k, \\ \lceil \log_2 p^n + 1 \rceil, & \text{если } p \neq 2^k \end{cases}$$

Пример 1.

Пусть $p = 2, n = 4$, тогда $l = 4$. Решению $s = [1, 0, 1, 1]'$ соответствует хромосома 1011.

Пусть $p = 3, n = 4$, тогда $l = 7$. Решению $s = [1, 0, 2, 1]'$ соответствует хромосома 0100010.

Следующим важным моментом при реализации генетического алгоритма является определение целевой функции, которая "направляет" поиск решения.

Целевая функция используется для оценивания приспособленности хромосомы. Заметим, что для вычисления значения целевой функции необходимо произвести декодирование хромосомы в соответствующее ей решение.

Пусть $s = [s_1, s_2, \dots, s_N]'$ - решение, которое необходимо оценить. Так как это решение представляет собой предполагаемое начальное состояние, то получаем $s(0) = s$. Осуществляя подстановку $s(t)$ в систему уравнений (1.1)-(1.2) при $t = \overline{0, t_{max}}$ где $t_{max} + 1$ равно длине диагностической последовательности, получим набор из $t_{max} + 1$ выходных векторов размерности $m : y_c(0), y_c(1), \dots, y_c(t_{max})$.

Целевую функцию определим следующим образом:

$$Q(s) = \sum_{t=0}^{t_{max}} Q_t(y_c(t)) \quad (20.9)$$

где

$$Q_t(y_c(t)) = \begin{cases} 0, & \text{если } y_c(t) \in y(t), \\ \sum_{i=1}^n \min(|\bar{y}(t) - y_c(t)|, |\underline{y}_i(t) - y_c(t)|), & \text{если } y_c(t) \notin y(t) \end{cases}$$

При таком выборе целевая функция будет неотрицательной во всей области допустимых значений и принимать значение 0 лишь на решениях интервальной диагностической задачи.

Однако в классическом понимании целевой функции как функции приспособленности необходимо, чтобы более приспособленные особи оценивались выше, чем менее приспособленные. Этого легко достичь, если произвести нормализацию целевой функции и обратить ее:

$$Q(s) = \sum_{t=0}^{t_{max}} Q'_t(y_c(t)) \quad (20.11)$$

где

$$Q'_t(y_c(t)) = 1 - \frac{Q_t(y_c(t))}{n(p-1)} \quad (20.12)$$

Таким образом, целевая функция $Q(s)$ будет принимать значения в диапазоне от 0 до 1. При этом большему значению целевой функции соответствует большая степень приспособленности хромосомы, т. е. данная целевая функция является "истинно" функцией приспособленности. Очевидно, что на решениях интервальной диагностической задачи целевая функция будет принимать значение 1.

Теперь коснемся вопроса определения вероятностей применения операторов ПГА. Как же их определять? На этот вопрос пока нет простого и однозначного ответа. В докторской диссертации де Джонга (1975), посвященной исследованию генетических алгоритмов, отмечается, что для хорошей работы ГА необходимо выбирать большую вероятность кроссинговера и малую вероятность мутации (например, обратно пропорционально размеру популяции). Следует заметить, что это правило не всегда оправдывает себя, например, для областей, связанных с исследованиями социальных адаптивных систем. Однако в большинстве технических задач это правило работает, и мы в нашей работе будем его использовать.

Размер популяции и максимальное количество поколений определяют объем вычислений программы. Чем они больше, тем больше вычислений проводится и большее количество хромосом будет просмотрено в поисках решения. Очевидно, что эти параметры надо делать максимальными. Однако с ростом этих параметров растет и время работы программы. Поэтому выбор их в большей степени определяется вычислительной мощностью и предельным временем, за которое программа должна найти решение.

В ходе исследования поставленной интервальной диагностической задачи была написана программа, которая решает эту задачу. Основу программы представляет реализация простого генетического алгоритма на языке C. Эта реализация имеет имя SGA-C v. 1.1 (Simple Genetic Algorithm on C), и представляет собой интерпретацию кода из книги Goldberg D.E. "Genetic Algorithms in Search, Optimization, and Machine Learning" (1989) на языке C.

SGA-C реализует простой генетический алгоритм, который содержит операторы отбора, скрещивания и мутации, и механизм задания вероятности применения этих операторов. При этом SGA-C позволяет исследователю сделать настройку для решения его специфических задач с помощью набора функций в файле app.c. Для решения поставленной интервальной диагностической задачи были написаны следующие функции:

- Функция начальной инициализации пользовательских данных: производит считывание основных параметров (характеристики поля P , длины диагностической последовательности), характеристических матриц ЛА, входных и выходных векторов.
- Функция приспособленности `objfunc ()`: производит вычисление функции $Q(s)$ по формулам (20.10)-(20.12).
- Функция отчета `app_report ()`: производит печать всех найденных в ходе работы программы решений и их количества.
- Функция `chrom2matrix ()`: производит декодирование хромосомы в матрицу.
- Функции `matrix_sum ()`, `matrix_mult ()`: реализуют операции сложения и умножения матриц.

Эта программа для нахождения решения интервальной диагностической задачи на основе генетического алгоритма размещена в Интернете по адресу: ссылка: <http://menace.cs.sgu.ru/sam/genetic/download/sga-c.auto.zip> .

Рассмотрим ряд примеров решения сформулированной диагностической задачи с применением данной программы.

Пример 2.

Решим интервальную диагностическую задачу для ЛА из примера 1.

Согласно описанной выше рекомендации, вероятность оператора кроссинговера P_c положим равной 0.1, а вероятность оператора мутации $p_m = 0.01$. Пусть популяция состоит из 20 хромосом, длина хромосомы равна 9 и максимальное количество поколений равно 100. При таких параметрах генетического алгоритма программой было найдено решение: $s(0) = [105]^T$. Как уже известно (пример 1), это решение является искомым решением интервальной диагностической задачи.

Как было сказано выше, вероятности операторов играют большую роль при работе генетического алгоритма. Чтобы проиллюстрировать это, данная задача решалась при разных вероятностях оператора скрещивания (P_c) и оператора мутации (P_m). В представленной ниже таблице содержится информация о нахождении решения интервальной диагностической задачи для разных значений P_c и P_m (в клетке таблицы стоит 0, если решение не найдено, 1 - если найдено).

Из этой таблицы можно сделать следующие выводы:

- для малых значений вероятности мутации (0.00-0.02) решение задачи находится далеко не всегда, что является свидетельством необходимости участия мутации в процессе работы генетического алгоритма;
- для достаточно большого числа наборов пар вероятностей (89 из 121, т. е. в 73 % случаев) алгоритм находит решение интервальной диагностической задачи.

$P_c P_m$	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09	0.10
-----------	------	------	------	------	------	------	------	------	------	------	------

0.00	0	0	1	0	1	1	0	1	0	1	1
0.05	0	1	0	1	1	0	0	1	1	1	1
0.10	0	1	1	1	0	1	1	1	1	1	1
0.15	0	1	1	1	1	1	1	1	1	1	1
0.20	0	0	1	1	1	1	1	1	1	1	1
0.25	0	1	0	1	1	1	1	1	1	1	1
0.30	0	1	0	1	1	1	1	1	1	1	1
0.35	0	0	0	0	1	1	1	1	1	1	1
0.40	0	0	0	1	1	1	1	1	1	1	1
0.45	0	0	0	1	1	1	1	1	1	1	1
0.50	0	0	1	1	0	0	1	1	1	1	1

Пример 3.

Рассмотрим ЛА из примера 1. Входные векторы оставим без изменений, а выходные векторы сделаем точными, т. е. рассмотрим случай вырожденных интервалов:

$$y(0) = \begin{bmatrix} [3, 3] \\ [4, 4] \end{bmatrix}, y(1) = \begin{bmatrix} [1, 1] \\ [3, 3] \end{bmatrix}, y(2) = \begin{bmatrix} [5, 5] \\ [5, 5] \end{bmatrix}$$

Заметим, что хотя выходные векторы заданы с помощью интервалов, однако представляют собой векторы с точными значениями. Таким образом, данный пример является примером классической диагностической задачи.

Используя таблицу из примера 2, положим вероятность оператора кроссинговера p_c равной 0.15, а вероятность оператора мутации $p_m = 0.07$. Пусть популяция состоит из 20 хромосом и максимальное количество поколений равно 200. При таких параметрах генетического алгоритма программой было найдено следующее решение:

$s(0) = [456]^T$. Найденное решение является одним из допустимых начальных состояний заданного ЛА, поэтому классическая диагностическая задача имеет решение, и это решение есть $s(0)$.

Таким образом, пример 3 иллюстрирует тот факт, что генетический алгоритм позволяет решать диагностическую задачу как в интервальной постановке, так и в классическом случае.

Главными преимуществами генетического алгоритма при решении разнообразных задач являются небольшой объем используемой памяти и высокая скорость получения решения. Например, при $p = 2, n = 30$, программа с применением ГА нашла решение менее чем за 1 секунду своей работы (был взят компьютер с процессором Celeron 900 MHz). Другая программа, которая осуществляла поиск решения полным перебором, это же решение нашла только через несколько десятков часов при использовании того же компьютера. Таким образом, генетический подход к решению интервальной диагностической задачи не уступает точному алгебраическому подходу как по скорости, так и по качеству решения поставленной задачи.

