

Ep. 6.7

674

ISSN 0136-3549

3134-3823

TALLINNA POLÜTEHNILISE INSTITUUDI

TOIMETISED

ТРУДЫ ТАЛЛИНСКОГО
ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

TRANSACTIONS OF TALLINN
TECHNICAL UNIVERSITY

МАШИННОЕ ПРОЕКТИРОВАНИЕ
ЭЛЕКТРОННЫХ УСТРОЙСТВ И
СИСТЕМ



TALLINN 1988

674

ALUSTATUD 1937

**TALLINNA POLÜTEHNILISE
INSTITUUDI TOIMETISED**

**ТРУДЫ ТАЛЛИНСКОГО
ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА**

**TRANSACTIONS OF TALLINN
TECHNICAL UNIVERSITY**

 **УДК 62.507+681.3+621.3**

**МАШИННОЕ ПРОЕКТИРОВАНИЕ
ЭЛЕКТРОННЫХ УСТРОЙСТВ И
СИСТЕМ**

Электротехника и автоматика XXXУ

TALLINN 1988



ТАЛЛИНСКИЙ ПОЛИТЕХНИЧЕСКИЙ ИНСТИТУТ

Труды ТПИ № 874

МАШИННОЕ ПРОЕКТИРОВАНИЕ ЭЛЕКТРОННЫХ УСТРОЙСТВ
И СИСТЕМ

Электротехника и автоматика XXXУ

На русском языке

Ответственный редактор Э. Рюстерн

Технический редактор Е. Зорина

Сборник утвержден коллегией Трудов ТПИ 29.06.88

Подписано к печати 25.10.88

МВ-07277

Формат 60x90/16

Печ. л. 11,0 + 0,5 приложение

Уч.-изд. л. 8,78

Тираж 300. Зак. № 652

Цена 1 руб. 80 коп.

Таллинский политехнический институт

200108 Таллин, Эхитаяте тее, 5

Ротапринт ТПИ

200006 Таллин, ул. Коскла, 2/9

© Таллинский политехнический институт, 1988

А. Кеэваллик, М. Круус, К. Салум

О ДЕКОМПОЗИЦИИ КОНЕЧНЫХ АВТОМАТОВ НА БАЗЕ
СДВИГОВЫХ РЕГИСТРОВ

Введение. В настоящее время снова повышается интерес к проблемам, связанным с реализацией цифровых управляющих автоматов (УА) на базе сдвиговых регистров (СР). Такая реализация существенно облегчает задачу контроля УА, давая возможность построить для него короткую контрольную последовательность [1]. Это объясняется тем, что если УА реализован на базе СР длиной K и выход СР непосредственно наблюдаемый по выходной функции автомата, то произвольная последовательность входных символов длиной K является диагностической [2] для этого автомата.

Задачу реализации УА на базе СР можно рассматривать как некоторую задачу декомпозиции исходного УА [3]. Такая декомпозиция строится по системе двухблочных разбиений

$P = \{\pi_i | 1 \leq i \leq n\}$ на множестве внутренних состояний автомата, причем каждое разбиение π_i соответствует одному разряду СР. В [3] были приведены необходимые и достаточные условия существования такой декомпозиции. В данной работе предложены алгоритмы реализации УА на базе СР с применением метода расщепления состояний исходного УА.

Основные понятия. Поскольку проблемы реализации выходной функции автомата можно в данном контексте считать второстепенными, то в качестве автоматной модели будем рассматривать автомат без выхода $A = (I, S, \delta)$, где I - входной алфавит, S - множество внутренних состояний и $\delta: S \times I \rightarrow S$ - функция переходов автомата.

Определение 1. Пару разбиений (π, π') называем симметричной, если и только если можно образовать взаимно одно-

значное соответствие $\varphi: \pi \leftrightarrow \pi'$ такое, что $\varphi(B) = \pi'[\delta(s, a)]$, где $B \in \pi, s \in B, a \in I$.

Определение 2. К-цепью называем цепь из симметричных пар разбиений (π_{i-1}, π_i) , $2 \leq i \leq k$.

Существование К-цепи является необходимым и достаточным условием для того, чтобы в реализации автомата использовать СР длиной К [3]. Если при этом система разбиений $\{\pi_1, \pi_2, \dots, \pi_k\}$ полная, то автомат реализуется на одном СР. (Систему разбиений называем полной, если $\pi_1 \cdot \pi_2 \cdot \dots \cdot \pi_k = O_\pi$, где O_π - разбиение, в каждом блоке которого один элемент). Пусть $\{(\pi_j, \pi'_j)\}$, $1 \leq j \leq l$ - множество всех возможных симметричных пар разбиений заданного автомата. Через $\pi^C = \Pi \pi_j, 1 \leq j \leq l$ и $\pi^T = \Pi \pi'_j, 1 \leq j \leq l$ обозначим наименьшие разбиения, образующие симметричную пару.

В [3] приведены алгоритмы нахождения разбиений π^C и π^T , а на основе их всех К-цепей для данного автомата. Заметим, что найденные К-цепи не состоят из двухблочных разбиений. Каждая такая К-цепь описывает структуру сети, в которой компонентные автоматы $A_1, \dots, A_k$ имеют больше, чем 2 состояния, т.е. для реализации $A_1, \dots, A_k$ необходимо использовать $\lceil \log_2 \#(\pi_i) \rceil$ СР длиной К, где через $\#(\pi_i)$ обозначено количество блоков в разбиении π_i , а $\lceil \omega \rceil$ - наименьшее целое число, которое больше, чем ω .

Реализация автомата на базе СР. В [3] не показано, как выбрать из множества всех К-цепей минимальное необходимое количество их для реализации автомата на базе СР (в дальнейшем просто СР-реализации). Предложим алгоритм выбора, позволяющий приблизительно минимизировать суммарную длину используемых СР.

Предположим, что некоторая К-цепь имеет t блоков в каждом разбиении. Очевидно, что на основе такой К-цепи можно построить (путем соответствующего укрупнения блоков разбиений) $2^{t-1} - 1$ различных К-цепей, которые состоят из двухблочных разбиений. В дальнейшем будем такую К-цепь называть двухблочной. Каждая двухблочная К-цепь в реализации автомата соответствует одному СР длиной К. Если количество блоков в разбиениях 2-, 3- и т.д. К-цепи обозначить через $t_2, t_3, \dots, t_k$ соответственно, то можем образовать T различных двухблочных К-цепей, где

$$T = \sum_{i=2}^K (2^{i-1} - 1). \quad (I)$$

На основе каждой двухблочной К-цепи можно построить некоторый подавтомат реализуемого конечного автомата. Для сравнения различных К-цепей целесообразно ввести количественную оценочную функцию. Для этого характеризуем каждую двухблочную К-цепь произведением ее разбиений:

$$\pi_t = \prod_{i=1}^K \pi_{t,i}, \quad 1 \leq t \leq T. \quad (2)$$

Очевидно, что $\pi_t \geq \pi^r \cdot \pi^c$. Различные разбиения $\pi_t, 1 \leq t \leq T$ могут быть между собой несравнимы. Чтобы оценить соответствующие К-цепи введем оценочную функцию [4]:

$$R(\pi_t) = M \cdot E(\pi_t) + (M - (\#(\pi_t))), \quad (3)$$

где M - количество состояний в разбиении π_t ;

$E(\pi_t)$ - количество состояний в максимальном блоке π_t ;

$\#(\pi_t)$ - количество блоков в разбиении π_t .

Если $\pi_t = 0_\pi$, то $R(\pi_t) = M$, а если $\pi_t = 1_\pi$ (разбиение, в котором один блок), то $R(\pi_t) = M^2 + M - 1$. Значит К-цепь реализует "большую часть" автомата, если значение функции $R(\pi_t)$ минимальное.

Задача нахождения СР-реализации сводится к задаче нахождения минимального количества двухблочных К-цепей так, чтобы $\prod \pi_t = \pi^c \cdot \pi^r$.

Опишем алгоритм нахождения полной системы разбиений для реализации автомата на базе СР.

Алгоритм I

Шаг 1. Найти все К-цепи [3].

Шаг 2. Образовать T (I) двухблочных К-цепей.

Шаг 3. Для каждой К-цепи вычислить произведение ее разбиений π_t и оценочную функцию $R(\pi_t)$ (3).

Шаг 4. $\tau = 1_\pi$.

Шаг 5. Найти $\min_t \{R(\tau \cdot \pi_t)\}$. К-цепь, соответствующую произведению π_t , включить в систему разбиений. Ес-

ли существуют К-цепи, для которых значение оценочной функции одинаково, выбрать цепь меньшей длины.

Шаг 6. $\tau = \tau \cdot \pi_+$.

Шаг 7. Если $\tau = \pi^r \cdot \pi^c$, то идти к шагу 10.

Шаг 8. Если $\#(\tau) = 2$, то идти к шагу 11.

Шаг 9. Идти к шагу 5.

Шаг 10. Если $\pi^c \cdot \pi^r = 0_\pi$, то идти к шагу 12.

Шаг 11. Дополнить систему разбиений до полного.

Шаг 12. Конец.

Наличие условия $\#(\tau) = 2$ на шаге 8 объясняется следующим образом.

Если $\#(\tau) = 2$, то систему разбиений можно дополнить до полной одним двухблочным разбиением. Тем самым мы включаем в СР — реализацию тривиальный СР длиной 1. Очевидно, что это решение является наилучшим в этой ситуации. На шаге 11 в систему разбиений добавляется одно разбиение с учетом полноты системы.

Пример. Для автомата А (рис. 1) существуют 2-, 3-, 4-, 5- и 6-цепи. Применяя предложенный алгоритм, выберем по минимальной оценочной функции первую 2-цепь, состоящую из разбиений π_1 и π_2 :

$$\pi_1 = \{ \overline{4, 7, 8, 9, 11, 12, 14, 17, 20, 21, 22, 23}; \\ \overline{1, 2, 3, 5, 6, 10, 13, 15, 16, 18, 19} \};$$

$$\pi_2 = \{ \overline{6, 8, 9, 10, 12, 13, 15, 18, 19, 21, 22, 23}; \\ \overline{1, 2, 3, 4, 5, 7, 11, 14, 16, 17, 20} \}.$$

На втором проходе алгоритма выбирается также 2-цепь из разбиений π_3 и π_4 :

$$\pi_3 = \{ \overline{3, 4, 6, 8, 15, 16, 17, 18, 19, 21, 22, 23}; \\ \overline{1, 2, 5, 7, 9, 10, 11, 12, 13, 14, 20} \};$$

$$\pi_4 = \{ \overline{1, 5, 6, 7, 9, 16, 17, 18, 20, 22, 23}; \\ \overline{2, 3, 4, 8, 10, 11, 12, 13, 14, 15, 19, 21} \};$$

Поскольку в максимальном блоке произведения $\pi_1 \cdot \pi_2 \cdot \pi_3 \cdot \pi_4$ 2-состояния, то система доводится до ортогонального выбором двухблочного разбиения, например, такого:

$$\pi_5 = \{ \overline{1, 2, 3, 4, 6, 7, 8, 9, 10, 11, 12, 15, 22} ; \overline{5, 13, 14, 16, 17, 18, 19, 20, 21, 23} \}.$$

Следовательно, данный автомат реализуется на двух СР длиной 2 и на одном дополнительном триггере. Заметим, что минимальное необходимое количество триггеров для реализации этого автомата также 5.

Метод расщепления состояний исходного автомата. Полная реализация произвольного конечного автомата на базе нетривиальных СР невозможна. В такой ситуации может быть использован подход, который заключается в том, что исходный автомат заменяется на эквивалентный ему автомат (путем расщепления его состояний), который лучше реализуется на базе СР.

Пусть даны разбиение π и покрытие φ на множестве состояний S исходного автомата.

Определение 3. Пару (π, φ) будем называть симметричной смешанной парой, если и только если между блоками π и φ можно образовать взаимно однозначное соответствие $\varphi : \pi \leftrightarrow \varphi$ такое, что

$$\varphi(B) = \varphi[\delta(s, a)],$$

где $B \in \pi, s \in B, a \in I$.

$s \backslash I$	a	b
1	2	3
2	3	4
3	5	5
4	6	6
5	4	4
6	7	7
7	8	10
8	9	9
9	10	12
10	11	11
11	12	19
12	13	13
13	14	14
14	15	15
15	16	16
16	17	17
17	18	18
18	1	1
19	20	20
20	21	21
21	18	22
22	23	23
23	22	18

Рис. 1. Автомат А.

Пусть симметричные пары разбиений (π_{i-1}, π_i) , $2 \leq i \leq k$, образуют К-цепь и (π_k, φ) – симметричная смешанная пара.

Определение 4. Цепь $(\pi_1, \pi_2)(\pi_2, \pi_3) \dots (\pi_{k-1}, \pi_k)(\pi_k, \varphi)$ будем называть расширенной К-цепью.

Пусть $\pi_k = \{B_1, B_2\}$. Покрытие φ , образующее симметричную смешанную пару с разбиением π_k , можем найти следующим образом:

$$\varphi = \delta(\pi_k) = \{\delta(B_1), \delta(B_2)\},$$

где

$$\delta(B_i) = \{\cup \delta(s, x) \mid s \in B_i, x \in I\}.$$

Расщепляя те состояния исходного автомата, которые находятся в обоих блоках покрытия φ , получим новый расщепленный автомат, для которого существует $(K+I)$ -цепь на множестве расщепленных состояний.

Предлагаем алгоритм нахождения двухблочной $(K+I)$ -цепи на множестве расщепленных состояний S' и получения расщепленного автомата.

Алгоритм 2.

Шаг 1. $\varphi = \delta(\pi_k)$. Образовать расширенную К-цепь.

Шаг 2. $V = \delta(B_1) \cap \delta(B_2)$ – множество расщепленных состояний.

Шаг 3. Для каждого состояния s_i образуем множество $D(s_i)$:

$$D(s_i) = \begin{cases} \{s_i \mid s_i \notin V\} \\ \{s_i^1, s_i^2 \mid s_i \in V\}. \end{cases}$$

Шаг 4. $S' = \cup_{s_i \in S} D(s_i)$.

Шаг 5. В разбиениях π_j , $1 \leq j \leq k$ заменить состояние $s_i \in V$ на множество состояний $D(s_i)$.

Шаг 6. Преобразуем покрытия φ в разбиение π_{k+1} .

Если $s_i \in \delta(B_1) \& s_i \in V$, то надо заменить s_i на s_i^1 .

Если $s_i \in \delta(B_2) \& s_i \in V$, то надо заменить s_i на s_i^2 .

Шаг 7. Определить переходную функцию так, чтобы разбиения π_k и π_{k+1} на множестве S' образовали симметричную пару разбиений.

Шаг 8. Конец.

Пример. Пусть для автомата В (рис. 2) вычислены разбиения π^c и π^r :

$$\pi^c = \{\overline{1,3}, \overline{2,5,6}, \overline{4}\};$$

$$\pi^r = \{\overline{1,5}, \overline{2,4}, \overline{3,6}\}.$$

3-цепи для данного автомата не существует. Выберем двухблочную 2-цепь по минимуму оценочной функции $R(\pi_1)$. Выбранная К-цепь состоит из разбиений

$$\pi_1 = \{\overline{1,3,4}, \overline{2,5,6}\} \quad \text{и}$$

$$\pi_2 = \{\overline{1,3,5,6}, \overline{2,4}\}. \quad \text{Применим теперь алгоритм 2.}$$

Вычислим $\varphi = \delta(\pi_2) = \{\overline{1,2,4,5}, \overline{2,3,4,6}\}$. Найдем множество $V = \{\overline{1,2,4,5}\} \cap \{\overline{2,3,4,6}\} = \{\overline{2,4}\}$. Определим $D(2) = \{2^1, 2^2\}$ и $D(4) = \{4^1, 4^2\}$. Расщепленное множество состояний состоит из 8 элементов: $S' = \{\overline{1,2^1}, \overline{2^2}, \overline{3}, \overline{4^1}, \overline{4^2}, \overline{5,6}\}$. После выполнения шагов 5 и 6 получим 3-цепь на множестве S' :

$$\pi_1 = \{\overline{1,3,4^1,4^2}, \overline{2^1,2^2,5,6}\};$$

$$\pi_2 = \{\overline{1,3,5,6}, \overline{2^1,2^2,4^1,4^2}\};$$

$$\pi_3 = \{\overline{1,2^1,4^1,5}, \overline{2^2,3,4^2,6}\}.$$

В результате выполнения шага 7 получим расщепленный автомат В' (рис. 3).

Поскольку $\pi_1 \cdot \pi_2 \cdot \pi_3 = 0_\pi$, то автомат В' реализуется полностью на базе одного СР длиной 3. Метод расщепления состояний можно при необходимости использовать повторно, т.е. принять расщепленный автомат как исходный и применить снова алгоритм 2. Расщепление состояний не является целесообразным,

s \ I	a	b
1	1	5
2	4	2
3	5	5
4	6	3
5	2	4
6	2	2

Рис. 2. Автомат В.

s \ I	a	b
1	1	5
2 ¹	4 ²	2 ²
2 ²	4 ²	2 ²
3	5	5
4 ¹	6	3
4 ²	6	3
5	2 ¹	4 ¹
6	2 ¹	2 ¹

Рис. 3. Автомат В'.

если хоть один из блоков покрытия φ совпадает с множеством S .

В заключение отметим, что использование предложенного метода расщепления состояний расширяет существенно класс автоматов, которые реализуются нетривиально на базе СР. Предложенные алгоритмы реализованы в рамках разработанной на кафедре ЭВМ системы автоматизированного проектирования дискретных управляющих устройств DILOS. Экспериментальное исследование, проведенное на потоке микропрограммных автоматов, позволило сделать следующие выводы. Во-первых, большинство исследуемых автоматов имели нетривиальную СР-реализацию. Во-вторых, примерно для половины автоматов реализация "улучшилась" после применения расщепления состояний. Исследование временных характеристик показало, что разработанные программные средства применимы в качестве части систем автоматизированного проектирования управляющих устройств.

Л и т е р а т у р а

1. Fujjwaga H., Kinoshita K. Design of sequential machines utilizing extra outputs // IEEE Trans. Comput. - 1974. - Vol. C-23, N 2. P. 138-145.

2. Г и л л А. Введение в теорию конечных автоматов. - М.: Наука, 1966. - 272 с.

3. К е з в а л л и к А., Л е й с П., К р у с М. О реализации микропрограммных автоматов на сдвиговых регистрах // Тр. Таллинск. политехн. ин-та. - 1984. - № 577. - С. 85-95.

4. O'Keeffe K.H., Johnson D.L. The application of shift registers to secondary state assignment. Part I, II // IEEE Trans. Comput. - 1968. - Vol. C-17, N 10. - P. 954-977.

Finite Automata Decomposition Using Shift Register
Realization

Abstract

A method of finite automata decomposition using shift register realization is presented. Every component of decomposition is realized on a one shift register. The method of state splitting is used for the number of shift registers minimization. The described algorithms are programmed using the language LJPAS - M in the automatic design system DILOS of the chair of Computer Sciences of the TPI.

A. Keevallik, M. Kruus, K. Salum

Lõplike automaatide dekompositsioon nihkeregistritele

Kokkuvõte

Artiklis käsitletakse juhtautomaatide projekteerimist nihkeregistritele. Kasutatakse erineva pikkusega nihkeregistreid. Ülesannet käsitletakse kui teatavat liiki dekompositsiooni, kus iga komponentautomaat on realiseeritud ühel nihkeregistril. Näidatakse ka olekute lõhustamise meetodi efektiivsust antud ülesande lahendamisel. Toodud algoritmid on realiseeritud elektronarvutite kateedri automaatprojekteerimise süsteemis DILOS.

А. Кезваллик, Б. Беркман

АЛГЕБРА ПАР НЕПОЛНОСТЬЮ ОПРЕДЕЛЕННЫХ РАЗБИЕНИЙ
И КРИТЕРИЙ ИНФОРЕСКУРСА ПРИ ДЕКОМПОЗИЦИИ
АВТОМАТОВ

В связи с ростом сложности проектируемых дискретных управляющих устройств (ДУУ) и переходом на новый базис реализации в виде БИС и СБИС возрастает значение задачи формализации и автоматизации логического проектирования. Одним из наиболее эффективных способов решения этой задачи является декомпозиция исходной автоматной модели в сеть из более простых компонентных автоматов, каждый из которых удовлетворяет заданным ограничениям и может быть реализован на отдельном БИС. В [1, 2] разработан парный метод декомпозиции, при котором декомпозиция строится по полному множеству разбиений на множестве состояний исходного автомата. Показано [3, 4], что различные задачи синтеза автоматов в современном элементном базисе сводятся к соответствующим задачам выбора полного множества разбиений. В качестве автоматной модели в этих работах использована модель микропрограммного автомата (МПА), при помощи которой могут быть описаны проектируемые ДУУ реальной сложности.

В целях сокращения большого перебора, возникающего при выборе разбиений, в [3] был разработан аппарат неполностью определенных разбиений (НОР), на основе которого производится выбор разбиений для задачи декомпозиции МПА с ограничением на количество внешних входов компонентных автоматов. Важной нерешенной проблемой как для этого частного случая, так и вообще при выборе разбиений является оценка и минимизация количества внутренних входов, посредством которых производится передача информации между компонентными автоматами сети.

В данной работе дается развитие аппарата НОР. Показано, что НОР образует алгебру пар [1]. Для того, чтобы уже на этапе выбора разбиений оценить количество внутренних входов компонентных автоматов и другие параметры будущей сети, предлагается использовать понятие инфоресурса разбиений [5], распространенного на НОР.

Неполностью определенным разбиением $\rho(S)$ на множестве $S = \{s_1, \dots, s_n\}$ вызывается $\{B_1, \dots, B_n\}$ - совокупность подмножеств (блоков) множества S , обладающая следующими свойствами:

$$(1) \quad B_i = \emptyset \quad (i=1, \dots, n);$$

$$(2) \quad B_i \cap B_j = \emptyset \quad (i, j=1, \dots, n; i \neq j);$$

$$(3) \quad \bigcup_{i=1}^n B_i \subseteq S.$$

Когда множество S ясно из контекста, будем опускать его и записывать $\rho(S)$ как ρ .

Множество $\bigcup_{i=1}^n B_i$ будем называть областью определения ρ (обозначается $D(\rho)$), а множество $S \setminus D(\rho)$ будем называть областью неопределенности ρ . Под $s_p \equiv s_q(\rho)$ будем понимать принадлежность s_p и s_q одному и тому же блоку ρ .

Ясно, что любое разбиение $\pi(S)$ будет являться НОР на множестве S , причем $D(\pi) = S$.

Определим на множестве НОР $\rho(S)$ отношение частичного порядка $\leq$:

$$\rho_1 \leq \rho_2 \iff (\forall B_i \in \rho_1) (\exists B'_j \in \rho_2) [B_i \subseteq B'_j].$$

Произведением $\rho_1 \cdot \rho_2$ НОР ρ_1 и ρ_2 будем называть НОР, состоящий из всевозможных непустых пересечений блоков ρ_1 с блоками ρ_2 .

Суммой $\rho_1 + \rho_2$ НОР ρ_1 и ρ_2 будем называть НОР $\rho_3(S)$, для которого $s_p \equiv s_q(\rho_3)$, если и только если существует цепочка $s_p = s^0, s^1, \dots, s^i = s_q$ такая, что $s^i = s^{i+1}(\rho_1)$ для четных i и $s^i = s^{i+1}(\rho_2)$ для нечетных i .

Нетрудно заметить, что множество всех НОР на множестве S образует дистрибутивную решетку L с частичным порядком $\leq$, причем

$$\sup(p_1, p_2) = p_1 + p_2 \text{ и } \inf(p_1, p_2) = p_1 \cdot p_2.$$

Нулем ($0_p(S)$) и единицей ($1_p(S)$) этой решетки являются соответственно $\text{НОР } \emptyset$ и $\{S\}$.

Решетка разбиений на множестве S с нулем $0_\pi(S)$ и единицей $1_\pi(S)$ является, очевидно, подрешеткой решетки НОР .

$\text{НОР } p(S)$ можно интерпретировать как компактную запись целой совокупности $G(p)$ разбиений на S . Причем

$$G(p) = \{\pi \mid \pi \geq p\}.$$

Легко доказать, что множеству $\text{НОР } P' = \{p_i \mid i=1, \dots, k\}$ можно поставить в соответствие полное множество разбиений $P = \{\pi_i \mid \pi_i \in G(p_i), i=1, \dots, k\}$, если и только если

$$p_i \in P', p_i \leq 0_\pi. \quad (I)$$

Будем называть множество $\text{НОР } P'$, обладающее свойством (I), полным.

Под микропрограммным автоматом будем понимать систему

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

где $S = \{s_1, \dots, s_N\}$ - множество состояний;

$X = \{x_1, \dots, x_\ell\}$ - множество двоичных входных переменных (входов), задающее входной алфавит $J = \{0, 1\}^\ell$;

$Y = \{y_1, \dots, y_n\}$ - множество двоичных выходных переменных (выходов), задающее выходной алфавит $O = \{0, 1\}^n$;

$\delta: SXJ \rightarrow S$ - функция переходов;

$\lambda: SXJ \rightarrow O$ (или $S \rightarrow O$) - функция выхода (соответственно модель Мили или Мура).

$(S - S)$ парой $\text{НОР } \langle p, p' \rangle$ МПА $A = (S, X, Y, \delta, \lambda)$ называется упорядоченная пара НОР на S такая, что из $s_p \equiv s_q(p)$ следует $\delta(s_p, z) \equiv \delta(s_q, z)(p')$ для любого z из J .

Нетрудно доказать, что имеет место следующая лемма.

Лемма. Если $\langle p_1, p'_1 \rangle$ и $\langle p_2, p'_2 \rangle$ - пары НОР МПА A , то:

1) $\langle p_1 \cdot p_2, p'_1 \cdot p'_2 \rangle$ - пара НОР автомата A ;

2) $\langle p_1 + p_2, p'_1 + p'_2 \rangle$ - пара НОР автомата A .

Очевидно, что для любого НОР $\rho < \rho, 1_\rho >$ и $< 0_\rho, \rho >$ - тривиальные пары НОР.

Из леммы следует, что для любого данного НОР ρ существует наименьшее НОР $m(\rho)$ такое, что $< \rho, m(\rho) >$ - пара НОР на A . НОР $m(\rho)$ определяется по формуле

$$m(\rho) = \prod \{ \rho_i \mid < \rho, \rho_i > \text{ - пара НОР на } A \}.$$

Аналогично для данного ρ существует наибольшее НОР $M(\rho)$ такое, что $< \rho, M(\rho) >$ - пара НОР на A .

$$M(\rho) = \sum \{ \rho_i \mid < \rho_i, \rho > \text{ - пара НОР на } A \}.$$

С информационной точки зрения (для данного НОР ρ) НОР $m(\rho)$ описывает наибольшее количество информации, которое мы можем вычислить о следующем состоянии МПА A , зная только ρ (т.е. блок ρ , который содержит текущее состояние A).

Аналогично (для данного ρ) НОР $M(\rho)$ описывает наименьшее количество информации, которое мы должны иметь о текущем состоянии МПА A для вычисления следующего состояния с точностью до блока ρ .

В соответствии с [II] будем называть подмножество Δ множества $L_1 \times L_2$ (где L_1 и L_2 - конечные решетки) алгеброй пар на $L_1 \times L_2$, если и только если выполняются два следующих условия:

1. Из наличия $< \rho_1, \rho_2 >$ и $< \rho_3, \rho_4 >$ в Δ следует наличие $< \rho_1 \cdot \rho_3, \rho_2 \cdot \rho_4 >$ и $< \rho_1 + \rho_3, \rho_2 + \rho_4 >$ в Δ .
2. Для любых ρ_1 в L_1 и ρ_2 в L_2 $< \rho_1, 1_\rho >$ и $< 0_\rho, \rho_2 >$ принадлежат Δ .

Из леммы и того факта, что $< \rho, 1_\rho >$ и $< 0_\rho, \rho >$ - пары НОР на A , следует, что множество всех (S-S) пар НОР на A - алгебра пар на $L \times L$, где L - решетка НОР на S . Отметим, что в алгебре пар НОР действительны все соотношения, доказанные в [II] для алгебры пар в общем случае.

(J-S) парой НОР $< \rho_1(J), \rho_2(S) >$ МПА A называется упорядоченная пара НОР такая, что для любого $s_\rho \in S$ из $z_i \equiv z_j(\rho_1)$ следует $\delta(s_\rho, z_i) \equiv \delta(s_\rho, z_j)(\rho_2)$.

Аналогичным образом можно определить также (S-0) и (J-0) пары НОР, соответствующие введенным в [II] парам разбиений.

Легко доказать, что все эти пары НОР также образуют алгебру пар.

Заметим, что специфика модели МПА придает особую важность $(J-S)$ парам НОР, в которых разбиения структурных входных алфавитов имеют специальный вид и приводят к освобождению зависимости компонентного автомата от одного или нескольких двоичных входов.

Пусть $\rho_i(J)$ – разбиение, каждый блок которого состоит из двух элементов алфавита J , совпадающих по значениям всех двоичных входов кроме i -го. Тогда работа компонентного автомата, входной алфавит которого равен $\rho_i(J)$, не будет зависеть от i -го входа и такой автомат сможет вычислять свое состояние с точностью до $m_{J-S}(\rho_i(J))$.

Верно и обратное. Любой компонентный автомат, вычисляющий свое состояние с точностью до $\pi(S)$, где $\pi(S) \in G(m_{J-S}(\rho_i(J)))$, не будет зависеть от i -го входа. Разбиение $m_{J-S}(\rho_i(J))$ будем называть первичным α -НОР $[3]$ и обозначать $\alpha_i(S)$.

Для оценки числа внутренних входов $[2]$, используемых для передачи информации между компонентными автоматами получаемой сети, распространим на множество $(S-S)$ пар НОР разработанное в $[5]$ понятие инфоресурса разбиений, равного

$$I(\pi_2/\pi_1) = \min_{\pi_1 \cdot \pi_3 \leq \pi_2} \|\pi_3\| = \max_{i=1, \dots, \|\pi_1\|} k_i,$$

где k_i – количество блоков π_2 , пересечение которых с i -ым блоком π_1 не пусто.

Инфоресурсом НОР $\rho_2(S)$ по отношению к НОР $\rho_1(S)$ будем называть величину

$$I(\rho_2/\rho_1) \stackrel{\text{Def}}{=} \min_{\pi_1 \in G(\rho_1)} (\max_{\pi_2 \in G(\rho_2)} I(\pi_2/\pi_1)) = \max_{\pi_2 \in G(\rho_2)} (\min_{\pi_1 \in G(\rho_1)} I(\pi_2/\pi_1) = I(\pi_2^*/\pi_1^*)).$$

Здесь π_1^* и π_2^* – наименьшие разбиения из $G(\rho_1)$ и $G(\rho_2)$ соответственно.

Из определения инфоресурса для НОР и приведенной выше формулы для $I(\pi_2/\pi_1)$ следует, что

$$I(p_2/p_1) = \log_2 \left(\max_{i=1, \dots, \| \pi_1 \|} (k_i + l_i) \right),$$

где k_i — количество блоков, пересечение которых с B_i (i -ым блоком p_1) не пусто;

$l_i = \| B_i \cup (S \setminus D(p_1)) \|$ — количество элементов области неопределенности p_2 , входящих в B_i .

Для заданного НОР $p(S)$ значение $I(M(p)/p)$ будет показывать наименьшее количество внутренних входов, которые требуются автомату A_π , соответствующему наименьшему разбиению π из $Q(p)$, для вычисления его следующего состояния.