Вопросы и упражнения

1. Сформулируйте диагностическую задачу для линейного автомата в интервальной постановке.
2. Опишите метод построения интервального диагностического дерева для ЛА с заданным множеством допустимых начальных состояний.
3. Укажите способ сведения решения задачи о поиске неизвестного начального состояния ЛА в интервальной постановке к решению системы линейных уравнений с интервальной правой частью.
4. Сформулируйте теорему о виде решения системы $AX = B$ с интервальной правой частью.
5. Опишите идею построения генетического алгоритма для решения диагностической задачи для ЛА в интервальной постановке.
6. Приведите вид целевой функции, используемой в описанном генетическом алгоритме.

Эксперименты с билинейными автоматами по распознаванию состояний

Объектом исследования являются билинейные автоматы. Для автоматов такого типа приведены критерии существования синхронизирующих, установочных и диагностических последовательностей, сформулированные в терминах их характеристических матриц. Предложены и обоснованы аналитические методы построения перечисленных последовательностей и исследованы их свойства.

Перейдем теперь к описанию используемой нами модели БА. Условимся, что для входных и выходных символов БА, а также для состояний БА будут сохранены те же обозначения, что и для ЛА.

Предполагается, что БА \tilde{A} задан над полем $GF(p) = \{0, 1, \dots, p-1\}$. Пространство состояний БА будем обозначать через S_n , где $|S_n| = p^n$, а число n будем, как и в случае ЛА, называть размерностью БА.

Далее рассматриваются БА, функционирование которых описывается следующими системами уравнений состояний и выходов соответственно:

$$\bar{s}(t+1) = A\bar{s}(t) + (\sum_{i=1}^l F_i u_i(t))\bar{s}(t) + B\bar{u}(t) \quad (21.1)$$

$$\bar{y}(t) = C\bar{s}(t) + (\sum_{i=1}^l C_i u_i(t))\bar{s}(t) + D\bar{u}(t) \quad (21.2)$$

где A, F_i - матрицы размерности $n \times n$, B - размерности $n \times l$; C, G_i - размерности $m \times n$; D - размерности $m \times l$.

Упомянутые матрицы будем называть далее характеристическими матрицами БА. Элементами всех этих матриц являются элементы поля $GF(p)$.

Введем следующие обозначения:

$$I(\bar{u}(t)) = \sum_{i=1}^t F_i u_i(t),$$

$$J(\bar{u}(t)) = \sum_{i=1}^t G_i u_i(t)$$

Пусть задана входная последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ и пусть $\bar{s}(0)$ есть начальное состояние БА. Методом математической индукции можно доказать, что конечное состояние БА и его реакция вычисляются соответственно по формулам

$$\bar{s}(t+1) = \left(\prod_{i=0}^t [A + I(\bar{u}(t-i))] \right) \bar{s}(0) + \left(\prod_{i=0}^t [A + I(\bar{u}(t-1))] \right) B\bar{u}(0) \dots + [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t)$$

$$y(t) = ([C + J(\bar{u}(t))]) \left(\prod_{i=0}^t [A + I(\bar{u}(t-i-1))] \right) \bar{s}(0) + [C + J(\bar{u}(t))] * \prod_{i=0}^t [A + J(\bar{u}(t-1))] B\bar{u}(0) + \dots + [C + J(\bar{u}(t))] B\bar{u}(t-1) + D\bar{u}(t)$$

Синхронизирующие последовательности

Применительно к БА \tilde{A} определение 1.1 СП $u(0), \bar{u}(1), \dots, \bar{u}(t)$ (в лекции 1) принимает следующий вид:

$$\forall \bar{s}_1, \bar{s}_2 \in S_n \left\{ \prod_{i=0}^t [A + I(\bar{u}(t-i))] \bar{s}_1 + \prod_{i=0}^t [A + I(\bar{u}(t-i))] B\bar{u}(0) + \dots \right.$$

$$+ [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t) = \left\{ \prod_{i=0}^t [A + I(\bar{u}(t-i))] \bar{s}_2 + \right.$$

$$+ [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t) = \prod_{i=0}^t [A + I(\bar{u}(t-i))] \bar{s}_2 +$$

$$+ \prod_{i=0}^{t-1} [A + I(\bar{u}(t-i))] B\bar{u}(0) + \dots + [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t)$$

Перенося в левую часть равенства (21.5) все выражения, стоящие в правой его части, получим

$$\forall \bar{s} \neq [0], |p_{i=0}^t[A + I(\bar{u}(t - i))]\bar{s} = [0] \quad (21.6)$$

где $[0]$ - нулевая матрица (в частности, вектор) соответствующей размерности.

Теорема 21.1. Для того чтобы входная последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ была СП для БА \tilde{A} , необходимо и достаточно, чтобы выполнялось равенство

$$\prod_{i=0}^t [A + I(\bar{u}(t - i))] = [0] \quad (21.7)$$

Доказательство. Докажем необходимость условия (21.7), поскольку его достаточность очевидна. Обозначим через $H = [h_{ij}]$ матрицу, представляющую собой левую часть (21.7). Пусть $u(0), \bar{u}(1), \dots, \bar{u}(t)$ есть СП, но тогда выполняется равенство (21.6). Поскольку \bar{s} - произвольное состояние, то положим $\bar{s} = [1, 0, \dots, 0]'$, при котором (21.6) становится равенством $[h_{11}h_{21} \dots h_{n1}]' = [0]$. Из этого равенства следует, что $h_{i1} = 0, i = \overline{1, n}$. Полагая далее \bar{s} равным векторам $[0, 1, 0, \dots, 0]', \dots, [0, 0, \dots, 1]'$, аналогичными рассуждениями приходим к заключению, что и все остальные столбцы матрицы H , как и ее первый столбец, нулевые. Таким образом, из (21.6) следует (21.7), что и требовалось доказать.

Из (21.7) следует справедливость следующего утверждения.

Теорема 21.2. Для того чтобы входная последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ была СП для БС \tilde{A} , достаточно, чтобы по крайней мере для одного из значений $i = 0, 1, \dots, t$ выполнялось равенство

$$[A + I(\bar{u}(t - i))] = [0] \quad (21.8)$$

Приведем еще одно достаточное условие существования СП. Предварительно напомним, что квадратная матрица называется верхней (нижней) треугольной, если все ее элементы, лежащие на

главной диагонали и ниже (выше) ее, нулевые.

Теорема 21.3. Если характеристические матрицы A и $F_i, i = \overline{1, l}$, БС \tilde{A} являются верхними (нижними) треугольными, то для этой БС существуют СП длины, не большей n , где n - число строк и столбцов упомянутых матриц.

Доказательство. Пусть A и $F_i, i = \overline{1, l}$, являются верхними треугольными матрицами. Тогда каждый сомножитель в произведении (21.7) также представляет собой матрицу того же типа. Условимся нумеровать диагонали этих матриц, лежащие выше главной диагонали и параллельные ей, числами $1, 2, \dots, n - 1$. Непосредственными вычислениями легко убедиться, что все элементы первой диагонали матрицы-произведения двух верхних треугольных матриц равны нулю. Если эту матрицу вновь умножить на верхнюю треугольную матрицу, т. е. взять произведение трех сомножителей из (21.7), то в результате получится матрица, у которой нулевыми будут элементы первой и второй диагонали. Отсюда методом индукции можно доказать, что произведение n штук верхних треугольных матриц даст нулевую матрицу. Тогда на основании теоремы 21.1 рассматриваемая БА имеет СП, длина которой не превосходит n .

Для нижних треугольных матриц соответствующее утверждение доказывается аналогично.

Поскольку БА представляет собой автомат Мили, для построения для нее УП (в частности, СП) и ДП применимы известные методы, описанные в [18], которые, однако, достаточно трудоемки, так как базируются на использовании конструкции дерева преемников. Ниже описываются аналитические методы построения различных типов последовательностей для БА, менее трудоемкие, чем методы из [18].

Опишем метод построения СП. Обратимся к равенству (21.7) и на его основе организуем пошаговый процесс поиска СП, начиная с СП длины 1 (при $t = 0$), затем СП длины 2 (при $t = 1$) и т. д. При варьировании величины t равенство (21.7) будет принимать последовательно следующий вид:

поля $GF(p)$, над которым рассматриваются такие системы, решение может быть найдено либо перебором, либо с помощью некоторого частного метода, который может быть предложен с учетом специфики конкретной системы.

Проиллюстрируем метод на примере БА над полем $GF(2)$, у которой $n = 3, l = 2$, а характеристические матрицы имеют следующий вид:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, F_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, F_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Первое равенство в (21.9) дает следующую систему уравнений:

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} u_1(0) + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} u_2(0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Легко убедиться, что она несовместна, и потому выписываем вторую систему из (21.9):

$$\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} u_1(0) + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} u_2(0) \right) \cdot \left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} u_1(1) + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} u_2(1) \right) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

После выполнения преобразований система примет вид

$$\begin{bmatrix} 1 + (u_1(0) + u_2(0))u_2(1) & 0 & u_1(1) + u_2(1) + u_2(1)(u_1(0) + u_2(0)) \\ 0 & u_1(0)u_1(1) & 0 \\ u_2(0) + u_2(0)u_2(1) & 0 & u_2(0)(u_1(1) + u_2(1)) + u_2(0)u_2(1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

В конечном счете, приравнявая каждый элемент последней матрицы

нулю, получим систему уравнений

$$\begin{aligned} (u_1(0)u_2(0))u_2(1) + 1 &= 0, \\ u_1(1) + u_2(1) + u_2(1)(u_1(0) + u_2(0)) &= 0, \\ u_2(0) + u_2(0)u_2(1) &= 0, \\ u_2(0)(u_1(1) + u_2(1)) + u_2(0)u_2(1) &= 0 \end{aligned}$$

Напомним, что решение мы ищем в поле $GF(2)$ и потому операция "+" - это сложение по модулю 2. Легко убедиться, что эта система имеет два решения:

1. $u_1(0) = 1, u_2(0) = 0, u_1(1) = 0, u_2(1) = 1$;
2. $u_1(0) = 0, u_2(0) = 1, u_1(1) = 0, u_2(1) = 1$.