Здесь и ниже $\lceil x \rceil$ — наименьшее целое, большее или равное x .

Предложенный для НОР критерий информресурса может быть использован также в целях сокращения перебора при генерировании разбиений для декомпозиций с различными ограничениями.

В качестве примера рассмотрим задачу декомпозиции МПА с ограничением на количество входов у компонентного МПА. Будем исходить из того, что задан МПА с l двоичными входами и требуется найти полное множество разбиений для декомпозиции МПА в сеть из возможно меньшего количества компонентных МПА, у каждого из которых не более l' двоичных входов ($l' \leq l$). Эта задача возникает, например, при реализации МПА на ПЛМ или при программной реализации МПА на однокристалльных микроЭВМ с ограничением на объем памяти.

В связи с большим перебором, требуемым для точного решения поставленной задачи, приведем эвристический алгоритм, основывающийся на следующих соображениях:

1. Для того, чтобы компонентный МПА, соответствующий разбиению π , не зависел от k внешних входов ($k \leq l$), необходимо и достаточно, чтобы π было больше или равно некоторой сумме k -го порядка из первичных α -НОР:

$$\pi \geq \sum_{i \in J'} \alpha_i, \quad \text{где } J' \subseteq \{1, \dots, l\}; \|J'\| = k.$$

2. Для оценки числа внутренних входов МПА, соответствующего выбираемому НОР p , используем $I(M(p)/p)$.

3. Для оценки близости множества НОР P' к полному, используем значение $I(O_p/p_P)$, где $p_P = \prod_{p_i \in P'} p_i$.

Предлагаем алгоритм генерации полного множества разбиений, который распадается на 3 этапа.

1 этап. Генерация первичных α -разбиений. Способ генерации первичных α -разбиений описан в [3].

2 этап. Генерация полного множества НОР.

На каждом (i -ом) шаге генерируемое множество НОР P' (вначале пустое) дополняется НОР $\rho_i = \sum_{j \in J_i} \alpha_j$ таким, что

$$(a) \ell - \|J_i\| + I(M(\rho_i)/\rho_i) = \ell'; \quad J_i \subseteq \{1, \dots, \ell\};$$

$$(b) I(O_P/\rho_i \cdot \prod_{\rho_j \in P, \rho_j \neq \rho_i} \rho_j) \rightarrow \min.$$

Этот этап заканчивается в результате возникновения одной из двух ситуаций:

1. На некотором шаге $I(O_P/\rho_i \cdot \prod_{\rho_j \in P, \rho_j \neq \rho_i} \rho_j) = 0$. Построено ортогональное множество НОР.

2. На некотором шаге $I(O_P/\rho_i \cdot \prod_{\rho_j \in P, \rho_j \neq \rho_i} \rho_j) = I(O_P/\prod_{\rho_j \in P, \rho_j \neq \rho_i} \rho_j)$.

Ортогональное множество НОР не может быть построено.

3 этап. Преобразование полного множества НОР в ортогональное множество разбиений.

а) Для каждого НОР ρ_i из множества P' выбирается разбиение $\pi_i = \min_{\pi \in G(\rho_i)} \pi$. Разбиение π_i получается в результате объединения ρ_i с одноэлементными блоками, составленными из элементов множества $S \setminus D(\rho_i)$.

В результате получается полное множество разбиений $P = \{\pi_i\}$.

б) Производим укрупнение разбиений π_i из множества P . При этом на каждом шаге укрупнения пара блоков разбиения π_i объединяется в один, если это, с одной стороны, не нарушает неравенства $\ell - \|J_i\| + I(M(\pi_i)/\pi_i) \leq \ell'$, а с другой - не нарушает полноты множества P .

Описанный алгоритм реализован на языке программирования ЛЯПАС-М [6] и работает в составе автоматизированной системы проектирования дискретных управляющих устройств DILOS.

Заметим, что планируется развитие описанного аппарата в следующих направлениях. С одной стороны, предполагается использовать имеющуюся свободу при выборе полного множества НОР и при его преобразовании в множество разбиений для учета дополнительных ограничений на базис и структуру получаемых сетей. С другой стороны, предполагается распространение аппарата алгебры пар НОР на покрытия. Алгебра пар непустых подмножеств может использоваться при выборе покрытий для декомпозиции автоматов в классе эквивалентных преобразований.

Л и т е р а т у р а

1. Hartmanis J., Stearns R.E. Algebraic structure theory of sequential machines. - N.-Y.: Prentice-Hall Inc., 1966. - 209 p.
2. К е з в а л л и к А.Э. Теорема декомпозиции конечных автоматов // Автоматика и вычислительная техника. - 1974. - № 1. - С. 17-24.
3. Б е р к м а н Б. Метод выбора разбиений для декомпозиции МПА с разделением входных переменных // Тр. Таллинск. политехн. ин-та. - 1982. - № 530. - С. 93-105.
4. С у д н и ц ы н А.В., В и й е с В.Р. Задача выбора разбиений в декомпозиционном синтезе дискретных управляющих автоматов // Тр. Таллинск. политехн. ин-та. - 1986. - № 626. - С. 18-28.
5. Б е р к м а н Б. Кодирование при параллельной передаче данных и его информационный критерий // Тр. Таллинск. политехн. ин-та. - 1986. - № 626. - С. 29-42.
6. З а к р е в с к и й А.Д., Т о р о п о в Н.Р. Система программирования ЛЯПАС-М. - Минск: Наука и техника, 1978. - 238 с.

A. Keevallik, B. Berkman

Incompletely Specified Partitions in
Finite Automata Decomposition

Abstract

Optimization of the finite automata decomposition is considered in terms of partition pair algebra approach. It is shown that the functional dependencies (between component automata) in automata network can be reduced using original informational complexity measure for incompletely specified partitions. To handle automata with real complexity the microprogram automata model is used.

A. Keevallik, B. Berkman

Mittetäielikult määratud tükelduste algebra ja
inforessursi kriteerium automaatide dekompositsiooniks

Kokkuvõte

Artiklis käsitletakse probleeme, mis on seotud mikroprogrammautomaatide optimaalse dekomponeerimisega. Välja on tootatud mittetäielikult määratud tükelduste aparaat, mis võimaldab oluliselt kahandada optimeerimisülesande lahendamise arvutusmahukust. Mittetäielikult määratud tükelduste kvantitatiivseks hindamiseks kasutatakse inforessursi kriteeriumi. Viimase rakendamisel baseerub ka väljatootatud algoritm mikroprogrammautomaadi dekomponeerimiseks etteantud sisendite arvuga komponentautomaatide võrguks.

ДЕКОМПОЗИЦИЯ МИКРОПРОГРАММНЫХ АВТОМАТОВ В КЛАССЕ ГОМОМОРФИЗМОВ С ИСПОЛЬЗОВАНИЕМ ИНФОРЕСУРСА

В работах, связанных с алгебраической структурной теорией автоматов [1, 2, 3], обычно рассматривается декомпозиция автоматов по полному множеству разбиений на множестве состояний исходного автомата. При этом строится сеть, результирующий автомат которой реализует исходный автомат в классе изоморфизмов, т.е. исходный автомат изоморфен некоторому подавтомату результирующего автомата сети.

Оказывается, однако, что зачастую декомпозиции с требуемыми свойствами (с заданными ограничениями на получаемую сеть) не могут быть найдены в рамках этого класса.

В работах [3, 4] рассматривается декомпозиция автоматов в классе эквивалентных преобразований (гомоморфизмов), позволяющая расширить класс декомпозируемых с заданными ограничениями автоматов. В этих работах используется так называемая техника "расщепления состояний" ("state splitting"), которая сводится к следующему. Для исходного автомата A строится его гомоморфный образ A' такой, что множество покрытий P_τ (необязательно полное) на множестве состояний S автомата A переводится в полное множество разбиений P_π на расщепленном множестве состояний S' автомата A' . Затем производится декомпозиция автомата A' по полному множеству разбиений P_π .

К сожалению, в имеющихся работах остаются нерешенными следующие проблемы:

1. Не формализована техника расщепления состояний, не решен вопрос о возможности этого расщепления. Так, одно и то же множество покрытий для одних автоматов может, а для

других — не может быть преобразовано к полному множеству разбиений.

2. Получаемый гомоморфный образ A' по сложности может значительно превышать исходный автомат A , что существенно усложняет процесс декомпозиции.

3. Отсутствуют методы декомпозиции в классе гомоморфизмов для модели микропрограммного автомата (МПА) [2], которая позволяет описывать управляющие устройства реальной сложности.

В настоящей работе будем рассматривать декомпозицию МПА только по полному множеству покрытий, благодаря чему удастся полностью формализовать процесс декомпозиции и избежать применения техники расщепления состояний в явном виде.

Определение 1. Под микропрограммным автоматом будем понимать систему

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

где $S = \{s_i \mid i = 1, \dots, m\}$ — множество внутренних состояний;

$X = \{x_1, \dots, x_p\}$ — множество входных двоичных переменных (каналов), из которых МПА получает входные наборы z — элементы булева пространства $\Psi(X) = \{0, 1\}^p$ (здесь и ниже под $\Psi(X)$ подразумевается $\{0, 1\}^{||X||}$;

$Y = \{y_1, \dots, y_n\}$ — множество выходных двоичных переменных (каналов), на которые МПА выдает выходные наборы w — элементы булева пространства $\Psi(Y)$;

$\delta: D(\delta) \rightarrow S$ — функция переходов,

где $D(\delta) \subseteq S \times \Psi(X)$ — область определения функции δ ;

$\lambda: D(\lambda) \rightarrow \Psi(Y)$ — функция выходов,

где $D(\lambda) \subseteq S$ — (модель Мура).

Определение 2. Под сетью МПА (Мура) будем понимать систему $N = (B, X, Y, F, g)$, где

I. $B = \{A_i \mid i \in J = \{1, \dots, k\}\}$ — множество компонентных МПА (базис сети);

$$A_i = (S_i, X_i' \cup X_i'', Y_i, \delta_i, \lambda_i).$$

Здесь $X'_i \subseteq X$ и X''_i — соответственно множества внешних и внутренних входов МПА A_i , задающие внешний и внутренний входные алфавиты $\Psi(X'_i)$ и $\Psi(X''_i)$; Y_i — множество выходов, задающее его выходной алфавит $\Psi(Y_i)$.

2. X — множество входов сети, задающее входной алфавит $\Psi(X)$.

3. Y — множество выходов сети, задающее выходной алфавит $\Psi(Y)$.

4. $F = \{f_i | i \in J\}$ — множество функций соединения (структура сети);

$$f_i: X \cap \Psi(Y_j^*) \rightarrow \Psi(X''_i); J_i \subseteq J; Y_j^* \subseteq Y_j.$$

5. $g: X \cap \Psi(Y_j^+) \rightarrow \Psi(Y)$ — функция выхода сети; $J_g \subseteq J; Y_j^+ \subseteq Y_j$.

Определение 3. Проекцией элемента z булева пространства $\Psi(X)$ на булево пространство $\Psi(X')$ ($X' \subseteq X$) называется элемент z' пространства $\Psi(X')$ (пишем: $z' = pr(z, X')$), значения координат которого равны значениям соответствующих им координат z .

Пример. Если $X = \{x_1, x_2, x_3, x_4\}$ и $X' = \{x_2, x_3\}$, то $pr((0101), X') = (10)$.

Определение 4. Под результирующим автоматом A_N сети МПА $N = \{\{A_i\}, X, Y, \{f_i\}, g\}$ следует понимать микропрограммный автомат $A_N = (S_N, X, Y, \delta_N, \lambda_N)$, где

1. X, Y — множества входных и выходных каналов автомата A_N , совпадающие с множествами входных и выходных каналов сети N соответственно.

2. $S_N = \bigcup_{i \in J} S_i$ — множество состояний результирующего автомата A_N .

3. $\delta_N: S_N \times \Psi(X) \rightarrow S_N$ — функция переходов автомата A_N , определяемая следующим образом:

$$\delta_N(s_{p_1}^1, \dots, s_{p_k}^k, z) = (\delta_1(s_{p_1}^1, z'_1 \times z''_1), \dots, \delta_k(s_{p_k}^k, z'_k \times z''_k)),$$

где

$$z \in \Psi(X); s_{p_i}^i \in S_i; z'_i = pr(z, X'_i); z''_i = f_i(X \cap pr(\lambda_j(s_{p_j}^j), Y_j^*)) \quad (i \in J).$$

4. $\lambda_N: S_N \rightarrow \Psi(Y)$ — функция выхода автомата A_N :

$$\lambda_N(s_{p_1}^1, \dots, s_{p_k}^k) = g(\underset{j \in J_g}{x}, \text{pr}(\lambda_j(s_{p_j}^j), Y_j^+)).$$

Определение 5. МПА $A' = \{S', X', Y', \delta', \lambda'\}$ называется подавтоматом МПА $A = (S, X, Y, \delta, \lambda)$, если и только если $S' \subseteq S$; $X' \subseteq X$; $Y' \subseteq Y$ и для каждого $s \in S'$ и $z' \in \Psi(X')$ существует $z \in \Psi(X)$ такой, что

$$z' = \text{pr}(z, X') \& \delta'(s, z') = \delta(s, z) \& \lambda'(s) = \text{pr}(\lambda(s), Y').$$

Определение 6. МПА $A' = (S', X, Y, \delta', \lambda')$ называется гомоморфным образом МПА $A = (S, X, Y, \delta, \lambda)$, если и только если существует сюръекция $\varphi: S' \rightarrow S$ такая, что для каждого $s_p \in S'$ и $z \in \Psi(X)$ имеет место

$$(1) \delta(\varphi(s_p), z) = \varphi(\delta'(s_p, z)); (2) \lambda(\varphi(s_p), z) = \lambda'(s_p, z).$$

Определение 7. МПА A' называется реализацией МПА A (пишем: $R(A) = A'$), если и только если у автомата A' существует гомоморфный автомату A подавтомат.

Под задачей общей декомпозиции МПА A понимается построение сети N такой, чтобы $R(A) = A_N$.

При построении таких декомпозиций будем исходить из имеющегося полного множества P_τ покрытий на множестве состояний S МПА A . Каждому покрытию τ_i будем ставить в соответствие компонентный МПА A_i сети, вычисляющий состояние A с точностью до блоков τ_i . Для вычисления своего следующего состояния автомату A_i необходима дополнительная информация о состоянии автомата A . Эта информация должна быть достаточной, чтобы, зная текущий блок τ_i , получить блок τ_j , в который осуществляется переход. Минимальное покрытие, которое требуется для вычисления перехода в блоки τ_i , равно $M(\tau_i)$ [3]. Минимальное количество битов для передачи недостающей в τ_i до $M(\tau_i)$ информации может быть определено при помощи введенного в [5] понятия инфоресурса.

Определение 8. Инфоресурсом $I(\tau_2/\tau_1)$ покрытия $\tau_2(S)$ относительно покрытия $\tau_1(S)$ называется величина

$$I(\tau_2/\tau_1) \stackrel{\text{Def}}{=} \log_2 \left(\min_{\tau_1, \tau_3 \leq \tau_2} I(\tau_3) \right), \quad (I)$$

причем инфоресурс покрытия τ_3 равен

$$I(\tau_3)^{\text{Def}} = \log_2(\min_{\pi_3 \leq \tau_3} \|\pi_3\|). \quad (\text{Ia})$$

Отсюда получим

$$I(\tau_2/\tau_1)^{\text{Def}} = \log_2(\min_{\tau_1 \cdot \pi_3 \leq \tau_2} \|\pi_3\|). \quad (\text{Iб})$$

Из всего сказанного следует, что минимальное количество внутренних входов у автомата, соответствующего τ_i (число бит для передачи недостающей в τ_i до $M(\tau_i)$ информации) равно $I(M(\tau_i)/\tau_i)$.

В [5] было отмечено, что среди множества подобных покрытий, соответствующих одному и тому же отношению толерантности, имеется одно максимальное (полное), представляющее собой совокупность клик этого отношения.

Полное покрытие, подобное заданному покрытию $\tau(S)$, будем обозначать через $\hat{\tau}$.

В [5] было показано, что нахождение инфоресурса для полных покрытий сводится к задаче раскраски графа

$$I(\hat{\tau}_2/\hat{\tau}_1) = \log_2(\text{chr}(R_1^t \cap R_2^t)),$$

где R_1^t и R_2^t — отношения толерантности, соответствующие покрытиям $\hat{\tau}_1$ и $\hat{\tau}_2$;

$\text{chr}(R)$ — хроматическое число графа отношения R .

К сожалению, этот результат не верен для произвольных покрытий.

Теорема I. Инфоресурс $I(\hat{\tau}_2/\tau_1)$ полного покрытия $\hat{\tau}_2(S)$ относительно произвольного покрытия $\tau_1(S)$ равен $I(\hat{\tau}_2/\hat{\tau}_1)$.

$$I(\hat{\tau}_2/\tau_1) = I(\hat{\tau}_2/\hat{\tau}_1). \quad (2)$$

Доказательство.

1. Докажем, что $I(\hat{\tau}_2/\tau_1) \leq I(\hat{\tau}_2/\hat{\tau}_1)$.

$$(\tau_1 \leq \hat{\tau}_1) \Rightarrow [\hat{\tau}_1 \cdot \pi_3 \leq \hat{\tau}_2 \Rightarrow (\tau_1 \cdot \pi_3 \leq \hat{\tau}_2)] \Rightarrow I(\hat{\tau}_2/\tau_1) \leq I(\hat{\tau}_2/\hat{\tau}_1).$$

2. Доказательство того, что $I(\hat{\tau}_2/\tau_1) \geq I(\hat{\tau}_2/\hat{\tau}_1)$, или равносильного утверждения $(\tau_1 \cdot \pi_3 \leq \hat{\tau}_2) \Rightarrow (\hat{\tau}_1 \cdot \pi_3 \leq \hat{\tau}_2)$, может быть произведено от противного и для краткости опускается.

Теорема 2. Инфоресурс $I(\tau_2/\pi_1)$ покрытия $\tau_2(S)$ относительно разбиения $\pi_1(S)$ равен

$$I(\tau_2/\pi_1) = \log_2(\max_{B_i \in \pi_1} k_i), \quad (3)$$

где k_i — мощность минимального покрытия $B_i \in \pi_1$ блоками из $\bar{\tau}_2(B_i)$ — проекции [5] покрытия $\tau_2(S)$ на блок B_i .

Доказательство

Согласно свойству инфоресурса [5], $I(\tau_2/\pi_1) = I(\tau_2 \cdot \pi_1/\pi_1)$. Таким образом, требуется доказать, что

$$\min_{\pi_1 \cdot \pi_3 \leq \tau_2 \cdot \pi_1} \|\pi_3\| = k^*, \text{ где } k^* = \max_{B_i \in \pi_1} k_i.$$

I. Докажем, как построить π_3 , для которого $\|\pi_3\| = k^*$ и $\pi_1 \cdot \pi_3 \leq \pi_1 \cdot \tau_2$.

Рассмотрим $\{\tau_3(B_i) | B_i \in \pi_1\}$, где $\tau_3(B_i)$ — минимальное покрытие блока B_i блоками проекции $\bar{\tau}_2(B_i)$. Очевидно, что для любого $B_i \in \pi_1$ имеет место неравенство $\tau_3(B_i) \leq \bar{\tau}_2(B_i)$, так как $\tau_3(B_i) \subseteq \bar{\tau}_2(B_i)$. Кроме того, $\max_{B_i \in \pi_1} \|\tau_3(B_i)\| = k^*$.

Построим для каждого $\tau_3(B_i)$ разбиение $\pi_3(B_i)$, для которого $\|\pi_3(B_i)\| = \|\tau_3(B_i)\|$ и $\pi_3(B_i) \leq \tau_3(B_i)$. Для этого каждый (j -й) блок C'_{ij} разбиения $\pi_3(B_i)$ выберем равным

$$C'_{ij} = C_{ij} \setminus \bigcup_{l=1}^{j-1} C_{il},$$

где C_{ij} — j -й блок покрытия $\tau_3(B_i)$ ($j=1, \dots, \|\tau_3(B_i)\|$). Пустых блоков при этом не получается, так как $\tau_3(B_i)$ — минимальное покрытие множества B_i . Таким образом,

$$\max_{B_i \in \pi_1} \|\pi_3(B_i)\| = k^*.$$

Примем, что $C'_{ij} = \emptyset$, если $\|\pi_3(B_i)\| < j \leq k^*$.

Теперь построим разбиение $\pi_3(S)$ следующим образом:

$$\pi_3 = \{D_j | D_j = \bigcup_{i=1}^{\|\pi_1\|} C'_{ij}; j=1, \dots, k^*\}.$$

Из способа построения разбиения $\pi_3(S)$ следует, что

$$\pi_1 \cdot \pi_3 = \bigcup_{B_i \in \pi_1} \pi_3(B_i) \leq \bigcup_{B_i \in \pi_1} \tau_3(B_i) \leq \bigcup_{B_i \in \pi_1} \tau_2(B_i) \leq \max_{B_i \in \pi_1} (\bigcup_{B_i \in \pi_1} \tau_2(B_i)) = \pi_1 \cdot \tau_2.$$

Итак, для разбиения $\pi_3(S)$ $\|\pi_3\| = k^*$ и $\pi_1 \cdot \pi_3 \leq \pi_1 \cdot \tau_2$.

2. Доказательство того, что для любого $\pi_3(S)$ из $\pi_1 \cdot \pi_3 \leq \pi_1 \cdot \tau_2$ следует $\|\pi_3\| \geq k^*$, может быть проведено от противного и для краткости опускается.

С целью нахождения $I(\tau_2/\tau_1)$ для произвольных покрытий воспользуемся тем, что, с одной стороны, $\tau_2 \leq \hat{\tau}_2$, а с другой — тем, что (как было доказано в теореме I) $(\tau_1 \cdot \pi_3 \leq \hat{\tau}_2) \Rightarrow (\hat{\tau}_1 \cdot \pi_3 \leq \hat{\tau}_2)$. Из этого следует, что

$$(\tau_1 \cdot \pi_3 \leq \tau_2) \Rightarrow (\tau_1 \cdot \pi_3 \leq \hat{\tau}_2) \Rightarrow (\hat{\tau}_1 \cdot \pi_3 \leq \hat{\tau}_2).$$

Таким образом, любое π_3 , удовлетворяющее неравенству $\tau_1 \cdot \pi_3 \leq \tau_2$, будет одновременно разбиением на одноцветные классы вершин для правильной раскраски графа отношения $R_1^t \cap \bar{R}_2^t$ (обратное не верно).

Отсюда, нахождение π_3 из минимального числа блоков, удовлетворяющего $\tau_1 \cdot \pi_3 \leq \tau_2$, сводится к специальной задаче раскраски графа отношения $R' = R_1^t \cap \bar{R}_2^t$, при которой для решения вопроса о раскраске некоторой вершины s_p i -й краской (о помещении состояния s_p в блок B_i разбиения π_3) необходимо, чтобы кроме обычного условия $(\forall s_q \in B_i) (\neg s_p R'_s q)$, выполнялось также условие

$$(\forall C_j \in \tau_1) (\exists D_i \in \tau_2) ((B_i \cup \{s_p\}) \cap C_j \in D_i). \quad (4)$$

Получение минимальной раскраски с учетом дополнительного условия (4), таким образом, и будет являться решением задачи нахождения инфоресурса для покрытий в общем случае.

Приведем теперь алгоритм построения сети МПА по ортогональному множеству покрытий. В этом алгоритме кодирование разбиений и покрытий и выбор кодовых разбиений предполагают использование инфоресурса [5], а под $\pi(y_j)$ подразумевается разбиение $\{B_1, B_2\}$, где $s_p \in B_i \Leftrightarrow pr(\lambda(s_p), \{y_j\}) = 0$.

Алгоритм декомпозиции МПА по полному множеству покрытий

I. Для каждого $\tau_i \in P_\tau$ строим множества X'_i и Y'_i соответственно, минимальное, требуемое i -му компонентному МПА множество входных переменных, и максимальное, вычисляемое i -ым МПА множество выходных переменных. Принадлежность переменных $y_u \in Y$ и $x_t \in X$ множествам X'_i и Y'_i соответственно определяется условиями:

$$y_u \in Y'_i \Leftrightarrow \tau_i \leq \pi(y_u);$$

$$x_t \in X'_i \Leftrightarrow (\exists z \in \Psi(X)) (\exists s_p \in S) \neg [\delta(s_p, z) \equiv \delta(s_p, z')(\tau_i)],$$

где $\bar{z}' \in \Psi(X)$; $\text{pr}(z', \{x_t\}) = \bar{1} \text{pr}(z, \{x_t\})$; $\text{pr}(z', X \setminus \{x_t\}) = \text{pr}(z, X \setminus \{x_t\})$.

2. Образует множество $Y' = \bigcup_{i=1}^K Y'_i$ внешних выходов, вы-

числяемых непосредственно компонентными автоматами сети:

3. Для каждого $y_j \in Y'$ исключаем y_j из всех Y'_i , кроме первого, которому они принадлежат; формируем

$$\pi'_i = \prod_{y_j \in Y'_i} \pi(y_j) \quad (i \in J);$$

4. Для каждого $i \in J$ ищем τ''_i из минимального количества блоков, для которого $\tau''_i \cdot \pi'_i = \tau_i$ (один возможный способ описан при доказательстве теоремы 2);

5. Для кодирования τ''_i составляем алфавиты

$$Y''_i \quad (\|Y''_i\| = I(\tau''_i)).$$

где каждый $y_j \in Y''_i$ соответствует кодовому покрытию $\tau(y_j)$ для формируемой кодировки τ''_i [5].

6. Для каждого τ_i ($i \in J$) строим вспомогательную функцию $F_i: S \times \Psi(X'_i) \rightarrow \tau_i$, где $F_i(s_p, z') = \tau_i[\delta(s_p, z)]$. Здесь $s_p \in S$; $z' \in \Psi(X'_i)$; $z \in \Psi(X)$; $z' = \text{pr}(z, X'_i)$.

При этом в качестве значения $F_i(s_p, z')$ выбирается один из таких блоков τ_i , которому принадлежат $\delta(s_p, z)$ для всех z , для которых $z' = \text{pr}(z, X'_i)$. Существование такого блока следует из способа построения X'_i .

7. Строим разбиения $M(\tau_i)$ [3, 4] как максимальные, для которых имеет место $s_p \equiv s_q (M(\tau_i)) \Leftrightarrow F_i(s_p, z') = F_i(s_q, z')$ для любого $z' \in \Psi(X'_i)$ ($i \in J$).

8. Строим разбиения (покрытия) $\tilde{\tau}_i$ ($i \in J$), определяемые следующим способом: $\tau_i \cdot \tilde{\tau}_i \in M(\tau_i)$; $I(\tilde{\tau}_i) = I(M(\tau_i) / \tau_i)$. Для поиска $\tilde{\tau}_i$ используем описанные выше методы раскраски графов.

9. Для кодирования $\tilde{\tau}_i$ ($i \in J$) составляем множества X''_i , где $\|X''_i\| = I(\tilde{\tau}_i)$. Выбор $x_j \in X''_i$ определяется при этом выбором соответствующего кодового покрытия $\tau(x_j)$. ($\prod_{x_j \in X''_i} \tau(x_j) = \tilde{\tau}_i$).

10. Образует для каждого $\tilde{\tau}_i$ ($i \in J$) функцию соединения $f_i: X \times \Psi(Y_j^*) \rightarrow \Psi(X''_i)$ или $f'_i: X \times \tau(y_r) \rightarrow \tilde{\tau}_i$, совпадающую с f_i и с точностью до переобозначения $y_r \leftrightarrow \tau(y_r)$. Здесь

$Y(f_i) \in Y^* = \bigcup_{h=1}^k (Y'_h \cup Y''_h)$. Выбор множества $Y(f_i)$ производится из соображений $\|Y(f_i)\| \rightarrow \min$ с учетом условия

$\prod_{y_r \in Y(f_i)} \tau(y_r) \leq \tilde{\tau}_i$. Последнее условие выполнимо, так как $\prod_{y_v \in Y^*} \tau(y_v) = \prod_{i=1}^k \tau_i = 0$. Значение функции f'_i на наборе $(B'_1, \dots, B'_t)$ равно B' , если и только если $\prod_{r=1}^t B'_r \in B'$ (здесь $B'_r \in \tau(y_r)$; $B' \in \tilde{\tau}_i$; $t = \|Y(f_i)\|$).

II. Аналогично образуем функции g_j для реализации внешних выходов сети y_j , не вычисляемых непосредственно компонентными автоматами ($y_j \in Y \setminus Y'$). $g_j: X \times \tau(y_h) \rightarrow \pi(y_j)$. Значение g_j равно:

$$g_i(B^1, \dots, B^u) = \begin{cases} B_1, & \text{если } \bigcap_{h=1}^u B^h \in B_1; \\ B_2, & \text{если } \bigcap_{h=1}^u B^h \in B_2. \end{cases} \quad (B^h \in \tau(y_h); y_h \in Y(g_i); u = \|Y(g_i)\|).$$

Прямое произведение всех g_i на тождественные функции $\pi(y_j) \rightarrow \pi(y_j)$ ($y_j \in Y'$) дает функцию $g: \prod_{j \in J_q} \Psi(Y_j^+) \rightarrow \Psi(Y)$ (с точностью до переобозначения $\Psi(\{y_j\}) \rightarrow \pi(y_j)$).

12. Компонентные автоматы A_i ($i \in J$) строим следующим образом:

- а) $S_i = \tau_i$;
- б) $X_i = X'_i \cup X''_i$;
- в) $Y_i = Y'_i \cup Y''_i$;
- г) $\delta_i(B_j, z) = F_i(s_p, pr(z, X'_i))$, где $B_j \in \tau_j$; $z \in \Psi(X_i)$; $s_p \in B_j \cap B_j^*$;

$$B_j^* = \bigcap_{h, \ell} B_{\ell}^h; B_{\ell}^h \in \tau(x_h) \quad \ell = \begin{cases} 1, & \text{если } pr(z, \{x_h\}) = 0; \\ 2, & \text{если } pr(z, \{x_h\}) = 1; \end{cases} \quad (x_h \in X''_j)$$

- д) $\lambda_1(B_j) = w$, где $w \in \Psi(Y)$; $B_j \in \tau_i$; $pr(w, Y'_i) = pr(\lambda(s_p)Y'_i)$; $s_p \in B_j$;

для любого $y_h \in Y''_i$ $pr(w, \{y_h\}) = \begin{cases} 0, & \text{если } B_j \in B_1^h \in \pi(y_h); \\ 1, & \text{если } B_j \in B_2^h \in \pi(y_h). \end{cases}$

Теорема 3. Результатом приведенного алгоритма декомпозиции МПА A является сеть N такая, что $A_N = R(A)$.

Доказательство. Для доказательства теоремы построим под-автомат A'_N автомата A_N , являющийся гомоморфным образом A .

1. Множество состояний S'_N автомата A'_N выберем равным

$$S'_N = \{s'_p | p=1, \dots, u\}, \text{ где } s'_p = (B_1^p, \dots, B_k^p); B_v^p \in \tau_v (v \in J); \bigcap_{v \in J} B_v^p = \emptyset; u -$$

количество комбинаций блоков B_v^p , пересечение которых не пусто.

Заметим, что для любого $s_p \in S$ существует хотя бы один $s'_p \in S'_N$ такой, что $\bigcap_{v \in J} B_v^p = s_p$ и для каждого $s'_p \in S'_N$ существует единственный такой $s_p \in S$ (это следует из полноты множества P_τ).