Таким образом, рассматриваемый БА имеет две СП:

$$\bar{u}(0) = [1, 0]', \bar{u}(1) = [0, 1]' \text{ и } \bar{u}(0) = [0, 1]', \bar{u}(1) = [0, 1]'.$$

Задачи синхронизации и установки автоматов - это разновидности задачи управления дискретной системой (ДС), которая в общем виде формулируется следующим образом: для рассматриваемой ДС найти такую входную последовательность, которая переводит ее из состояния \bar{s}_1 в состояние \bar{s}_2 . Для автоматов, заданных графом переходов, эта задача сводится к задаче поиска путей на графе между двумя заданными вершинами, методы решения которой известны. Однако для автоматов с большим числом состояний проблема построения графа переходов сама по себе является трудоемкой. Покажем, что для БА может быть предложен аналитический метод решения упомянутой задачи, не требующий знания графа переходов.

Поиск требуемой входной последовательности осуществим следующим образом. В (21.1) положим $t = 0$ и вместо $\bar{s}(0)$ подставим состояние \bar{s}_1 , а вместо $\bar{s}(1)$ - состояние \bar{s}_2 . Полученное выражение будем рассматривать как систему линейных алгебраических уравнений относительно неизвестных $u_1(0), \dots, u_l(0)$. Очевидно, что если эта система совместна, то ее решению соответствует искомая входная последовательность. Если таких решений будет несколько, то это

говорит о существовании нескольких путей перехода из $\bar{s}_1 >$ в \bar{s}_2 , а отсутствие решений - о невозможности требуемого перехода с помощью входной последовательности длины 1. В последнем случае сделаем попытку найти соответствующую последовательность длины 2, положив в (21.3) $t = 1$ и заменив $\bar{s}(t)$ и $\bar{s}(t + 1)$ соответственно на \bar{s}_1 и \bar{s}_2 . Если полученная система вновь окажется несовместной, то продолжим описанный процесс далее аналогичным образом. Если до значения $t = \omega - 2$ включительно, где ω - число состояний БА, все последовательно получаемые системы линейных уравнений окажутся несовместными, то это означает, что переход из \bar{s}_1 в \bar{s}_2 невозможен, поскольку если соответствующий путь существует, то длина его не может превосходить $\omega - 1$.

Проиллюстрируем описанный метод на примере БА над полем $GF(2)$, уже рассмотренном выше, где характеристическая матрица B имеет вид

$$B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Пусть требуется найти входную последовательность, переводящую заданную БС из состояния $\bar{s}_1 = [1, 1, 1]'$ в состояние $\bar{s}_2 = [0, 0, 1]'$.

Выпишем систему уравнений, используя выражение (6.1):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} u_1(0) + \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix} u_2(0) \right) + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \times \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Выполнив соответствующие преобразования, в итоге получим систему

$$\begin{bmatrix} u_2(0) + 1 \\ u_1(0) + u_2(0) \\ u_1(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

Эта система имеет единственное решение $u_1(0) = 1, u_2(0) = 1$, следовательно, искомая входная последовательность есть $[1, 1]'$.

Установочные последовательности

Условия существования такой последовательности даются следующей теоремой.

Теорема 21.4. Для того чтобы последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ была УП для БА \tilde{A} , необходимо и достаточно, чтобы для каждого ненулевого состояния $\bar{s} \in S_n$ выполнялось по крайней мере одно из двух условий:

1. $\bigwedge_{d=0}^t [C + J(\bar{u}(d))] \prod_{i=0}^{d-1} [A + I(\bar{u}(d - i - 1))] \bar{s} \neq 0$;
2. $\prod_{i=0}^t [A + I(\bar{u}(t - i))] \bar{s} = [0]$

Здесь знак $\bigwedge_{d=0}^t (\bigvee_{d=0}^t)$ означает дизъюнкцию (конъюнкцию) выражений, стоящих за этим знаком и получаемых при изменении индекса d от 0 до t . Условимся о следующем: если при вычислении $I(\bar{u}(\nu))$ и $J(\bar{u}(\nu))$ окажется, что $\nu:0$ при некоторых i и d , то в соответствующем выражении сомножители $[A + I(\bar{u}(\nu))]$ и $[C + J(\bar{u}(\nu))]$ полагаются равными единичной матрице.

Доказательство. Применительно к БА \tilde{A} определение 1.2 УП принимает следующий вид: последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ называется УП, если

$$\begin{aligned}
& \forall \bar{s}_1, \bar{s}_2 \in S_n, \wedge_{d=0}^t \{ [C + J(\bar{u}(d))] \prod_{i=0}^{d-1} [A + I(\bar{u}(d-i-1))] \} s_1 + \dots \\
& \dots + [C + J(\bar{u}(d))] B\bar{u}(d-1) + D\bar{u}(d) = [C + J(\bar{u}(d))] \prod_{i=0}^{d-1} [A + I(\bar{u}(d-i-1))] \bar{s}_2 + \dots \\
& \dots = [C + j(\bar{u}(d))] B\bar{u}(0) + \dots + [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t) = \\
& = \prod_{i=0}^{t-1} [A + I(\bar{u}(t-i))] B\bar{u}(0) + \dots + [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t) = \\
& = \prod_{i=0}^{t-1} [A + I(\bar{u}(t-i))] \bar{s}_2 + \prod_{i=0}^{t-1} [A + I(\bar{u}(t-i))] B\bar{u}(0) + \dots + [A + I(\bar{u}(t))] B\bar{u}(t-1) + B\bar{u}(t)
\end{aligned}$$

Преобразуем каждое в отдельности равенство, стоящее до и после знака импликации, перенеся все в левые их части и осуществив соответствующие сокращения, в результате чего получим

$$\begin{aligned}
& \forall \bar{s}_1, \bar{s}_2 \in S_n, [\wedge_{d=0}^t [C + J(\bar{u}(d))] \prod_{i=0}^t [A + I(\bar{u}(d-i-1))] (\bar{s}_1 - \bar{s}_2) = [0]] \rightarrow \\
& \rightarrow \prod_{i=0}^t [A + I(\bar{u}(t-i))] (\bar{s}_1 - \bar{s}_2) = [0]
\end{aligned}$$

Поскольку \bar{s}_1 и \bar{s}_2 - произвольные не совпадающие между собой состояния из S_n , то и состояние $\bar{s} = \bar{s}_1 - \bar{s}_2$ также может быть любым ненулевым состоянием из S_n . Отсюда следует, что последний предикат эквивалентен предикату

$$\forall \bar{s} \neq [0] [\wedge_{d=0}^t [C + J(\bar{u}(d))] \prod_{i=0}^{d-1} [A + I(\bar{u}(d-i-1))] \bar{s} = [0]] \rightarrow \prod_{i=0}^t [A + I(\bar{u}(t-i))] \bar{s} = [0]$$

Поскольку $X \rightarrow Y$ эквивалентно $\bar{X} \vee Y$, то последний предикат эквивалентен следующему предикату

$$\forall \bar{s} \neq [0] [\vee_{d=0}^t [C + J(\bar{u}(d))] \prod_{i=0}^{d-1} [A + I(\bar{u}(d-i-1))] \bar{s} \neq [0]] \vee \prod_{i=0}^t [A + I(\bar{u}(t-i))] \bar{s} = [0]$$

Из этого предиката и следует справедливость теоремы.

Проиллюстрируем предложенный метод на примере БА над полем $GF(2)$, у которого $n = 3, l = 2, m = 2$, а характеристические матрицы таковы:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, F_1 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, F_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \\ C = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, D = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, G_1 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, G_2 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Первое условие теоремы 21.4 при $d = 0$ порождает систему уравнений

$$[C + J(\bar{u}(0))] \bar{s} = \begin{bmatrix} u_1(0) & u_2(0) & u_1(0) + u_2(0) + 1 \\ u_2(0) & u_1(0) + 1 & u_1(0) + u_2(0) + 1 \end{bmatrix} * \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = [0]$$

Рассмотрим теперь все возможные векторы $\bar{u}(0) : [0, 0]', [1, 0]', [1, 0]', [1, 1]'$. Их подстановка в выписанную нелинейную систему порождает четыре следующих системы:

$$begin\ cases\ s_3 = 0,$$

$$s_2 + s_3 = 0, \ begin\ cases\ s_2 = 0,$$

$$s_1 + s_2 = 0, \ begin\ cases\ s_1, \ 0 = 0, \ begin\ cases\ s_1 + s_2 + s - 3 = 0,$$

$$s_1 + s_3 = 0$$

Каждая из этих систем относительно неизвестных s_1, s_2, s_3 имеет не единственное решение, и поэтому условие 1 теоремы 21.4 для всех входных последовательностей длины 1 не выполняется. Перейдем теперь к проверке условия 2 той же теоремы, которое при $d = 0$ порождает систему

$$[A + I(\bar{u}(0))] \bar{s} = \begin{bmatrix} 1 & 0 & u_1(0) + u_2(0) \\ 0 & u_1(0) + 1 & 0 \\ u_2(0) & 0 & u_2(0) + 1 \end{bmatrix} * \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = [0]$$

Легко проверить, что эта система имеет единственное (нулевое) решение при любом $\bar{u}(0)$, т. е. условие 2 также не выполняется. Поэтому перейдем к поиску УП длины 2, полагая $d = 1$.

Условие 1 теоремы 21.4 порождает следующую систему уравнений:

$$\begin{aligned} \text{begincases} C + J(\bar{u}(0)) \end{aligned} \bar{s} = [0], \\ C + J(\bar{u}(1)) [A + I(\bar{u}(0))] \bar{s} = [0]$$

В координатной записи эта же система имеет следующий вид:

$$\begin{bmatrix} u_1(0) & u_2(0) & u_1(0) + u_2(0) + 1 \\ u_2(0) & u_1(0) + 1 & u_1(0) + u_2(0) + 1 \end{bmatrix} * \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = [0],$$

$$\begin{bmatrix} [u_1(1) + u_2(1) + 1]* & u_2(1)[u_1(0) + 1] & u_1(1)[u_1(0) + u_2(0)] + \\ *u_2(0) + u_1(1) & & + u_1(1) + u_2(1) + 1][u_2(0) + 1] \\ [u_1(1) + u_2(1) + 1]* & [u_1(1) + 1]* & u_2(1)[u_1(0) + u_2(0)] + \\ *u_2(0) + u_2(0)1 & *[u_1(0) + 1] & + [u_1(1) + u_2(1) + 1][u_2(0) + 1] \end{bmatrix} * \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = [0]$$

Теперь необходимо рассмотреть все возможные пары векторов $\bar{u}(0)$, $\bar{u}(1)$ и, подставляя их в выписанную выше систему, получить множество систем линейных уравнений относительно неизвестных s_1, s_2, s_3 . Вычисления показывают, что, например, для пар входных векторов $\bar{u}(0) = [1, 1]'$, $\bar{u}(1) = [0, 0]'$ и $\bar{u}(0) = [0, 0]'$, $\bar{u}(1) = [1, 0]'$ первая подсистема $[C + J(\bar{u}(0))] \bar{s} = [0]$ имеет ненулевое решение ($\bar{s} = [1, 0, 1]'$ и $\bar{s} = [1, 0, 0]'$ соответственно), а вторая подсистема для этих же пар входных векторов имеет единственное (нулевое) решение. Последний факт означает, что условие 1 теоремы выполняется, следовательно, обе приведенные входные последовательности являются для рассматриваемого БА установочными. Заметим, что этот БА имеет и другие УП длины 2, которые находятся аналогично.

Для подтверждения сказанного, в приведенной таблице показаны реакции рассматриваемого БА на две найденные входные последовательности и соответствующие конечные состояния. Состояния БА в таблице закодированы в виде десятичного эквивалента двоичного представления вектора-состояния. Так, например, состояние $\bar{s} = [1, 0, 0]'$ закодировано числом 4, состояние $\bar{s} = [1, 0, 1]'$ - числом 5 и т. д.

Начальное состояние	Реакция БА	Конечное состояние	Реакция БА	Конечное состояние
0	3, 2	7	0, 2	5
1	0, 2	7	3, 2	0
2	1, 2	7	1, 2	5
3	2, 2	7	2, 2	0
4	0, 1	2	0, 0	1
5	3, 1	2	3, 0	5
6	2, 1	2	1, 0	1
7	1, 1	2	2, 0	4

Аналогичная кодировка в таблице принята для входных и выходных векторов. Так, УП $[0, 0]'$, $[1, 0]'$ закодирована как последовательность 0, 2; а УП $[1, 1]'$, $[0, 0]'$ - как последовательность 3, 0.

Диагностические последовательности

Условия существования такой последовательности даются следующей теоремой.

Теорема 21.5. Для того чтобы последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ была ДП для БА \tilde{A} , необходимо и достаточно, чтобы

$$\text{rank} \begin{bmatrix} C + J(\bar{u}(0)) \\ [C + J(\bar{u}(1))][A + I(\bar{u}(0))] \\ \dots \\ [C + J(\bar{u}(t))]\prod_{i=0}^{t-1}[A + I(\bar{u}(t-i-1))] \end{bmatrix} = n, \text{ где } n \text{ размерность БА}$$

Доказательство. Применительно к БА определение 1.3 ДП принимает следующий вид: последовательность $u(0), \bar{u}(1), \dots, \bar{u}(t)$ называется ДП, если

$$GF(p).$$

Выписанную совокупность соотношений можно интерпретировать как систему линейных уравнений относительно неизвестных $s_1(0), \dots, s_n(0)$, являющихся координатами вектора-состояния $\bar{s}(0)$. Однозначность восстановления начального состояния $\bar{s}(0)$ по наблюдаемой реакции $\bar{y}(0), \bar{y}(1), \dots, \bar{y}(t)$ при известной входной последовательности $u(0), \bar{u}(1), \dots, \bar{u}(t)$ означает, что представленная система линейных уравнений должна иметь единственное решение. Из [33] известно, что для этого необходимо и достаточно, чтобы ранг этой системы был равен числу ее неизвестных. Отсюда и следует справедливость теоремы.

Опишем теперь метод нахождения ДП для заданного БА, основанный на теореме 21.5. Метод состоит в переборе всевозможных входных слов $u(0), \bar{u}(1), \dots, \bar{u}(t)$ последовательно для $t = 1, 2, 3, \dots$ и вычислении для очередного проверяемого слова ранга матрицы, фигурирующей в теореме 21.5. Если этот ранг окажется равным n , то проверяемое входное слово является искомым ДП. В противном случае указанный процесс продолжается далее до тех пор, пока значение параметра t не достигнет своего предельного значения, равного верхней границе длины минимальной ДП для рассматриваемого БА. Если и при этом значении t ДП не будет найдена, то для рассматриваемого БА ее не существует. Проиллюстрируем этот метод на примере БА, для которого выше осуществлялся поиск УП. Для этого БА приведем некоторые "составляющие" матрицы в условиях теоремы 21.5:

$$[C + J(\bar{u}(0))] = \begin{bmatrix} u_1(0) & u_2(0) & u_1(0) + u_2(0) + 1 \\ u_2(0) & U_1(0) + 1 & u_1(0) + u_2(0) + 1 \end{bmatrix},$$

$$C + J(\bar{0})$$

$$[A + I(\bar{u}(0))] = \begin{bmatrix} [u_1(1) + u_2(1)] * u_2(0) + u_1(1) & u_2(1)[u_1(0) + 1] & u_1(1)[u_1(0) + u_2(0)] + [u_1(1) + u_2(1) + 1][u_2(0) + 1] \\ [u_1(1) + u_2(1)] * u_2(0) + u_1(1) & [u_1(1) + 1][u_1(0) + 1] & u_2(1)[u_1(0) + u_2(0)] + [u_1(1) + u_2(1) + 1][u_2(0) + 1] \end{bmatrix}$$

Размерность матрицы $[C + J(\bar{u}(0))]$ равна 2×3 , следовательно, ее ранг не может превышать 2. Но тогда, поскольку для рассматриваемой БС $n = 3$, условие теоремы 21.5 при $t = 0$ не может быть выполнено.

В связи с этим перейдем к перебору всевозможных входных слов длины 2 (полагая $t = 1$). Рассмотрим, например, входное слово $\bar{u}(0), \bar{u}(1) = 3, 0$. Для него матрица из условия теоремы 21.5 имеет вид

$$\begin{bmatrix} C + J(u(0)) \\ [C + J(\bar{u}(0))][A + I(\bar{u}(0))] \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

Поскольку ранг этой матрицы равен 3, проверяемое слово 3, 0 есть ДП. В этом можно убедиться по данным, приведенным в таблице. Для другого входного слова $\bar{u}(0), \bar{u}(1) = 0, 2$ та же матрица из условий теоремы равна

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Ранг ее также равен 3, следовательно, входное слово 0, 2 также является ДП для рассматриваемого БА. Таким образом, оба входных слова 3, 0 и 0, 2 являются для рассматриваемого БА одновременно УП и ДП.

Вопросы и упражнения

1. Приведите определение билинейного автомата и выпишите уравнения состояний и выходов, описывающие его функционирование.
2. Дайте определения синхронизирующей, установочной и диагностической последовательностей билинейного автомата в терминах его характеристических матриц.
3. Сформулируйте и докажите критерий существования синхронизирующей последовательности для билинейного автомата.
4. Опишите аналитический метод построения синхронизирующей последовательности для билинейного автомата.
5. Сформулируйте и докажите критерий существования

установочной последовательности для билинейного автомата.

6. Опишите аналитический метод построения установочной последовательности для билинейного автомата.
7. Сформулируйте и докажите критерий существования диагностической последовательности для билинейного автомата.
8. Опишите аналитический метод построения диагностической последовательности для билинейного автомата.

Разновидности экспериментов с билинейными автоматами

Исследованы эксперименты по распознаванию неизвестного входного слова билинейного автомата и некоторые модификации понятия билинейного автомата без потери информации. Рассмотрены различные типы билинейных автоматов с запаздыванием и для них приведены критерии существования синхронизирующих, установочных и диагностических последовательностей.

Эксперименты по распознаванию неизвестного входного слова

БА \tilde{A} назовем БА без потери информации из состояния \bar{s} (БПИ- \bar{s}), если, зная это состояние и наблюдая выходную последовательность на любую неизвестную входную последовательность, последнюю можно определить однозначно.