Поэтому можно построить сюръекцию $\varphi: S'_N \rightarrow S$ следующим образом:

$$\varphi(s'_p) = s_p \iff \bigcap_{v \in J} B_v^p = s_p.$$

2. Докажем, что для любых $s'_p \in S'_N$ и $z \in \Psi(X)$ имеет место

$$\varphi(\delta_N(s'_p, z)) = \delta(\varphi(s'_p), z).$$

Обозначим $s_p = \varphi(s'_p) = \bigcap_{v \in J} B_v^p$ и $s_q = \delta(s_p, z)$. Тогда, поскольку

$\delta_N(s'_p, z) = (\delta_1(B_1^p, z_1), \dots, \delta_k(B_k^p, z_k))$, требуется доказать, что для любого $i \in J$, $s_q \in \delta_i(B_i^p, z_i)$. Здесь $z_i \in \Psi(X_i)$ определяется так:

$$z'_i = pr(z_i, X'_i) = pr(z, X'_i); z''_i = pr(z_i, X''_i) = f'_i(\bigcap_{y_r \in Y(f_i)} B'_r),$$

где в свою очередь, $B'_r \in \tau(y_r)$; $\bigcap_{y_r \in Y(f_i)} B'_r$ включает в себя $\bigcap_{v \in J} B_v^p = \{s_p\}$.

Таким образом, $\delta(B_i^p, z_i) = F_i(s_p, pr(z, X'_i))$ и $s_q \in \delta_i(B_i^p, z_i)$.

3. Докажем теперь, что для любого $s'_p \in S'_N$ ($s'_p = (B_1^p, \dots, B_k^p)$) $\lambda_N(s'_p) = \lambda(\varphi(s'_p))$.

Обозначим $s_p = \varphi(s'_p)$; $w' = \lambda_N(s'_p)$; $w = \lambda(s_p)$. Докажем, что $w' = w$.

а) Для любого $y_h \in Y'$ существует A_i такой, что $y_h \in Y'_i$. При этом $\pi(y_h) \geq \pi_i$. Откуда $pr(w', \{y_h\}) = pr(\lambda_i(B_i^p), \{y_h\}) = pr(w, \{y_h\})$.

б) Для любого $y_j \in Y \setminus Y'$ имеет место следующее. Пусть $pr(w, \{y_j\}) = 0$. Тогда $s_p \in B_1$. Откуда $g'_i(B_1^p, \dots, B_k^p) = B_1$ (здесь

Исходный МПА А:

s_p	s_q	$X(s_p, s_q)$	$Y(s_p)$
1	1	x_1	$y_1 y_3$
	2	$\bar{x}_1$	
2	4	1	y_3
3	4	x_4	$y_2 y_3$
	5	$\bar{x}_4$	
4	2	$x_2 x_4$	$y_1 y_2 y_3$
	3	$x_2 \bar{x}_4$	
	4	$\bar{x}_2$	
5	1	$x_1 \bar{x}_3$	$y_1 y_2$
	4	$\bar{x}_1 \bar{x}_3$	
	5	x_3	

Декомпозиционные покрытия:

$$\tau_1 = \overline{1, 5}; \overline{2, 3}; \overline{4, 5} = \{B_1, B_2, B_3\};$$

$$\tau_2 = \overline{1, 2, 4}; \overline{3, 4}; \overline{5} = \{C_1, C_2, C_3\};$$

Существенные входы и выходы:

$$XMIN(A, \tau_1) = \{x_1, x_2\};$$

$$XMIN(A, \tau_2) = \{x_3, x_4\};$$

$$YMAX(A, \tau_1) = \{y_1\};$$

$$YMAX(A, \tau_2) = \{y_3\}; Y' = \{y_1, y_3\};$$

Кодовые разбиения и покрытия:

$$\pi(y_1) = \overline{2, 3}; \overline{1, 4, 5};$$

$$\pi(y_2) = \overline{1, 2}; \overline{3, 4, 5};$$

$$\pi(y_3) = \overline{5}; \overline{1, 2, 3, 4};$$

$$\tau(y_4) = \overline{1, 2, 3, 5}; \overline{4, 5};$$

$$\tau(y_5) = \overline{1, 2, 4}; \overline{3, 4, 5};$$

Функция выхода для $Y \setminus Y'$:

$$g_2: y_2 = y_4 \vee y_5.$$

Вспомогательные функции F_1 и F_2 :

F_1 :

s_p	B_i	$X(s_p, B_i)$
1	B_1	x_1
	B_2	$\bar{x}_1$
2	B_3	1
3	B_3	1
4	B_2	x_2
	B_3	$\bar{x}_2$
5	B_1	x_1
	B_3	$\bar{x}_1$

F_2 :

s_p	C_i	$X(s_p, C_i)$
1	C_1	1
2	C_1	1
3	C_1	x_4
	C_3	$\bar{x}_4$
4	C_1	x_4
	C_2	$\bar{x}_4$
5	C_1	$\bar{x}_3$
	C_3	x_3

$$\text{Разбиения } M(\tau_1) = \overline{1}; \overline{2, 3}; \overline{4, 5}; \quad M(\tau_2) = \overline{1, 2}; \overline{3}; \overline{4, 5}.$$

Функции соединения: $f_1: x_5 = y_3$; $f_2: x_6 = y_4$.

$$(I(\tilde{\pi}_1) = I(\tilde{\pi}_2) = I(M(\tau_1)/\tau_1) = I(M(\tau_2)/\tau_2) = 1;$$

$$\tilde{\pi}_1 = \overline{5}; \overline{1, 2, 3, 4}; \quad \tilde{\pi}_2 = \overline{1, 2, 3, 5}; \overline{4, 5}.$$

Рис. 1. Декомпозиция МПА. Пример.

Компонентные МПА сети A_1 и A_2 :

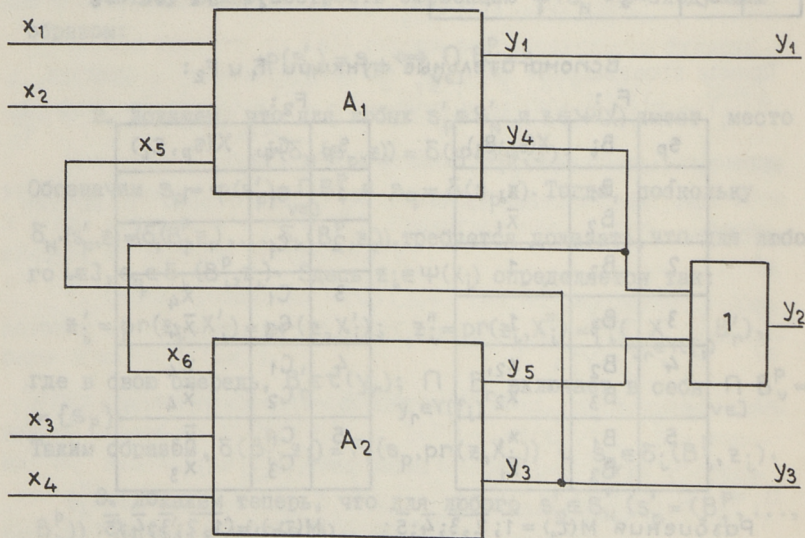
A_1 :

B_i	B_j	$X(B_i, B_j)$	$Y(B_i)$
B_1	B_1	x_1	y_1
	B_2	$\bar{x}_1 x_5$	
	B_3	$\bar{x}_1 \bar{x}_5$	
B_2	B_3	1	—
B_3	B_1	$x_1 \bar{x}_5$	$y_1 y_4$
	B_2	$x_2 x_5$	
	B_3	$\bar{x}_1 \bar{x}_5$	
	B_3	$\bar{x}_2 x_5$	

A_2 :

C_i	C_j	$X(C_i, C_j)$	$Y(C_i)$
C_1	C_1	$\bar{x}_6$	y_3
	C_1	$x_4 x_6$	
	C_2	$\bar{x}_4 x_6$	
C_2	C_1	x_4	$y_3 y_5$
	C_2	$\bar{x}_4 x_6$	
	C_3	$\bar{x}_4 \bar{x}_6$	
C_3	C_1	$\bar{x}_3$	y_5
	C_3	x_3	

Сеть из 2-х компонентных МПА:



$$S'_N = \{(B_1, C_1), (B_1, C_3), (B_2, C_1), (B_2, C_2), (B_3, C_1), (B_3, C_2), (B_3, C_3)\};$$

$$\varphi: (B_1, C_1) \rightarrow 1; (B_1, C_3) \rightarrow 5; (B_2, C_1) \rightarrow 2; (B_2, C_2) \rightarrow 3;$$

$$(B_3, C_1) \rightarrow 4; (B_3, C_2) \rightarrow 4; (B_3, C_3) \rightarrow 5.$$

Рис. 2. Декомпозиция МПА. Пример. (Продолж.)

$B^h \in \tau(y_h); y_h \in Y(q_j); u = \|Y(q_j)\|$), так как для любого $h = 1, \dots, v$ имеем $\varepsilon_p \in B^h$, а $B^1 \cap \dots \cap B^u \subseteq q'_j(B^1, \dots, B^u)$. Таким образом, $pr(w', \{y_j\}) = q_j(B^1, \dots, B^u) = 0 = pr(w, \{y_j\})$.

Аналогично доказывается, что $pr(w', \{y_j\}) = 1 \Leftrightarrow pr(w, \{y_j\}) = 1$. Итак, A'_N — φ -гомоморфный образ A . $A'_N = R(A)$. Теорема доказана.

Пример декомпозиции автомата A приведен на рис. 1, 2.

Отметим, что приведенный алгоритм, благодаря использованию критерия информресурса, минимизирует количество входов у компонентных автоматов получаемой сети.

Заключение. В настоящей статье разработан алгоритм декомпозиции МПА по полному множеству покрытий, не требующий отдельного этапа расщепления состояний и упрощающий функциональные блоки сети. Для дальнейшего развития методов декомпозиции в классе гомоморфизмов предполагается использование свободы при образовании множества выходных переменных Y''_i и при кодировании разбиений $\mathcal{N}_i$ в целях минимизации функций выхода и соединения сети. Требуют развития также вопросы выбора множества покрытий для декомпозиции, поэтому предполагается использовать аппарат неполностью определенных разбиений и покрытий.

Л и т е р а т у р а

1. К е е в а л л и к А.Э. Теорема декомпозиции конечных автоматов // Автоматика и вычислительная техника. — 1974. — № 1. — С. 17–24.

2. J a c o b s o n G., К е е v a l l i k A., L e i s P. Some aspects of the construction of the microprogram automata networks // Proceedings of the Second International Symposium on Discrete Systems. Dresden. — 1977. — Vol. 1. — P. 120–128.

3. H a r t m a n i s J., S t e a r n s R.E. Algebraic structure theory of sequential machines. N.-Y.: Prentice-Hall Inc., — 1966. — 209 p.

4. C u r t i s A.H. Generalized decomposition theory of finite sequential machines // NASA Technical Note. Washington, — 1976. — 46 p.

5. Б е р к м а н Б. Кодирование при параллельной передаче данных и его информационный критерий // Тр. Таллинск. политехн. ин-та. - 1986. - № 626. - С. 29-42.

B. Berkman

Informational Criterion for Microprogram Automata
Decomposition in the Class of Homomorphisms

Abstract

A method for microprogram automata decomposition in the class of their homomorphic transformations has been developed. The decomposition can be obtained through the full set of covers on the set of states of automata. To minimize the number of inputs of component automata the informational criterion for cover pairs has been used.

B. Berkman

Mikroprogrammautomaatide dekomponeerimine
homomorfismide klassis inforessursi alusel

Kokkuvõtte

Artiklis kirjeldatakse meetodit mikroprogrammautomaatide dekomponeerimiseks homomorfismide klassis. Dekompositsioon leitakse automaadi olekute hulgal määratud katete täieliku süsteemi alusel. Võrreldes tuntud dekompositsioonimeetoditega automaadi isomorfismide klassis, võimaldab väljatöötatud meetod laiendada dekomponeeritavate automaatide klassi. Inforessursi alusel minimeeritakse komponentautomaatide sisendite arvu.

ВЫБОР РАЗБИЕНИЙ ДЛЯ ИТЕРАТИВНОЙ ДЕКОМПОЗИЦИИ КОНЕЧНЫХ АВТОМАТОВ

1. Введение. Большинство известных методов декомпозиции конечных автоматов основано на разложении исходного автомата в каскадную сеть компонентных автоматов, то есть в сеть без обратной связи, где соединения между компонентами не образуют направленных циклов. В то же время известно, что возможности сетей автоматов с обратными связями превосходят возможности автоматов, построенных только по принципу параллельно-последовательного или каскадного соединения компонентов [4, 5]. Из этого следует, что суммарное число состояний компонентов каскадной сети, реализующей данный автомат, заметно увеличивается по сравнению с числом состояний исходного автомата.

Предлагаемый метод итеративной декомпозиции основан на использовании возможностей обратной связи для сокращения шагов разложения автомата на сеть. В начале декомпозиции в данном случае неизвестна структура сети, а полученные компоненты связываются между собой так, чтобы максимально использовать уже имеющуюся информацию о соединениях.

Данная работа является продолжением [7], где разработан метод линейной декомпозиции автомата с использованием алгоритма деления покрытий по [6]. К сожалению, мало автоматов, позволяющих разложение в линейную сеть и к тому же эффективность использования такого метода слишком сильно зависит от начального покрытия.

2. Необходимые понятия и определения. В этой работе рассматриваются автоматы без выхода, поэтому определим конечный автомат как тройку:

$$A = (I, S, \delta),$$

где I - входной алфавит;

S - множество состояний;

$\delta : I \times S \rightarrow S$ - функция переходов.

Функцию δ можно расширить до $S \times I^*$, где I^* - множество всех непустых слов алфавита I .

Определим сеть автоматов $N = (U, I, \Phi)$:

1) $U = \{A_i \mid i \in J\} = \{1, \dots, n\}$ - множество компонентных автоматов, где $A_i = (S_i, I_i, \delta_i)$. Здесь $I_i = I \times I'_i$, где I'_i - внутренний входной алфавит.

2) $\Phi = \{\varphi_i \mid i \in J\}$ - множество функций соединения, где

$$\varphi_i : x A_j \rightarrow I'_i; \quad j \in J.$$

Сети N функционально можно сопоставить эквивалентный ей автомат A_N . Сеть N называется декомпозицией автомата A , если и только если у автомата A_N имеется подавтомат A_N , изоморфный автомату A [2].

Разбиение произвольного конечного множества $X = \{x_1, x_2, \dots, x_m\}$ на блоки $B_1^1, B_1^2, \dots, B_1^{m_i}$ обозначим через $\pi_i(x)$. Нулевое (поэлементное) разбиение обозначим через $\emptyset_x$, а единичное (одноблочное) через I_x . Сужением разбиения $\pi_i(x)$ на $X' \subset X$ будем называть разбиение $\bar{\pi}_i(x') \equiv \{B_i^\alpha \cap X' \mid B_i^\alpha \in \pi_i; B_i^\alpha \cap X' \neq \emptyset\}$.

Запись $x_i \equiv x_j(\pi)$ означает, что элементы x_i и x_j находятся в одном и том же блоке разбиения π .

Для любых разбиений $\pi_i(x)$ и $\pi_j(x)$ примем, что

$$\pi_i \cdot \pi_j \equiv \{B_i^\alpha \cap B_j^\beta \mid B_i^\alpha \in \pi_i; B_j^\beta \in \pi_j; B_i^\alpha \cap B_j^\beta \neq \emptyset\}$$

$$\pi_i + \pi_j \equiv \bigcap \{\pi_k(x) \mid \pi_k \geq \pi_i, \pi_j\}, \text{ где } \pi_i \leq \pi_j \Leftrightarrow \pi_i \cdot \pi_j = \pi_i.$$

Разбиение $\Pi(x)$ можно интерпретировать как информацию о базисном множестве X .