По формуле (21.2) имеем

$$\bar{s}(0) = C\bar{s}(0) + \left(\sum_{i=1}^l G_i u_i(0) \right) \bar{s}(0) + D\bar{u}(0)$$

Преобразуем это равенство следующим образом:

$$\sum_{i=1}^l G_i \bar{s}(0) u_i(0) + D\bar{u}(0) = \bar{u}(0) - C\bar{s}(0) \quad (22.1)$$

Будем интерпретировать (22.1) как систему линейных уравнений относительно неизвестных $u_1(0), \dots, u_l(0)$, являющихся координатами вектора $\bar{u}(0)$. Условимся матрицу этой системы обозначать через $L(\bar{s}(0))$ и называть матрицей распознавания для состояния $\bar{s}(0)$.

Будем говорить, что состояние \bar{s}_j БА \tilde{A} достижимо из состояния \bar{s}_i , если существует такая входная последовательность, которая переводит

БА из состояния \bar{s}_j в \bar{s}_j . Обозначим через $R(\bar{s}_j)$ множество тех состояний БА, которые достижимы из состояния \bar{s}_j .

Теорема 22.1. Для того чтобы БА \tilde{A} был БА БПИ- $\bar{s}(0)$, необходимо и достаточно, чтобы для любого состояния $\bar{s} \in R(\bar{s}(0))$ выполнялось условие $L(\bar{s}) = l$.

Доказательство. Очевидно, что неизвестный входной вектор $\bar{u}(0)$ БА \tilde{A} , стартующего в начальном состоянии $\bar{s}(0)$, может быть распознан тогда и только тогда, когда система (22.1) имеет единственное решение. Из [33] известно, что необходимым и достаточным условием для этого является равенство ранга этой системы числу l . Если входной вектор $\bar{u}(0)$ найден, то по формуле (21.1) легко вычислить следующее состояние $\bar{s}(1)$ этого БА. Зная его и выходной вектор $\bar{y}(1)$, можно найти вектор $\bar{u}(1)$ из системы $\sum_{i=1}^l G_i \bar{s}(1) + D\bar{u}_i(1) = \bar{y}(1) - C\bar{s}(1)$ поскольку $\bar{s}(1)$ достижимо из $\bar{s}(0)$ и по условию теоремы ранг матрицы последней системы равен l . Доказательство леммы получается отсюда по индукции.

Следуя [18], БА \tilde{A} назовем БА БПИ, если для него возможно однозначно определить неизвестную входную последовательность по известному начальному состоянию и наблюдаемой реакции независимо от начального состояния и подаваемой входной последовательности.

Теорема 22.2. Для того чтобы БА \tilde{A} был БА БПИ, необходимо и достаточно, чтобы для любого состояния $\bar{s} \in S_n L(\bar{s})$ был равен l .

Доказательство теоремы очевидно.

Отметим, что каждый БА БПИ является БА БПИ- $\bar{s}(0)$ для любого $\bar{s}(0)$, однако, если БА не является сильно связным, то, вообще говоря, не каждый БА БПИ- $\bar{s}(0)$ будет одновременно и БА БПИ.

Приведем одно достаточное условие принадлежности БА классу БА

БПИ.

Теорема 22.3. БА \tilde{A} является БА БПИ, если $\sum_{i=1}^l G_i = [0]$ и $rank D = l$.

В этом случае система (22.1) вырождается в систему

$$D\bar{u}(0) = \bar{y}(0) - C\bar{s}(0)$$

откуда и следует справедливость теоремы.

Опишем теперь процедуру нахождения множества $R(\bar{s}(0))$ всех состояний БА \tilde{A} , достижимых из состояния $\bar{s}(0)$. На первом шаге, полагая в формуле (21.1) $t = 0$ и подставляя в (21.1) все возможные входные векторы $\bar{u}(0)$, общее число которых равно p^l , получим множество всех состояний БА, достижимых из $\bar{s}(0)$ входными последовательностями длины 1, которое обозначим через $R_1(\bar{s}(0))$. Построим множество $R(\bar{s}(0)) := \{s_0\} \cup R_1(\bar{s}(0))$. На втором шаге положим в (21.1) $t = 1$ и для каждого $\bar{s} \in R_1(\bar{s}(0))$ и всех возможных входных векторов \bar{u} построим множество $R_2(\bar{s}(0)) = \bigcup R(\bar{s})$. Положим теперь множество $R(\bar{s}(0)) : R(\bar{s}(0)) \cup_{\bar{s} \in R_1(\bar{s}(0))} R(\bar{s})$. Если мощность только что построенного множества $R(\bar{s}(0))$ равна мощности множества $R(\bar{s}(0))$, полученного на предыдущем шаге, то процедура заканчивается - искомое множество найдено. В противном случае выполнение процедуры продолжается далее аналогичным образом до $t = p^n - 1$, поскольку любое состояние, достижимое из $\bar{s}(0)$, может быть достигнуто входной последовательностью, длина которой не превышает числа состояний БА.

Введем еще одну разновидность автоматов БПИ. БА \tilde{A} назовем БА существенно без потери информации (СБПИ), если, зная только входную последовательность, можно всегда однозначно определить неизвестную входную последовательность, породившую наблюдаемую

реакцию.

Сформулируем одно достаточное условие принадлежности БА классу БА СБПИ.

Теорема 22.4. БА \tilde{A} является БА СБПИ, если $\sum_{i=1}^l G_i = [0]$, $C = [0]$ и $D = I$.

При выполнении условий теоремы система (22.1) вырождается в систему $D\bar{u}(0) = \bar{y}(0)$, которая не зависит от начального состояния $\bar{s}(0)$ БА. Отсюда и следует справедливость утверждения.

Эксперименты с билинейными автоматами с запаздыванием

По аналогии с [17] приведем классификацию билинейных автоматов с запаздыванием.

Начнем с общих билинейных систем с распределенным запаздыванием.

Запаздывание по состоянию

Уравнение состояния имеет вид

$$\bar{s}(t+1) = A_0 \bar{s}(t) + A_1 \bar{s}(t-1) + \dots + A_h \bar{s}(t-h) + \left(\sum_{i=1}^l F_i^{(0)} u_i(t) \right) \bar{s}(t) + \dots + \left(\sum_{i=1}^l F_i^{(h)} u_i(t) \right) \bar{s}(t-h) + B\bar{u}(t)$$

Уравнение выхода имеет вид

$$\bar{y}(t) = C_0 \bar{s}(t) + C_1 \bar{s}(t-1) + \dots + C_h \bar{s}(t-h) + \left(\sum_{i=1}^l G_i^{(0)} u_i(t) \right) \bar{s}(t) + \dots + \left(\sum_{i=1}^l G_i^{(h)} u_i(t) \right) \bar{s}(t-h) + D\bar{u}(t)$$

где

$A_0, A_1, \dots, A_h, F_1^{(0)}, \dots, F_l^{(0)}, F_1^{(1)}, \dots, F_l^{(1)}, \dots, F_1^{(h)}, \dots, F_l^{(h)}$ - матрицы размерности $n \times n$;

B - матрица размерности $n \times l$;

$C_0, C_1, \dots, C_h, G_1^{(0)}, \dots, G_l^{(0)}, G_1^{(1)}, \dots, G_l^{(1)}, \dots, G_1^{(h)}, \dots, G_l^{(h)}$
 - матрицы размерности $m \times n$;

D - матрица размерности $m \times l$;

h - натуральное число (запаздывание).

Все элементы матриц и векторов берутся из основного поля $GF(p)$.

Запаздывание по управлению

Уравнение состояния имеет вид

$$\begin{aligned} \bar{s}(t+1) = & A\bar{s}(t) + \left(\sum_{i=1}^l F_i^{(0)} u_i(t) \right) \bar{s}(t) + \left(\sum_{i=1}^l F_i^{(1)} u_i(t-1) \right) \bar{s}(t) + \dots + \\ & + \left(\sum_{i=1}^l F_i^{(h)} u_i(t-h) \right) \bar{s}(t) + B_0 \bar{u}(t) + B_1 \bar{u}(t-1) + \dots + B_h \bar{u}(t-h) \end{aligned}$$

Уравнение выхода имеет вид

$$\begin{aligned} \bar{y}(t) = & C\bar{s}(t) + \left(\sum_{i=1}^l G_i^{(0)} u_i(t) \right) \bar{s}(t) + \left(\sum_{i=1}^l G_i^{(1)} u_i(t-1) \right) \bar{s}(t) + \dots + \\ & + \left(\sum_{i=1}^l G_i^{(h)} u_i(t-h) \right) \bar{s}(t) + D_0 \bar{u}(t) + D_1 \bar{u}(t-1) + \dots + D_h \bar{u}(t-h) \end{aligned}$$

Приведем теперь классификацию билинейных автоматов с запаздыванием.

БА с запаздыванием по состоянию

Уравнение состояния имеет вид

$$\bar{s}(t+1) = A_0 \bar{s}(t) + A_h \bar{s}(t-h) + \left(\sum_{i=1}^l F_i^{(0)} u_i(t) \right) \bar{s}(t) + \left(\sum_{i=1}^l F_i^{(h)} u_i(t) \right) \bar{s}(t-h) + B \bar{u}(t)$$

Уравнение выхода имеет вид

$$\bar{y}(t) = C_0 \bar{s}(t) + C_h \bar{s}(t-h) + \left(\sum_{i=1}^l G_i^{(0)} u_i(t) \right) \bar{s}(t) + \left(\sum_{i=1}^l G_i^{(h)} u_i(t) \right) \bar{s}(t-h) + D \bar{u}(t)$$

БА с запаздыванием по управлению

Уравнение состояния имеет вид

$$\bar{s}(t+1) = A \bar{s}(t) + \left(\sum_{i=1}^l F_i^{(0)} u_i(t) \right) \bar{s}(t) + \left(\sum_{i=1}^l F_i^{(h)} u_i(t-h) \right) \bar{s}(t) + B_0 \bar{u}(t) + B_h u(t-h)$$

Уравнение выхода имеет вид

$$\bar{y}(t) = C \bar{s}(t) + \left(\sum_{i=1}^l G_i^{(0)} u_i(t) \right) \bar{s}(t) + \left(\sum_{i=1}^l G_i^{(h)} u_i(t-h) \right) \bar{s}(t) + D_0 \bar{u}(t) + D_h \bar{u}(t-h)$$

Очевидно, что билинейные автоматы с запаздыванием являются частным случаем общих билинейных систем с распределенным запаздыванием, где соответствующие матрицы (запаздывание на $1, 2, \dots, h-1$ такт) являются нулевыми. Таким образом, все результаты, полученные для билинейных автоматов с распределенным запаздыванием, легко распространить на билинейные автоматы с запаздыванием. Поэтому далее рассматриваются только билинейные автоматы с распределенным запаздыванием.

Для однозначности определения состояний для $t > 0$ необходимо задать начальные данные: для билинейных автоматов с распределенным запаздыванием по состоянию необходимо задать состояния в моменты времени $t = 0, -1, -2, \dots, -h$; для билинейных автоматов с распределенным запаздыванием по управлению необходимо задать начальное состояние в момент времени $t = 0$, а также входные символы в моменты времени $t = 0, -1, -2, \dots, -h$.

Для изучения свойств билинейных автоматов с распределенным запаздыванием применим метод, использованный в [71] и уже примененный нами в лекции 3 (часть II) для линейных автоматов с

запаздыванием. Он состоит в том, что более сложная система с запаздыванием сводится к эквивалентной ей системе, но уже без запаздывания, однако с большим числом состояний. Так, для билинейного автомата с распределенным запаздыванием необходимо найти эквивалентный ему билинейный автомат. Далее, учитывая специфику характеристических матриц билинейных автоматов, можно получить необходимые и достаточные условия существования синхронизирующих, установочных и диагностических последовательностей для исходного билинейного автомата с распределенным запаздыванием.

Рассмотрим билинейный автомат с распределенным запаздыванием по состоянию. Его состояние в произвольный момент времени t определим как вектор $\bar{z}(t) = [\bar{s}(t)\bar{s}(t-1)\dots\bar{s}(t-h)]^T$. Тогда следующее состояние после подачи очередного входа $\bar{u}(t)$ будет

$$\bar{z}(t+1) = [\bar{s}(t+1)\bar{s}(t)\dots\bar{s}(t-h+1)]^T$$

$$\text{Начальное состояние } \bar{z}(0) = [\bar{s}(0)\bar{s}(-1)\dots\bar{s}(-h)]^T.$$

Тогда билинейный автомат с распределенным запаздыванием по состоянию можно описать следующим образом:

$$\bar{z}(t+1) = A * \bar{z}(t) + \left(\sum_{i=1}^l F_i * u_i(t) \right) \bar{z}(t) + B * \bar{u}(t) \quad (22.6)$$

$$\bar{y}(t) = C * \bar{z}(t) + \left(\sum_{i=1}^l G_i * u_i(t) \right) \bar{z}(t) + D * \bar{u}(t) \quad (22.7)$$

Здесь матрицы со звездочкой имеют следующую блочную структуру:

$$A* = \begin{bmatrix} A_0 & A_1 & \dots & A_h \\ E_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & E_n & 0_n \end{bmatrix}$$

где в роли блоков выступают матрицы размерности $n \times n$, а число блочных строк и столбцов равно $h+1$;

$$F_i^* = \begin{bmatrix} F_i^{(0)} & F_i^{(1)} & \dots & F_i^{(h)} \\ 0_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & 0_n & \dots & 0_n \end{bmatrix}$$

где в роли блоков выступают матрицы размерности $n \times n$, а число блочных строк и столбцов равно $h + 1$;

$$B^* = \begin{bmatrix} B \\ [0] \\ \dots \\ [0] \end{bmatrix}$$

где в роли блоков выступают матрицы размерности $n \times l$, а число блочных строк равно $h + 1$;

$$C^* = [C_0 C_1 \dots C_h]$$

где в роли блоков выступают матрицы размерности $m \times n$, а число блочных столбцов равно $h + 1$;

$$G_i^* = [G_i^{(0)} G_i^{(1)} \dots G_i^{(h)}]$$

где в роли блоков выступают матрицы размерности $m \times n$, а число блочных столбцов равно $h + 1$;

$$D^* = D$$

В справедливости (22.6) и (22.7) можно убедиться, выполнив непосредственную подстановку. Поскольку соответствующие выкладки являются громоздкими, но не представляют принципиальной сложности, здесь они не приводятся.

Рассмотрим билинейный автомат с распределенным запаздыванием по управлению.

Пусть входным вектором в момент времени t является

$\tilde{u}(t) = [\bar{u}(t)\bar{u}(t-1)\dots\bar{u}(t-h)]^T$. Тогда размерность $\tilde{u}(t)$ равна $(h+1)l \times 1$.

Начальный входной вектор $\tilde{u}(0) = [\bar{u}(0)\bar{u}(-1)\dots\bar{u}(-h)]^T$.

Тогда билинейный автомат с распределенным запаздыванием по управлению можно описать следующим образом:

$$\bar{s}(t+1) = A * \bar{s}(t) + \left(\sum_{i=1}^{(h+1)l} F_i * \tilde{u}_i(t) \right) \bar{s}(t) + B * \tilde{u}(t) \quad (22.8)$$

$$\bar{y}(t) = C * \bar{s}(t) + \left(\sum_{i=1}^{(h+1)l} G_i * \tilde{u}_i(t) \right) \quad (22.9)$$

Здесь

$A* = A; F_1*, F_2*, \dots, F_{(h+1)l}* \quad \text{равны} \quad \text{соответственно}$
 $F_1^{(0)}, \dots, F_l^{(0)}, \dots, F_1^{(h)}, \dots, F_l^{(h)}; B* = [B_0 B_1 \dots B_h]$ - вектор
 блочной структуры, где в роли блоков выступают матрицы размерности
 $n \times l$, а число блочных столбцов равно
 $h+1; C* = C; G_1*, G_2*, \dots, G_{(h+1)l}* \quad \text{равны} \quad \text{соответственно}$
 $G_1^{(0)}, \dots, G_l^{(0)}, \dots, G_1^{(h)}, \dots, G_l^{(h)}; D* = [D_0 D_1 \dots D_h]$ - вектор
 блочной структуры, где в роли блоков выступают матрицы размерности
 $m \times l$, а число блочных столбцов равно $h+1$.

В справедливости (22.8) и (22.9) можно убедиться, выполнив непосредственную подстановку.

В результате предложенных выше преобразований исходный билинейный автомат с распределенным запаздыванием по состоянию (по управлению) сведен к билинейному автомату без запаздывания, вектор состояния (входной вектор) которого имеет большую размерность, чем исходный автомат. Последнее означает, что с точки зрения функционального поведения оба эти автомата являются эквивалентными.

Приведенные ниже теоремы являются аналогами доказанных выше

теорем 12.1, 12.4, 12.5 и т. д.

Теорема 22.5. Для того чтобы входная последовательность $\bar{u}(0) \dots \bar{u}(t)$ была синхронизирующей для билинейного автомата \tilde{A} с распределенным запаздыванием по состоянию, необходимо и достаточно, чтобы выполнялось равенство

$$\prod_{\nu=0}^t \begin{bmatrix} A_0 + \sum_{i=1}^l F_i^{(0)} u_i(t-\nu) & A_1 + \sum_{i=1}^l F_i^{(1)} u_i(t-\nu) & \dots & A_h + \sum_{i=1}^l F_i^{(h)} u_i(t-\nu) \\ E_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & E_n & 0_n \end{bmatrix} = [0]$$

Теорема 22.6. Для того чтобы входная последовательность $\bar{u}(0) \dots \bar{u}(t)$ была синхронизирующей для билинейного автомата \tilde{A} с распределенным запаздыванием по управлению, необходимо и достаточно, чтобы выполнялось равенство

$$\prod_{\nu=0}^t [A + \sum_{k=0}^h \sum_{i=1}^l F_i^{(k)} u_i(t-\nu)] = [0]$$

Теорема 22.7. Для того чтобы входная последовательность $\bar{u}(0) \dots \bar{u}(t)$ была установочной для билинейного автомата \tilde{A} с распределенным запаздыванием по состоянию, необходимо и достаточно, чтобы для каждого ненулевого состояния $\bar{s} \in S_{(h+1)n}$ выполнялось по крайней мере одно из условий:

$$1. \quad \bigwedge_{d=0}^t \prod_{\nu=0}^{d-1} \begin{bmatrix} (C_0 + \sum_{i=1}^l G_i^{(0)} u_i(d) \dots C_h + \sum_{i=1}^l G_i^{(h)} u_i(d)) \times \\ \left[A_0 + \sum_{i=1}^l F_i^{(0)} u_i(d-\nu-1) \quad \dots \quad A_h + \sum_{i=1}^l F_i^{(h)} u_i(d-\nu-1) \right] \\ E_n \quad \quad \quad 0_n \quad \dots \quad 0_n \\ \dots \quad \quad \quad \dots \quad \dots \quad \dots \\ 0_n \quad \quad \quad \dots \quad E_n \quad \quad \quad 0_n \end{bmatrix} \bar{s} \neq 0$$

$$2. \quad \begin{matrix} \prod_{\nu=0}^t \\ \bar{s} = [0] \end{matrix} \left[\begin{array}{cccc} A_0 + \sum_{i=1}^l F_i^{(0)} u_i(t-\nu) & A_1 + \sum_{i=1}^l F_i^{(0)} u_i(t-\nu) & \dots & A_h + \sum_{i=1}^l F_i^{(h)} u_i(t-\nu) \\ E_n & 0_n & \dots & \dots 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots E_n & \dots & 0_n \end{array} \right] =$$

Здесь знак $\bigwedge_{d=0}^t$ означает дизъюнкцию выражений, стоящих за этим знаком и получаемых при изменении индекса от 0 до t .