Определим энтропию для разбиения π_i :

$$H(\pi_i) \equiv - \sum_{\alpha=1}^{m_i} q(B_i^\alpha) \log_2 q(B_i^\alpha); \quad q(B) = \frac{|B|}{|X|}.$$

Примем, что условная энтропия $H(\pi_j / \pi_i) \equiv H(\pi_i \cdot \pi_j) - H(\pi_i)$.

Систему разбиений $P(x) = \{\pi_i(x) \mid i \in J\}$ назовем полной, если и только если $\bigcap_{i \in J} \pi_i = \emptyset_x$.

Любую пару разбиений $\rho_i(x) \in \langle \pi_{i1}(x), \pi_{i2}(x) \rangle$ будем называть каналом на X . Энтропия канала определяется как

$$H(\langle \pi_i, \pi_k \rangle) \equiv H(\pi_k / \pi_i).$$

Назовем канал $\rho_i = \langle \pi_i(s), \tau_i(s) \rangle$ на множестве состояний S автомата $A = (I, S, \delta)$ h -каналом, если и только если

$$(\forall x \in I) (s' \equiv s''(\tau_i) \Rightarrow \delta(s', x) \equiv \delta(s'', x)(\pi_i)).$$

h -каналы представляют собой гомоморфные образы δ_h : $I \times \tau(s) \rightarrow \Pi(s)$ функционального соответствия $\delta: I \times S \rightarrow S[I]$.

Для любого $\pi_i \in P(s)$ существует максимальное разбиение $M(\pi_i) = \sum \tau_j$, образующее h -канал с разбиением π_i . $H(\langle \pi_i(s), M(\pi_i(s)) \rangle)$ определяет минимальное количество информации, необходимое для A_i от компонентов сети $N[3]$.

Аналогично, для любого $\pi_i \in P(s)$ существует минимальное разбиение $m(\pi_i) = \prod \tau_j$, так что $\langle \tau_j, \pi_i \rangle$ образуют h -канал. $H(\langle m(\pi_i(s)), \pi_i(s) \rangle)$ определяет максимальное количество информации, получаемой сетью N от компонента A_i [1].

Структурной матрицей сети N называется матрица m_{ij} при $i, j \in J$, определяемая следующим образом:

$$m_{ij} = \begin{cases} 1 & \text{если } j \in J_i \\ \emptyset & \text{если } j \notin J_i. \end{cases}$$

Содержательно $m_{ij} = 1$ интерпретируется как соединение между компонентными автоматами A_i и A_j . Матрице $\|m_{ij}\|$ с минимальным числом единиц соответствует связь N с минимальным числом связей [2].

3. Деление разбиений. Деление разбиений является удобным средством для определения информации, требуемой для функционирования компонента, от других компонентов сети. Если канал $\langle \pi, M(\pi) \rangle$ определяет полностью необходимую информацию, тогда $\pi' \leq \frac{M(\pi)}{\pi}$ исключает из результата информацию, которая имеется на выходе данного компонента и может быть получена компонентом через локальную обратную связь. Метод деления по алгоритму нахождения фактор-множества толерантности выдает в общем случае результатом покрытие [6].

Учитывая то, что частное определяется неоднозначно, в покрытии сохраняются варианты выбора всевозможных разбиений, которые могут быть частными, однако в дальнейших итерациях такая лишняя информация скапливается и является фак-

тором, мешающим выбору. Более подходящим следует, видимо, считать метод деления, где используется понятие сужения.

4. Алгоритм нахождения частного $\pi_y \leq \frac{\tau}{\pi_0}$. Для каждого блока $B_0^i \subset \pi_0$ находим всевозможные сужения на блоках разбиений τ .

$$\bar{\tau}(B_0^i) \triangleq \{B_\tau^j \cap B_0^i \mid B_\tau^j \in \tau\}.$$

Полученная информация является основой для нахождения частного с учетом следующих ограничений:

1) блоки в одном сужении не могут быть подмножествами одного и того же блока результата;

2) чтобы результат являлся разбиением, необходимо использовать все блоки из сужения только один раз.

Приведем метод, позволяющий вычислить π_y с минимальной энтропией по сравнению с другими разбиениями, удовлетворяющими вышеприведенным ограничениям.

Первый блок B_y^1 образуем так, что выберем из каждого $(\bar{\tau}(B_0^i) \mid \bar{\tau}(B_0^i) \neq \emptyset)$ блок с наибольшей мощностью, и соединим их, то есть

$$B_y^1 = \bigcup_{\alpha=1}^m B_0^\alpha,$$

где $B_0^\alpha \in \bar{\tau}(B_0^i)$ и $\forall B \in \bar{\tau}(B_0^i), |B^\alpha| \geq |B|$.

Использованные блоки удалим из сужений и повторим действие для нахождения следующих блоков для π_y пока $\forall \bar{\tau}(B_0^i) = \emptyset$.

5. Итеративный алгоритм нахождения декомпозиций. Предлагаемый в дальнейшем алгоритм позволяет декомпозировать исходный автомат A в совокупность h -каналов. Успешность алгоритма зависит от эффективности преобразований информации h -каналами. Если $H(m(\pi))$ и $H(M(\pi))$ сравнимые, то решение задачи почти тривиально.

6. Определения к алгоритму

$|\pi|_B$ — число блоков в разбиении π ;

$P_j\{\pi\}$ — некоторое произведение из j разбиений (без повторения), входящих в множество $\{\pi\}$, $j \leq |\{\pi\}|$;

$n_{\max}$ — заданное ограничение на число состояний в компонентных автоматах.

Декомпозируем $A = (I, S, \delta)$ на компоненты $A_i = (S_i, I_i, \delta_i)$,

где $|S_i| = |\Pi_i|_B \leq n_{\max}$.

1) $i = 1$. Возьмем начальное разбиение $\Pi_1, |\Pi_1|_B \leq n_{\max}$.

2) Если $\bigcap_{j=1}^i \Pi_j = \emptyset_S$, то конец алгоритма.

3) Образует канал $\langle \Pi_i, M(\Pi_i) \rangle$. Таким способом по [3] определим минимальное количество информации, необходимое для компонента A_i от сети N .

4) Если $|M(\Pi_i)|_B \leq n_{\max}$, то $\Pi_{i+1} = M(\Pi_i)$, $i = i + 1$, продолжить от шага 2.

5) Попробуем удовлетворить потребность компонента A_i в информации, используя продуцированную им информацию, то есть введем локальную обратную связь. Для этого найдем решение для уравнения $\Pi_i \cdot \Pi_{i+1} \leq M(\Pi_i)$, откуда $\Pi_{i+1} \leq \frac{M(\Pi_i)}{\Pi_i}$.

Если $|\Pi_{i+1}|_B \leq n_{\max}$, то $i = i + 1$, идти к шагу 2.

6) Иначе, если i -й компонентный автомат требует дополнительной информации, решим систему

$$\begin{cases} \Pi_{i+1} \leq \frac{M(\Pi_i)}{P_j\{\Pi\}} \\ |\Pi_{i+1}|_B \leq n_{\max} \end{cases}$$

Если имеется решение, то компоненту требуется информация от предстоящих компонентов, определяемая $P_j\{\Pi\}$, $i = i + 1$, идти к шагу 2.

7) Разделение. Использование информации от предыдущих компонентов не позволяет найти решение, удовлетворяющее второму уравнению системы. Для решения надо разделить Π_{i+1} на части $\Pi_{i+1}^1 \dots \Pi_{i+1}^k$ так, чтобы

$$\bigcap_{j=1}^k \Pi_{i+1}^j \leq \Pi_{i+1} \text{ и } \forall |\Pi_{i+1}^j|_B \leq n_{\max}.$$

Далее находим решение для всех ветвей. $i = i + 1$, идти к шагу 2.

При таком разделении на части задача декомпозиции затрудняется и это, с одной стороны, показывает, что автомат сильно связанный, а с другой стороны, возможно, было неудачно выбрано начальное разбиение.

Примечание 1. Если на некотором шаге i получим, что $M(\pi_i) \geq P_j \{ \pi \}$, то найдем СП-компонент $[I]$. Если все-таки $\bigcap_{j=1}^l \pi_j > \emptyset_S$, то надо выбрать новое начальное разбиение, исходя из уже полученных компонентов $\pi_{i+1} \geq \frac{\emptyset_S}{\bigcap_{j=1}^l \pi_j}$.

Примечание 2. Если у полученного компонента $|\pi_i|_B < n_{\max}$, то можно его заменить расширенным компонентом π'_i , потребность в информации у которого такая же, найденная следующим образом:

$$\begin{cases} m(M(\pi_i)) \leq \pi'_i \leq \pi_i \\ |\pi'_i|_B \leq n_{\max}. \end{cases}$$

Возможность такого расширения обосновывается в [1].

Пример. Пусть задан конечный автомат $A = (I, S, \delta)$ и $n_{\max} = 2$.

$\delta \backslash I$	a	b	c	d
1	2	7	3	6
2	2	5	1	6
3	7	6	3	3
4	5	6	3	3
5	1	4	8	6
6	1	2	6	6
7	8	1	3	3
8	6	1	3	3

Пусть $\pi_1 = \{ \overline{1,2,3,4}; \overline{5,6,7,8} \}$, тогда $M(\pi_1) = \{ \overline{1,2}; \overline{3,4}; \overline{5,6}; \overline{7,8} \}$. $|M(\pi_1)|_B > n_{\max}$

$$\frac{M(\pi_1)}{\pi_1} = \{ \overline{1,2,5,6}; \overline{3,4,7,8} \}. \text{ Из } \left| \frac{M(\pi_1)}{\pi_1} \right|_B = n_{\max} \text{ следует}$$

что это частное можно выбрать в качестве разбиения π_2 .

$$M(\pi_2) = \{ \overline{1,5}; \overline{2,6}; \overline{3,7}; \overline{4,8} \}. |M(\pi_2)|_B > n_{\max}$$

$$\frac{M(\pi_2)}{\pi_2} = \{ \overline{1,5,3,7}; \overline{2,6,4,8} \} \text{ и выбираем его в качестве раз-}$$

биения Π_3 . $M(\Pi_3) = \{1,2; 3,4; 5,6; 7,8\} = \pi_1 \cdot \pi_2 \mid \prod_{i=1}^3 \pi_i =$
 $= \emptyset_s$, то есть автомат разложен в сети из трех компонентных автоматов.

Структурная матрица сети:

$$\|m_{ij}\| = \begin{vmatrix} I & I & \emptyset \\ \emptyset & I & I \\ I & I & \emptyset \end{vmatrix}$$

Заключение. Разработан метод итеративной декомпозиции конечных автоматов. Полученная сеть из компонентных автоматов не имеет заданной структуры, последняя возникает в процессе декомпозиции по информационным каналам. Проведенные эксперименты показывают эффективность метода при декомпозиции мало- и среднесвязанных автоматов. Дальнейшие исследования сосредоточены на выборе более эффективных начальных разбиений для декомпозиции.

Л и т е р а т у р а

1. H a r t m a n i s J., S t e a r n s R.E. Algebraic structure theory of sequential machines. Englewood Cliffs, New York: Prentice-Hall. Inc., 1966.

2. К е з в а л л и к А.Э. Теорема декомпозиции конечных автоматов // Автоматика и вычислительная техника.-1974. № 1.-С. 17-24.

3. К е з в а л л и к А.Э., Л а у с м а а Т. Декомпозиция конечных автоматов на основе информационного критерия // Тр. Таллинск. политехн. ин-та.-1985.- № 601.- С. 3-II.

4. F i s m a n G.S. Synergy in machines // Journal of Pure and Applied Algebra. North-Holland Publishing Company. - 1983. - Vol. 28. - P. 155-177.

5. G e l l e r D.P. Realization with feedback encoding I: Analogies of the classical theory // SIAM J. Comput. March, 1975. - Vol. 4, N 1.

6. Z e c h K.A. Feedback-orientierte Kodierung endlicher Automaten // Elektronische Information - Verarbeitung und Kybernetik. - 1973. 9, - N 3. - S. 632-650.

7. Лейс П., Салум К. Итеративный метод декомпозиции конечных автоматов // Тр. Таллинск. политехн. ин-та. - 1983.- № 550. - С. 81-89.

K. Tammemäe

Iterative Decomposition Method of Finite Automaton

Abstract

An iterative decomposition method of finite automaton is presented in this paper. Structure of network realising the source automaton is generated in the process of iteration. Active use of possibilities of feedback loops makes it possible to reduce the steps of iteration and the summary number of network states.

K. Tammemäe

Tükelduste süsteemi valik abstraktsete automaatide iteratiivsel dekompositsioonil

Kokkuvõte

Artiklis käsitletakse abstraktsete automaatide iteratiivse dekompositsiooni üht meetodit. Koostatud on algoritm täieliku түкelduste süsteemi valikuks, mis lubab lähteautomaadi realiseerida võrguna. Tүкelduste valikul kasutatakse jagamistehet ja infokanalite mõistet. Dekompositsiooni käigus arvestatakse aktiivselt tagasiside võimalusi, minimeerimaks võrgu summaarset olekute arvu. On määratletud antud meetodi abil eriti efektiivselt dekomponeeritavate automaatide omadused.

ФУНКЦИОНАЛЬНЫЙ ПОДХОД К СОЗДАНИЮ МАТЕМАТИЧЕСКИХ МОДЕЛЕЙ ДИСКРЕТНЫХ ОБЪЕКТОВ НА ОБОБЩЕННЫХ АЛЬТЕРНАТИВНЫХ ГРАФАХ

Функциональными методами построения математических моделей для объектов диагностирования (ОД) называются такие методы, которые исходят из описания внешнего поведения ОД, а его внутренняя структура при этом считается неизвестной или чисто гипотетической. Эти методы являются незаменимыми в тех случаях, когда внутренняя структура ОД неизвестна или когда ее описание является слишком объемным. Особую важность приобретают такие методы также на начальных этапах проектирования дискретных объектов, например, с целью получения математической модели для генерации верифицирующих тестов для ожидаемых структурных решений.

Вопросы, связанные с синтезом математических моделей на бинарных альтернативных графах (БАГ), основательно рассмотрены в [1]. Методы автоматического синтеза моделей ДО на обобщенных альтернативных графах до сих пор отсутствуют.

Модели на обобщенных альтернативных графах (ОАГ) представляют объект через функциональные связи между его переменными. Эти модели, по сравнению с моделями на БАГ, позволяют работать на более высоком уровне абстракции и, как правило, намного компактнее. Соответственно уменьшается и трудоемкость вычислений над ними.

Каждое ОАГ моделирует поведение какой-нибудь переменной ОД, которая называется его главной переменной. ОАГ представим в виде направленного графа с одной исходной вершиной. К каждой вершине приписана какая-нибудь весовая функция. Переменные, являющиеся аргументами весовых функций в пределах одного графа, называются переменными графа.

Вершины, которые имеют выходящие из них дуги, являются управляющими. Значение весовой функции на управляющей вершине определяет дальнейший путь движения по графу.

Вершины, из которых пути не выходят, называются терминальными. Процесс моделирования можно представить как движение по графу. При достижении терминальной вершины значение его весовой функции присваивается главной переменной данного графа.

Объекты, требующие вычисления нескольких переменных, представляются в виде системы ОАГ. Семантика моделей на системе ОАГ рассмотрена в [2]. Приведем некоторые характерные свойства, на которые будем ссылаться в дальнейшем:

1) весовые функции управляющих вершин должны быть дискретные, для весовых функций при терминальных вершинах требуется лишь их определенность при всех встречаемых значениях аргументов, кроме неизвестного значения;

2) из любой управляющей вершины может выходить любое конечное число дуг;

3) в качестве переменных графа можно рассматривать как разные непересекающиеся подмножества реальных булевых переменных объекта, так и переменные более абстрактной природы (например, переменные времени);

4) с целью отражения неопределенности главной переменной при заданных условиях в качестве весовой функции при соответствующей терминальной вершине используется специальная константа (обозначение в дальнейшем — " $*$ ").