Если при вычислениях индекс у u_i меньше нуля при некоторых d и ν (условие 1), то выражение $\sum_{i=1}^l F_i^{(0)} u_i(d-\nu-1)$ положим равным единичной матрице.

Теорема 22.8. Для того чтобы входная последовательность $\bar{u}(0) \dots \bar{u}(t)$ была установочной для билинейного автомата \tilde{A} с распределенным запаздыванием по управлению, необходимо и достаточно, чтобы для каждого ненулевого состояния $\bar{s} \in S_n$ выполнялось по крайней мере одно из условий:

1. $\bigwedge_{d=0}^t \left[\left[C + \sum_{k=0}^h \sum_{i=0}^l G_i^{(k)} u_i(d) \right] \prod_{\nu=0}^{d-1} \left[A + \sum_{k=0}^h \sum_{i=0}^l F_i^{(k)} u_i(d-\nu-1) \right] \bar{s} \neq [0] \right]$
2. $\prod_{\nu=0}^t \left[A + \sum_{k=0}^h \sum_{i=0}^l F_i^{(k)} u_i(t-\nu) \right] \bar{s} = [0]$.

Здесь знак $\bigwedge_{d=0}^t$ означает дизъюнкцию выражений, стоящих за этим знаком и получаемых при изменении индекса от 0 до t .

Если при вычислениях индекс у u_i меньше нуля при некоторых d и ν (условие 1), то выражение $\sum_{i=1}^l F_i^{(k)} u_i(d-\nu-1)$ положим равным единичной матрице.

Теорема 22.9. Для того чтобы входная последовательность $\bar{u}(0) \dots \bar{u}(t)$ была диагностической для билинейного автомата \tilde{A} с распределенным запаздыванием по состоянию, необходимо и достаточно, чтобы

$$\text{rank}_{(h+1)n} \begin{bmatrix} \left[C_0 + \sum_{i=1}^l l G_i^{(0)} \dots C_h + \sum_{i=1}^l G_i^{(h)} u_i(0) \right] \\ \left[C_0 + \sum_{i=1}^l G_i^{(0)} u_i(1) \dots C_h + \sum_{i=1}^l G_i^{(h)} u_i(1) \right] (A * + \sum_{i=1}^l F_i * u_i(0)) \\ \left[C_0 + \sum_{i=1}^l G_i^{(0)} u_i(t) \dots C_h + \sum_{i=1}^l G_i^{(h)} u_i(t) \right] \prod_{\nu=0}^t \left[A * + \sum_{i=1}^l F_i * u_i(t - \nu - 1) \right] \end{bmatrix} =$$

Здесь матрицы со звездочкой имеют следующую блочную структуру:

$$A * = \begin{bmatrix} A_0 & A_1 & \dots & A_h \\ E_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & \dots & E_n & 0_n \end{bmatrix}$$

где в роли блоков выступают матрицы размерности $n \times n$, а число блочных строк и столбцов равно $h + 1$;

$$F_i * = \begin{bmatrix} F_i^{(0)} & F_i^{(1)} & \dots & F_i^{(h)} \\ 0_n & 0_n & \dots & 0_n \\ \dots & \dots & \dots & \dots \\ 0_n & 0_n & \dots & 0_n \end{bmatrix}$$

где в роли блоков выступают матрицы размерности $n \times n$, а число блочных строк и столбцов равно $h + 1$.

Теорема 22.10. Для того чтобы входная последовательность $\bar{u}(0) \dots \bar{u}(t)$ была диагностической для билинейного автомата \tilde{A} с распределенным запаздыванием по управлению, необходимо и достаточно, чтобы

$$\text{rank}_n \begin{bmatrix} \left[C + \sum_{k=0}^h \sum_{i=0}^l G_i^{(k)} u_i(0) \right] \\ \left(C + \sum_{k=0}^h \sum_{i=0}^l G_i^{(k)} u_i(1) \right) \left(A + \sum_{k=0}^h \sum_{i=0}^l F_i^{(k)} u_i(0) \right) \\ \left(C + \sum_{k=0}^h \sum_{i=0}^l G_i^{(k)} u_i(t) \right) \prod_{\nu=0}^t \left[A + \sum_{k=0}^h \sum_{i=0}^l F_i^{(k)} u_i(t - \nu - 1) \right] \end{bmatrix} =$$

Вопросы и упражнения

1. Дайте определение билинейного автомата без потери информации из некоторого состояния.
2. Дайте определение билинейного автомата без потери информации.
3. Сформулируйте условие принадлежности автомата классу билинейных автоматов БПИ.
4. Приведите определения билинейных автоматов с распределенным запаздыванием и просто с запаздыванием по состоянию и управлению.
5. Покажите, что билинейный автомат с распределенным запаздыванием всегда можно преобразовать в билинейный автомат без запаздывания.
6. Приведите критерии существования синхронизирующей, установочной и диагностической последовательностей для билинейных автоматов с распределенным запаздыванием.

Список литературы

1. Агафонов В. Н. , Спецификация программ: понятийные средства и их организации, Новосибирск: Наука, Сиб. отд-е, 1990. — 224 с
2. Агибалов Г. П., Юфат Я. Г., О простых экспериментах для линейных инициальных автоматов, Автоматика и вычислительная техника. № 2. 1972. С. 17–19
3. Аксенова Г. П., Согомонян Е. С. , Синтез схем встроенного контроля для автоматов с памятью , Автоматика и телемеханика. № 9. 1971. С. 170–179
4. Алефельд Г., Херцбергер Ю., Введение в интервальные вычисления, М.: Мир, 1987
5. Батищев Д. И., Генетические алгоритмы решения экстремальных задач, Воронеж. гос. техн. ун-т; Нижегородский гос. ун-т. — Воронеж, 1995
6. Барашко А. С., Скобцов Ю. А., Сперанский Д. В. , Моделирование и тестирование дискретных устройств, Киев: Наукова думка, 1992. — 288 с
7. Березин И. С., Жидков Н.П. , Методы вычислений В 2-х т, М.: Гос. изд-во физ.-мат. лит., 1960. Т. 2
8. Богомолов А. М., Барашко А. С., Грунский И. С. , Эксперименты с автоматами, Киев: Наукова думка, 1973. — 144 с
9. Богомолов А. М., Грунский И. С., Сперанский Д.В., Контроль и преобразования дискретных автоматов, Киев: Наукова думка, 1975. — 176 с
10. Богомолов А. М., Твердохлебов В. А., Диагностика сложных систем, Киев: Наукова думка, 1974. — 128 с
11. Богомолов А.С., Сперанский Д.В., Об одном оптимальном синхронизирующем эксперименте с линейными автоматами , Известия РАН. Теория и системы управления. № 3. 2002. С. 64–70
12. Богомолов А. С., Сперанский Д. В. , Синхронизирующие эксперименты с интервальными линейными системами , Автоматика и телемеханика. № 6. 2002. С. 166–178
13. Богомолов А. С., Сперанский Д. В. , Оптимальные синхронизирующие эксперименты с автоматами , Автоматика и телемеханика. № 10. 2001. С. 203–208
14. Бол्यानский В. Г. , Оптимальное управление дискретными системами , М.: Наука, 1973
15. Венцель Е. С. , Элементы динамического программирования , М.: Наука, 1964
16. Визинг В. Г., Гольдберг М. К. , О длине обхода сильно связанного графа , Кибернетика. № 1. 1969. С. 79–82
17. Габасов Р., Кириллова Ф. М., Краютко В. В. , Теория управляемости линейных дискретных систем, Определяющее уравнение Дифференциальные уравнения. Т. 8, № 5. 1972. С. 767–773
18. Гилл А. , Введение в теорию конечных автоматов , М.: Наука, 1966. — 272 с
19. Гилл А. , Линейные последовательностные машины , М.: Наука, 1974
20. Глушков В. М. , Абстрактная теория автоматов , Успехи матем. наук. Т. 16, вып. 5 (101). 1961. С. 3–62
21. Глушков В. М. , Синтез цифровых автоматов , М.: Физматгиз, 1962
22. Грунский И. С. , Анализ поведения конечных автоматов , Луганск: Изд-во Луганск. гос. пед. ун-та, 2003. — 318 с
23. Грунский И. С., Рубанович Ю. А. , Об одном алгоритме построения выходящих контрольных точек , Автоматика и телемеханика. № 11. 1967. С. 145–150