Попытаемся сформулировать класс объектов, которые могут описываться моделью ОАГ. Введем следующие конечные подмножества переменных, которые могут присутствовать в системе ОАГ:

I — переменные, присутствующие только как аргументы весовых функций;

O — переменные, присутствующие только как главные переменные;

S — главные переменные, являющиеся также аргументами весовых функций.

Обозначениями соответствующих алфавитов будут $A(I)$, $A(O)$ и

$A(S)$ соответственно. Тогда графами можно определить отображения в виде:

$$\sigma : A(I) \times A(S) \rightarrow A(S) \quad (1)$$

$$\lambda : A(I) \times A(S) \rightarrow A(O), \quad (2)$$

и ОД определен пятеркой:

$$ОД = \{ I, S, O, \sigma, \lambda \}. \quad (3)$$

(3), (1), (2), вместе взятые, являются каноническим представлением конечного автомата.

Следствие 1. Все конечные автоматы представимы конечной системой ОАГ.

Следствие 2. Учитывая свойство I и то, что аппаратура ОАГ сама по себе не требует дискретности аргументов весовых функций (переменных графа), можно сказать, что данная математическая модель позволяет описывать и некоторые объекты, операционная часть которых является недискретной.

Пример 1. Рассмотрим в качестве ОД синхронный пиковый детектор со сбросом (рис. 1а). ОАГ, моделирующий его аналоговый выход, приведен на рис. 1б.

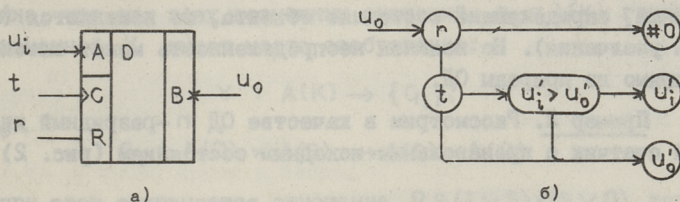


Рис. 1.

Эффективность представления ОД на модели ОАГ во многом зависит от природы ОД, от допускаемого уровня абстракции и, в некоторой степени, от формы исходного описания ОД.

Исходим из представления ОД в виде конечного автомата (3). Так как при его моделировании требуется вычисление всех переменных

$$z_k \in S \cup O,$$

а графы для переменных из этих двух множеств принципиально не отличаются, мы можем в дальнейшем не различать переходные и выходные функции друг от друга. Введем обобщенное преобразование

$$G : A(I) \times A(S) \rightarrow A(S) \times A(O). \quad (4)$$

В общем случае это преобразование определено не полностью. Можно сказать, что для каждой отдельной переменной z_k из S или O это преобразование выражается через функции от n_k аргументов

$$\varphi_k(\underline{z}_k) \quad \text{и} \quad f_k(\underline{z}_k), \quad (5)$$

так, что

$$z_k = g(z_k) = \begin{cases} f_k(\underline{z}_k), & \text{если } \varphi_k(\underline{z}_k) = 0 \\ " * " , & \text{если } \varphi_k(\underline{z}_k) \neq 0 \end{cases} \quad (6)$$

и

$$\underline{z}_k = \{z_{k,1}, \dots, z_{k,n}\} \in I \cup S.$$

При этом $f_k()$ называем функцией неопределенности для переменной z_k . $\varphi_k()$ может задаваться как явно, так и неявно. Примером задания $\varphi_k()$ в явном виде может быть описание R-S триггера. Неявная неопределенность встречается в случаях, когда в исходном описании поведение объекта в некоторых ситуациях не определяется (считается безразличной). При этом существует принцип, что в этих ситуациях внутренние переменные, определяющие состояние объекта, не изменяются (принцип умолчания). Но неявная неопределенность может вытекать и прямо из природы ОД.

Пример 2. Рассмотрим в качестве ОД n -разрядный двоичный счетчик с произвольным исходным состоянием (рис. 2).

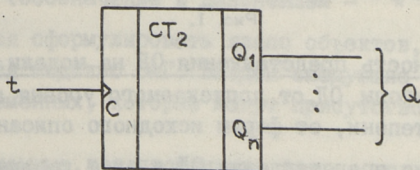


Рис. 2.

Пусть Q — его обобщенная выходная переменная и t — абстрактная переменная, такая, что значению $t = 1$ соответствует попадание тактового импульса на вход счетчика, а значению 0 — отсутствие импульса. При этом

$$f(t, Q) = t((Q' + 1) \oplus 2^n) \vee t Q',$$

а функции неопределенности на базе заданных переменных явно задавать не удастся. Получаемые ОАГ для данного объекта могут иметь вид:

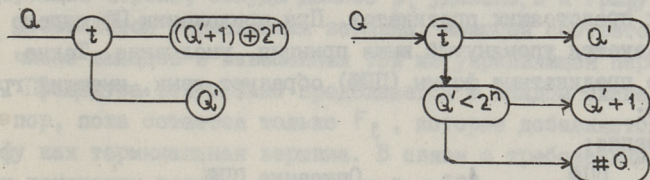


Рис. 3.

Процесс синтеза тестов на модели ОАГ становится тем эффективнее, чем меньше переменных, присутствующих одновременно в управляющих и функциональных вершинах графов, т.е. чем лучше можно разделить функции управляющей и операционной части ОД.

Введем разбиение входных переменных ОД на два непересекающихся подмножества:

$$I = K \cup D : K \neq \emptyset \quad (7)$$

такие, что для всех возможных значений $c_i \in A(K)$ групповой переменной K имеет место отображение:

$$\gamma : A(K) \rightarrow \{Q_i\}, \quad (8)$$

где

$$Q_i : A(D) \times A(S) \rightarrow A(S) \times A(O) \quad (9)$$

и при этом сохраняется отношение $R \subseteq (I \cup S) \times (S \cup O)$, заданное отображением (4).

Определение 3. Если существует разбиение (7) при вышеуказанных условиях, то значит, что ОД является регулярным в функциональном смысле.

Групповую переменную K будем называть управляющей переменной, а переменные из D — операндами. Довольно распространенной формой представления для отображения (8) является его запись системой т.н. предикатных форм (ПФ). С целью уменьшения объема описания часто является целесообразным дальнейшее разнесение управляющих переменных. Предикатные формы тогда будут иметь вид:

$$(\dots, p_i, \dots) (\dots, z_k = g_{k,i}(z_{k,i}), \dots),$$

где предикатом p обычно служит сравнение управляющей переменной с константой и условием выполнения присваивания значения переменной. Z_k считается одновременная истинность всех предстоящих предикатов. При приложении ПФ широко используется упомянутый выше принцип умолчания. Такие простые предикатные формы (ППФ) образуют язык, имеющий грамматику:

Описание:

ППФ		Описание ППФ	(10)
;			
ППФ:			
Присваивания			(10)
	Предикат	ППФ	
;			

Как видно из продукции (10), допускаются и безусловные действия с некоторыми переменными объекта. Язык ППФ неудобен для описания сложных объектов, но он является часто применяемой формой промежуточного представления при синтезе математической модели.

Процесс построения САГ по описанию ОД на языке ППФ завершается в два этапа.

На первом этапе все предикаты преобразуются в одинаковую форму в виде пар:

$$p_i = (x_i, r_i), \quad (11)$$

где x_i — управляющая переменная;

r_i — значение переменной,

при которой предикат истинный. В дальнейшем эти значения будем называть решениями. После этого подставим во все переменные Z_k , для которых имеет место присваивание нового значения, в соответствии с блоком состоящим из строк в виде:

$$(\dots, p_{i,j}, \dots) F_1. \quad (12)$$

где F_1 — значение, присваиваемое Z_k при данных условиях.

На втором этапе ведется построение ОАГ отдельно для всех переменных Z_k с ранжированием управляющих переменных с целью минимизации числа управляющих вершин на графе. Процедуры, выполняющие первый этап, тривиальны и здесь не рассматриваются. При построении графов ведется разбиение мно-

жества строк (I_2) в текущем блоке по значениям некоторой управляющей переменной. При этом образуются новые блоки, содержащие строки, откуда данное p_i удалено, и к графу для Z_k добавляется управляющая вершина, имеющая соответствующее число выходов и взвешенная той же управляющей переменной. Процедура рекурсивно продолжается в каждом блоке до тех пор, пока остается только F_l , которое добавляется к графу как терминальная вершина. В связи с требованиями методики генерации тестов, приведенной в [3], идентичные подграфы не выделяются, а идентичные терминальные вершины в границах одного графа не дублируются.

Введем следующие обозначения:

X — множество переменных x_i , предикаты которых присутствуют во всех строках данного блока;

A_i — алфавит для x_i ;

B_i — множество всех решений r_i в пределах блока;

$\{K_j\}$ — множество всех возможных клик.

Кликкой называется подмножество всех строк в блоке, которые после удаления из них предикатов по $x_i \in X$

а) содержат предикаты по одной и той же переменной, или

б) не содержат больше предикатов и состоят из идентичных присваиваний;

NK — количество клик;

NL — число строк в блоке.

Предлагается алгоритм вычисления оценки эффективности разбиения по x_i . Используя данную оценку при выборе переменной для очередной управляющей вершины создаваемого графа, число возникающих подграфов будет минимальным.

Алгоритм I

1) для $j: 1$ до NK взять множество $C_j := B_i$;

2) $l := 0$; $h := 0$;

3) если $l \leq NL$; идти к 4, иначе идти к 7;

4) найти j такое, что $l \in K_j$;

5) удалить $r_{i,l}$ из C_j ;

6) $l := l + 1$; идти к 3;

7) для $j := 1$ до NK взять $h := h + |C_j| * |K_j|$;

8) возврат.

Предложенный алгоритм основывается на предположении, что разбиение по x_i тем эффективнее, чем сильнее корреляция между значениями x_i и переходами между кликами. Построение ОАГ можно провести по следующему алгоритму:

Алгоритм 2

- 1) удалить все кратные строки;
- 2) найти $NL, X, \{K_j\}, NK$;
- 3) $i := 1$; $max := 0$; $m := 0$; $NI := X$;
- 4) если $NI = 0$, то идти к I3, иначе идти к 5;
- 5) если $i \leq NI$, то идти к 6, иначе идти к I2;
- 6) найти B_i ; вызвать алгоритм I;
- 7) если $h > 0$ или $B_i \neq A_i$ идти к 9, иначе идти к 8;
- 8) удалить все предикаты по x_i ; идти к 2;
- 9) если $h < max$ идти к II, иначе идти к I0;
- I0) $max := h$; $m := i$;
- II) $i := i + 1$; идти к 5;
- I2) если $m = 0$ идти к I3, иначе идти к I4;
- I3) добавить в граф вершину F_1 ; возврат;
- I4) добавить в граф вершины по x_m ;
- I5) удалить все предикаты по x_m ;
- I6) рекурсия по всем разбиениям блока по B_m ;
- I7) возврат.

Работа приведенных алгоритмов иллюстрируется на следующем примере.

Пример 3. Представим исходный блок для какой-нибудь переменной Z_k в виде таблицы на рис. 4а. Здесь a, b, c и d — управляющие переменные, а номера в таблицах — номера уникальных решений для соответствующих предикатов. Отсутствие номера означает отсутствие предиката в данной строке. Идентичные присваивания обозначены буквами А, В.

При выполнении алгоритмов I и 2 для исходного блока получим:

a)	ℓ	a	b	c	d	F_ℓ
	1	0	0	0	0	A
	2	0	1	1		B
	3	1	0	0		B
	4	1	1	1		A
	5	2	0			A
	6	2	1			A
	7	3	0			B
	8	3	1	1		B

b)	b	c	d	
I	0	0	0	A
	1	1		B
II	0	0		B
	1	1		A
III	0			A
	1			A
IV	0			B
	1	1		B

Рис. 4.

$$X = \{a, b\};$$

$$\{K_j\} = \{(1, 2, 3, 4), (1, 8), (5, 6), (7)\};$$

$$h_a = 2*4 + 2*2 + 2*2 + 3*1 = 19,$$

$$h_b = 0*4 + 0*2 + 1*2 + 1*1 = 3.$$

Первое разбиение ведется по переменной a . Новые блоки, получаемые при этом, показаны на рис. 4б. Модель, получаемая для z_k , показана на рис. 5.

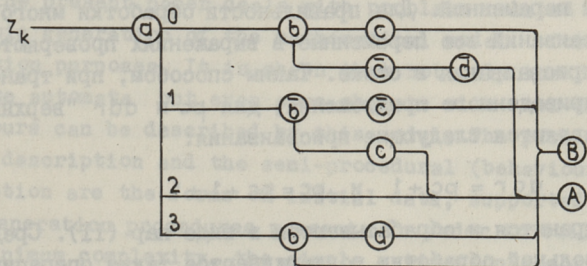


Рис. 5.

Полупроцедурная форма описания широко применяется для представления сложных объектов, при которых ясно выражается их функциональная регулярность (например, микропроцессорные ВИС и системы). По структуре основных языковых кон-

струкций эта форма близка к предикатным формам, но последовательность операций в ней можно интерпретировать во времени. На кафедре ЭВМ ТПИ разработан и реализован язык такого типа для описания широкого класса сложных ОД. В языке имеются средства для введения и обработки макроопределений, благодаря чему даже весьма сложные объекты можно описывать довольно компактно и наглядно. В качестве примера приведем фрагмент из описания микропроцессора Intel-8080:

```
if(comm == ADI){pcgo; pcgo; res=in+a; a=res; flags(res);} ,
```

где `pcgo` и `flags` есть макроопределения:

```
pcgo() {adr=pc; pc=pc+1;}
```

```
flags(x){if(x==0) f_z=1; else f_z=0; f_s=x&080h;...}.
```

При трансляции предложений языка все найденные макро-вызовы расширяются примерно так же, как при работе макро-ассемблера. Так как за каждым условием может стоять множество действий и других условий, имеет смысл говорить о путях трансляции. Для того, чтобы результатом первичной трансляции были блоки ПФ, при каждом условии транслятор рекурсивно генерирует и обрабатывает все возможные пути дальнейшей трансляции. Каждое действие или условие на пути ставят в стек. Ветвь рекурсии заканчивается, когда путь кончается или находится условие, противоречащее уже имеющемуся в стеке. Тогда начинается дополнение блоков ПФ. При этом из стека берется только последнее условие и (или) присваивание по каждой переменной. Для правильности обработки многократных присваиваний все переменные в выражениях проверяются на наличие присваивания в стеке. Таким способом, при трансляции вышеприведенного предложения, для `pc` и `adr` "верхними" в стеке являются следующие присваивания:

$adr = pc + 1$ и $pc = pc + 1 + 1$.

Условия хранятся и обрабатываются в виде пар (II). Средствами символьной обработки поддерживаются такие операции, как нахождение взаимных перекрытий, слияние и инвертирование условий. Построение системы ОАГ ведется по алгоритмам I и 2.

Л и т е р а т у р а

1. У б а р Р.Р. Тестовая диагностика цифровых устройств. Учебное пособие.- Ч. I. - Таллин, 1980.- III с.

2. У б а р Р.Р. Обобщенная модель альтернативных графов для синтеза тестов цифровых схем // Тр. Таллинск. политехн. ин-та. - 1983. - № 550. - С. 97-109.

3. К о н ь т Т.К. Подход к автоматизации построения тестов для микропроцессорных БИС // См. наст. сб., с. 55-64.

V. Alango

Generalized Alternative Graph Model Synthesis for Numerical Objects.

Functional Approach

Abstract

The present paper deals with problems related to the automatic generation of the mathematical model for test generation purposes. It is shown that not only the finite discrete automata but even some objects with non-discrete behaviours can be described by this model. The pure functional description and the semi-procedural (behavioural) description are the forms of initial data, supported by graph generation procedures present. In order to build graphs with minimum complexity, the simple estimation algorithm is proposed.

Diskreetsete objektide matemaatiliste mudelite saamine
üldistatud alternatiivsetel graafidel.