24. Гэри М., Джоисон Д., Вычислительные машины и труднорешаемые задачи, М.: Мир, 1982. — 416 с
25. Зыков А. А., Теория конечных графов, 1, Новосибирск; Наука, 1969
26. Калмыков С. А., Шокии Ю. И., Юлдашев З. Х., Методы интервального анализа, Новосибирск: Наука, Сиб. Отд-е, 1986
27. Колесов Н. В., Построение проверяющего теста для линейного конечного автомата, Автоматика и телемеханика. № 2. 1982. С. 61–66
28. Корбут А. А., Финкельштейн Ю. Ю., Дискретное программирование, М.: Наука, 1969
29. Корчинский А. П., Тоценко В. Г., Синтез автоматов оперативного контроля дискретных устройств с конечной памятью, Автоматика и телемеханика. № 10. 1973. С. 184–193
30. Кудрявцев В. Б., Алешин С. В., Подколзин А. С., Введение в теорию автоматов, М.: Наука, 1985. — 320 с
31. Куприянова Л. В., Сперанский Д. В., Самойлов В. Г., Интервальная арифметика над полем $GF(p)$, Вычислительные технологии. Т. 7, № 6. 2002. С. 54–64
32. Курейчик В. М., Генетические алгоритмы и их применение, Таганрог: Изд-во ТРТУ, 2002
33. Курош А. Г., Курс высшей алгебры, М.: Наука, Главная ред. физ.-мат. лит., 1976
34. Майника Э., Алгоритмы оптимизации на графах и сетях, М.: Мир, 1981. — 454 с
35. Мишу М., Математическое программирование, М.: Наука, 1980
36. Мур Э., Умозрительные эксперименты с последовательностными машинами, М.: Л., 1956. С. 179–210
37. Основы технической диагностики, Энергия, 1976. — 464 с
38. Первозванский А. А., Курс теории автоматического управления, М.: Наука, 1986
39. Петренко А. Ф., Эксперименты над протокольными объектами, Автоматика и вычислительная техника. № 1. 1987. С. 16–21
40. Поитрагии Л. С., Математическая теория оптимальных процессов, М.: Наука, 1976
41. Понтрягин Л. С., Болтянский В. Г. и др., Математическая теория оптимальных процессов, М.: Физматгиз, 1961
42. Пропой А. И., Элементы теории оптимальных линейных процессов, М.: Наука, 1973
43. Сперанский Д. В., Автоматы существенно без потери информации, Автоматика и телемеханика. № 10. 1971. С. 93–99
44. Сперанский Д. В., О связи между устойчивостью и синхронизируемостью линейных автоматов над различными полями, Автоматика и вычислительная техника. № 1. 2000. С. 63–75
45. Сперанский Д. В., О тестировании линейных автоматов, Автоматика и телемеханика. № 5. 2000. С. 157–165
46. Сперанский Д. В., Об автоматах существенно без потери информации конечного порядка, Автоматика и телемеханика. № 6. 1972. С. 86–92
47. Сперанский Д. В., Обобщенная синхронизация линейных последовательностных машин, Кибернетика и системный анализ. № 3. 1998. С. 17–25
48. Сперанский Д. В., Обобщенные автоматы без потери информации конечного порядка, II, Кибернетика и системный анализ. № 4. 1994. С. 174–178

49. Сперанский Д. В. , Обобщенные автоматы без потери информации, I, Кибернетика и системный анализ. № 3. 1994. С. 63–69
50. Сперанский Д. В. , Обобщенные линейные автоматы без потерн информации , Известия РАН. Теория и системы управления. № 1. 1998. С. 64–71
51. Сперанский Д. В. , Распознавание состояний нестационарных линейных автоматов , Известия РАН. Теория и системы управления. № 6. 2000. С. 82–89
52. Сперанский Д. В., Связь между устойчивостью и синхронизируемостью нестационарных линейных автоматов над различными полями , Электронное моделирование. Т. 23, № 5. 2001. С. 22–32
53. Сперанский Д. В. , Синтез тестов с минимальным числом перепадов сигналов для линейных автоматов , Автоматика и вычислительная техника. № 4. 2002. С. 70–78
54. Сперанский Д. В. , Синхронизация линейных последовательностных машин , Автоматика и телемеханика. № 5. 1996. С. 141–149
55. Сперанский Д. В. , Установочные и диагностические последовательности для линейных автоматов , Автоматика и телемеханика. № 5. 1997. С. 133–141
56. Сперанский Д. В., Огнева М. В. , Обобщенные установочные и диагностические эксперименты для линейных автоматов , Автоматика и вычислительная техника. № 3. 2001. С. 3–13
57. Сперанский Д. В., Самойлов В. Г., О диагностической задаче для линейных автоматов в интервальной постановке , Информац.-керуючі системи на залізничному транспорті. № 4. 2003. С. 117–119
58. Сперанский Д. В., Сперанский И. Д. , Об одной задаче для сети из линейных автоматов без потери информации , Автоматика и телемеханика. № 1. 1999. С. 140–147
59. Сперанский Д. В., Сперанский И. Д. , Разрешающие возможности диагностических экспериментов с линейными автоматами , Кибернетика и системный анализ. № 3. 2000. С. 62–65
60. Сперанский Д. В., Сперанский И. Д. , Эксперименты с билинейными системами , Автоматика и телемеханика. № 6. 2000. С. 176–189
61. Сперанский Д. В., Сперанский И. Д. , Эксперименты с линейными дискретными системами, Электронное моделирование. № 4. 1999. С. 64–73
62. Спивак М. А. , Введение в абстрактную теорию автоматов , Саратов: Изд-во Саратов. ун-та, 1970
63. Спивак М. А. , Обобщенные задачи диагноза и установки для конечных автоматов , Известия АН СССР, Технич. киб-ка. № 3. 1969. С. 82–87
64. Стрейц В. , Методы пространства состояний в теории линейных дискретных систем управления , М.: Наука, 1985
65. Твердохлебов В.А. , Логические эксперименты с автоматами , Саратов: Изд-во Саратов. ун-та, 1988. — 184 с
66. Фараджев Р. Г. , Линейные последовательностные машины , М.: Сов. радио, 1965
67. Хиббард Т. Н. , Точные верхние границы длин минимальных экспериментов, определяющих заключительное состояние, для двух классов последовательных машин , Кибернетический сборник. Вып. 2. 1966. С. 7–23
68. Хорн Р., Джонсон Ч. , Матричный анализ , М.: Мир, 1989
69. Яблонский С. В. , Введение в дискретную математику , М.: Наука, 1997. — 384 с
70. Яблонский С. В. , Надежность и контроль управляющих систем , Математические вопросы кибернетики. Вып. 1. 1988. С. 5–25

71. Янушевский Р. Т., Управление объектами с запаздыванием, М.: Наука, 1978
72. Bhattacharyya A., Checking experiments in sequential machines, New York: J. Wiley and Sons, 1989. — 155 p
73. Breuer M., Friedman A., Diagnosis and reliable design of digital systems, Computer Science Press. — 300 p
74. Even S., On information lossless automata of finite order, IEEE Trans. Elect. Comput. Vol. C-14, № 4. 1965. P. 561–569
75. Ginzburg A., Algebraic Theory of Automata, New York — London, Academic Press, 1968
76. Huffman D. A., Canonical forms for information-lossless finite-state logical machines, IRE Trans. Circuit Theory. Special supplement. Vol. CT-6. 1959. P. 41–59
77. Henney F. C., Finite-state models for logical machines, New York, 1968
78. Kohavi Z., Switching and finite automata theory, New York: Mc Graw Hill, 1970. — 522 p
79. Kohavi Z., Lavalle P., Design of sequential machines with fault detecting capabilities, IEEE Trans. Electron Comput. V. 16. № 4. 1967. P. 171–184
80. Lorens C. S., Invertible Boolean function, IEEE Trans. Elect. Comput. EC-13, V. 5. 1964. P. 529–541
81. Shary S. P., Algebraic approach to the linear static identification, tolerance and control problem, or One more application of Caucher arithmetic, Reliable Computing. Vol. 2, № 1. 1996. P. 3–33
82. Yannakakis M., Lee D., Principles and methods of testing finite-state machines/ – A survey /, Proc. IEEE. V. 84. № 8. 1996. P. 1090–1123

Содержание

Титульная страница	2
Выходные данные	3
Лекция 1. Эксперименты с автоматами, имеющими взвешенный входной алфавит	4
Лекция 2. Синтез экспериментов методами динамического программирования	19
Лекция 3. Обобщенные автоматы без потери информации	33
Лекция 4. Обобщенные автоматы без потери информации конечного порядка	47
Лекция 5. Преобразования автоматов в автоматы без потери информации	59
Лекция 6. Эксперименты по контролю функции выходов инициального автомата	82
Лекция 7. Контроль функции выходов неинициального автомата с использованием простого безусловного эксперимента	103
Лекция 8. Контроль функции выходов инициального автомата с использованием кратного безусловного эксперимента	113
Лекция 9. Преобразование автомата для упрощения функционального контроля	127
Лекция 10. Синхронизирующие эксперименты с линейными автоматами	136
Лекция 11. Установочные и диагностические эксперименты со стационарными и нестационарными линейными автоматами	152
Лекция 12. Эксперименты в пространстве обобщенных состояний и с линейными автоматами с запаздыванием	171

Лекция 13. Синхронизация и устойчивость дискретных линейных систем	187
Лекция 14. Эксперименты по распознаванию неисправностей линейных автоматов	207
Лекция 15. Линейные автоматы существенно без потери информации	224
Лекция 16. Обобщенные линейные автоматы без потери информации	236
Лекция 17. Минимизация времени восстановления неизвестных входных сигналов в сети из автоматов без потери информации	251
Лекция 18. Оптимальные эксперименты с линейными автоматами	265
Лекция 19. Интервальная арифметика над конечным полем и ее приложения к теории экспериментов с автоматами	283
Лекция 20. Диагностическая задача в интервальной постановке	299
Лекция 21. Эксперименты с билинейными автоматами по распознаванию состояний	320
Лекция 22. Разновидности экспериментов с билинейными автоматами	337
Список литературы	350