Funktsionaalne lähenemisviis

Kokkuvõte

Artiklis käsitletakse küsimusi, mis on seotud keerukate numbriliste objektide esitamisega üldistatud alternatiivsete graafide abil. Näidatakse, et antud mudel lubab esitada mitte ainult diskreetseid lõplikke automaate, vaid ka mõningaid osaliselt mittediskreetseid objekte. Esitatakse algoritmid graafide sünteesiks objekti funktsionaalsest kirjeldusest lähtudes ning graafi tippude optimaalse järjestyse leidmiseks.

ПОДХОД К АВТОМАТИЗАЦИИ ПОСТРОЕНИЯ ТЕСТОВ ДЛЯ МИКРОПРОЦЕССОРНЫХ БИС

1. Введение. Характерной чертой современной вычислительной техники является интенсивное применение микропроцессорных БИС. Высокая степень интеграции и ограниченное количество контрольных точек в устройствах, построенных на микропроцессорных наборах БИС делает невозможным применение традиционных методов генерирования тестов, основанных на вентильном представлении контролируемых объектов. Это приводит к использованию интуитивных функциональных тестов, осуществляющих проверку каждой операции микропроцессорных БИС на некотором множестве операндов или же конкретных программ пользователей. Такие средства тестирования далеко не универсальны и не обеспечивают требуемой полноты покрытия неисправностей.

В данной статье рассматривается метод автоматизации построения структурированных функциональных тестов для микропроцессорных БИС и систем. Основывается он на модели обобщенных графов АГ [1, 2] и представляет единый способ описания и универсальные алгоритмы синтеза тестов для широкого класса дискретных объектов, включая микропроцессорные БИС и системы.

При этом учитываются некоторые особенности рассматриваемого класса объектов, в частности, возможности ручного задания установочных и транспортирующих последовательностей в виду их относительной простоты.

2. Модель объекта диагностирования. Модель объекта диагностирования (ОД) основывается на описании системы команд. Модель ОД состоит из совокупности АГ

$$G_i = (S_i, \Gamma_i, M_i), \quad M_i \in M,$$

описывающих поведение функциональных компонентов ОД (регистры, счетчики, АЛУ, мультиплексоры) на программных или микропрограммных уровнях представления. Для каждого компонента ОД строится свой частный тест. Здесь S_i - множество вершин АГ, взвешенных переменными $m \in M_i$, а M - множество всех переменных, используемых для описания ОД.

$$M = M_x \cup M_s \cup M_y \quad \text{где } M_x, M_s \text{ и } M_y$$

определяют соответственно подмножества входных, внутренних и выходных переменных. Отображение $\Gamma_i(s)$, $s \in S_i$ задает множество последователей вершины s .

Множество $S_i = S_i^H \cup S_i^T$, где $S_i^H = \{s \mid \Gamma_i(s) \neq \emptyset\}$ множество внутренних вершин, а $S_i^T = \{s \mid \Gamma_i(s) = \emptyset\}$ множество терминальных вершин.

3. Вершины АГ-модели ОД. Вершинам $s \in S_i$ сопоставлены веса $e(s)$, принимающие значения из их областей определения $A(e(s))$. Весами внутренних вершин $s^H \in S_i^H$ могут быть переменные $m \in M_s$, а весами терминальных вершин $s^T \in S_i^T$ переменные $m \in M_x \cup M_y$ и функции $f(M)$. Функции $f(M)$ описывают функциональные компоненты операционной части микропроцессора. Терминальные вершины, взвешенные функциями, называются функциональными вершинами F . Аргументами этих функций можно использовать переменные $R_i, R_j \in M_s$ и $R_n \in M$, где R_i и R_j регистровые переменные, а R_n - некоторая промежуточная переменная. Значение промежуточных переменных определяется с помощью отдельного АГ G_n .

Переменным $m \in M_x \cup M_y$ на терминальных вершинах сопоставлены подпрограммы ввода (установки) входных данных и подпрограммы вывода (транспортировки) выходных данных. На такой вершине происходит активизация соответствующей подпрограммы ввода-вывода. Такие вершины называются специальными вершинами. Весовые переменные внутренних вершин $s \in S_i^H$ разделяются на булевые переменные X_i , для которых $A(X_i) = \{0, 1\}$ и переменные управления I_i , для которых $A(I_i) = \{0, 1, \dots, n\}$.

Вершины с переменными X_i называются бинарными вершинами и имеют два выхода, причем одному из них приписано значение 1, а другому значение 0. X вершины можно рассматривать как триггера условий.

Вершины с переменными I_j называются вершинами управления и имеют n выходов. I вершины можно рассматривать как дешифраторы команд. Совокупность управляющих и бинарных вершин описывает управляющий автомат ОД. Совокупность функциональных вершин описывает операционный автомат ОД. Промежуточными переменными на терминальных вершинах описываются мультиплексорные функции.

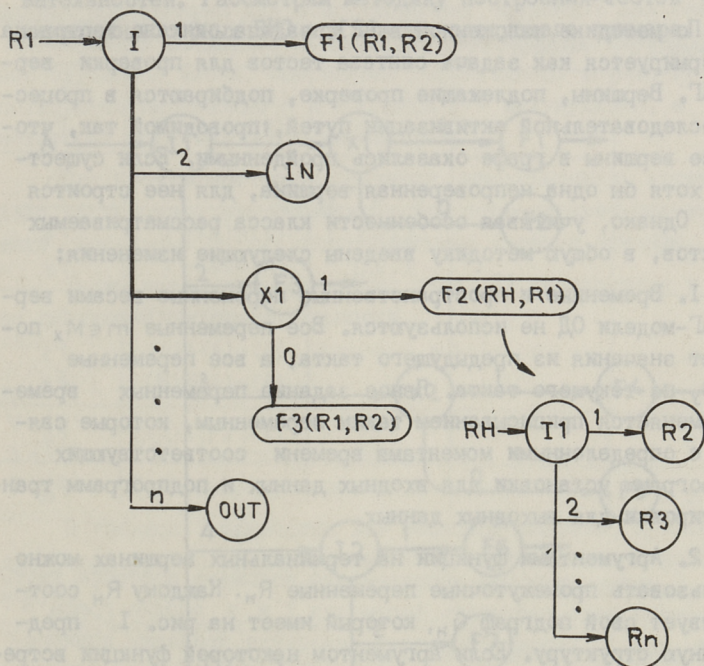


Рис. 1.

На рис. 1 представлена АГ-модель, описывающая поведение регистра R_1 некоторого дискретного объекта. Здесь I - регистр команд с дешифратором, X_1 - триггер условий, F_j - функции АЛУ. IN и OUT соответственно подпрограммы установки входных данных в ДО транспортировки результатов на выход ДО. Промежуточная графа G_n описывает устройство адресации ДО. Здесь I_1 - дешифратор режима адресации, а переменные R_j образуют блок регистров общего назначения.

4. Методика построения тестов. Генерирование тестов для микропроцессорных БИС включает следующие этапы:

1) синтез установочного сегмента исходных данных, синтез тестового сегмента для проведения проверяемой операции, синтез транспортировочного сегмента для транспортировки результатов теста к выходу объекта;

2) синтез таких операндов, чтобы удовлетворялись условия для обнаружения заданных дефектов.

По методике, изложенной в [2], задача синтеза тестов на АГ формируется как задача синтеза тестов для проверки вершин АГ. Вершины, подлежащие проверке, подбираются в процессе последовательной активизации путей, проводимой так, чтобы все вершины в графе оказались пройденными. Если существует хотя бы одна непроверенная вершина, для нее строится тест. Однако, учитывая особенности класса рассматриваемых объектов, в общую методику введены следующие изменения:

1. Временные и пространственные переменные весами вершин АГ-модели ОД не используются. Все переменные $m \in M_x$ получают значения из предыдущего такта, а все переменные $m \in M_y$ из текущего такта. Явное задание переменных времени заменяется приписыванием таким переменным, которые связаны с определенными моментами времени соответствующих подпрограмм установки для входных данных и подпрограмм транспортировки для выходных данных.

2. Аргументами функции на терминальных вершинах можно использовать промежуточные переменные R_n . Каждому R_n соответствует свой подграф G_n , который имеет на рис. 1 предложенную структуру. Если аргументом некоторой функции встречается промежуточная переменная R_n , то необходимо выйти из этой вершины в подграф G_n , задающий условия, при выполнении которых R_n принимает конкретное значение. Каждая промежуточная переменная требует своего теста.

3. Тестирование вершин, которые взвешенные переменными типа X, производится одновременно с вершинами типа I.

4. На одном и том же пути переменные повторяться не могут.

5. Переменные вершин на путях графа имеют следующее упорядочение: I, X, F.

6. Синтез установочных и транспортировочных последовательностей на АГ-модели не происходит. Поскольку нахождение этих последовательностей для внутренних переменных достаточно просто, то они подготавливаются заранее вручную в виде подпрограмм.

7. Процесс синтеза теста производится на основе только проверяемого графа. Отпадают рекурсивные обращения из графов в графы при поиске подходящих установочных последовательностей. Рассмотрим методику построения тестов для микропроцессорных БИС при помощи следующего примера.

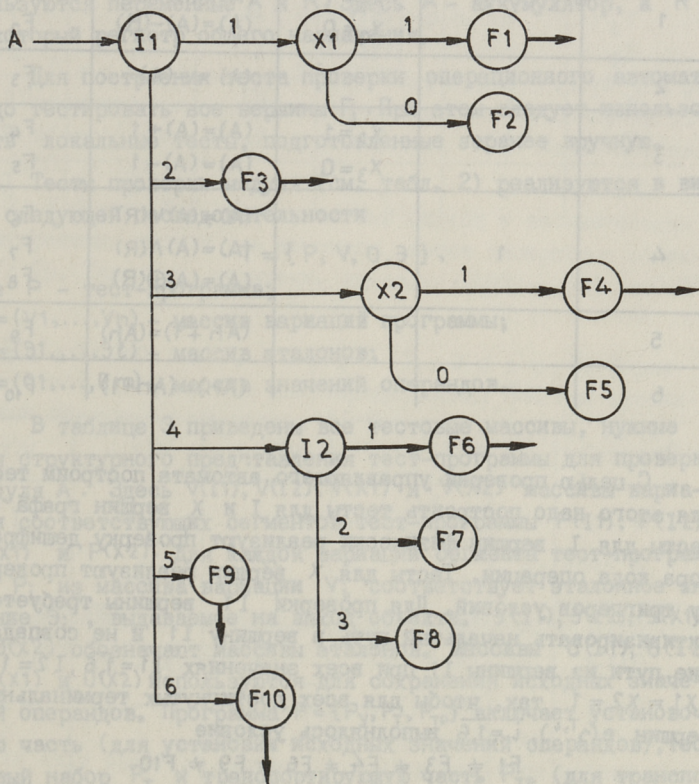


Рис. 2.

8. Пример. Рассмотрим модель AG , представленную на рис. 2 и описывающую поведение регистра A в некотором гипотетическом микропроцессоре при выполнении некоторого подмножества команд (см. табл. I). Здесь вершины I_1, I_2 , взвешенные переменными команд, описывают управляющий автомат микропроцессора, вершины X означают условия и F_j описывают функции операционного автомата микропроцессора.

Т а б л и ц а I

Команда I_1	Команда I_2	Условие	Операция	Функция
1		$X_1 = 1$ $X_2 = 0$	$(A) = (A) + (R)$ $(A) = (A) - (R)$	F_1 F_2
2			$(A) = -(A)$	F_3
3		$X_2 = 1$ $X_3 = 0$	$(A) = (A) + 1$ $(A) = (A) - 1$	F_4 F_5
4	1		$(A) = (A) \vee (R)$ $(A) = (A) \wedge (R)$ $(A) = (A) \oplus (R)$	F_6 F_7 F_8
5			$(Ah + 1) = (Ah)$	F_9
6			$(An) = (An + 1)$	F_{10}

С целью проверки управляющего автомата построим тест. Для этого надо построить тесты для I и X вершин графа A . Тесты для I вершин физически реализуют проверку дешифратора кода операции. Тесты для X вершин реализуют проверку триггеров условий. Для проверки I_1 вершины требуется активизировать начальный путь в вершину I_1 и не совпадающие пути из вершины I_1 при всех значениях $I_1 = \overline{1,6}, I_2 = 1$ и $X_1 = X_2 = 1$ так, чтобы для всех достигнутых терминальных вершин $e(s^{T,i})$, $i = \overline{1,6}$ выполнялось условие

$$F_1 \neq F_3 \neq F_4 \neq F_6 \neq F_9 \neq F_{10}.$$

Для проверки I_2 вершины необходимо активизировать начальный путь в вершину I_2 и все несовпадающие пути из вершины I_2 . После этого решать неравенство $F_6 \neq F_7 \neq F_8$.

Для проверки вершин x_1 и x_2 необходимо активизировать дополнительные пути $x_1 = x_2 = 0$ и решить неравенства соответственно $F_1 \neq F_2$ и $F_4 \neq F_5$.

Полученный тест приведен в таблице 2. В последней строке таблицы приведена структурированная форма теста, где D обозначает изменяющуюся часть набора. Тест выполняется циклически для всех значений I_1, I_2, x_1 и x_2 так, чтобы при варьировании значений одной переменной значения других оказались постоянными для всех выбранных операндов для каждого F_j . Значения операндов для F_j получаются путем генерирования случайных чисел. Операндами терминальных вершин F_j используются переменные A и R . Здесь A — аккумулятор, а R — некоторый регистр общего назначения.

Для построения теста проверки операционного автомата надо тестировать все вершины F_j . При этом следует использовать локальные тесты, подготовленные заранее вручную.

Тесты проверки модуля (см. табл. 2) реализуются в виде следующей последовательности

$$T = \{ P, V, O, \mathcal{E} \},$$

где P — тест-программа;

$V = (V_1, \dots, V_n)$ — массив вариаций программы;

$\mathcal{E} = (\mathcal{E}_1, \dots, \mathcal{E}_l)$ — массив эталонов;

$O = (O_1, \dots, O_m)$ — массив значений операндов.

В таблице 3 приведены все тестовые массивы, нужные для структурного представления тест-программы для проверки модуля A . Здесь $V(I_1), V(I_2), V(x_1)$ и $V(x_2)$ массивы вариации соответствующих сегментов тест-программы $P(I_1), P(I_2), P(x_1)$ и $P(x_2)$. Для каждой вариации сегмента тест-программы P_i из массива вариации V_i соответствует эталонное значение $\mathcal{E}_i$, выдаваемое на выход объекта. $\mathcal{E}(I_1), \mathcal{E}(I_2), \mathcal{E}(x_1)$ и $\mathcal{E}(x_2)$ обозначают массивы эталонов. Массивы $O(I_1), O(I_2), O(x_1)$ и $O(x_2)$ используются для сохранения исходных значений операндов. Программа $P = \{ P_y, P_t, P_{tr} \}$ включает установочную часть (для установки исходных значений операндов), тестовый набор P_t и транспортирующую часть P_{tr} (для транспортировки результатов теста к выходу объекта). Массивы $O(I_1) \dots O(x_2)$ используются для модификации P_t . Массивы $\mathcal{E}(I_1) \dots \mathcal{E}(x_2)$ используются в P_{tr} .

Таблица 2

T	I_1	I_2	X_1	X_2	F
T_1	1	1	1	1	F_1
T_2	2	1	1	1	F_3
T_3	3	1	1	1	F_4
T_4	4	1	1	1	F_6
T_5	5	1	1	1	F_9
T_6	6	1	1	1	F_{10}
T_7	4	11	1	1	F_6
T_8	4	12	1	1	F_7
T_9	4	13	1	1	F_8
T_{10}	1	1	0	1	F_2
T_{11}	1	1	0	0	F_5
TD	D	D	D	D	D

Таблица 3

Операнды	$V(I_1)$		$V(I_2)$	$V(X_1)$	$V(X_2)$
A	1	$\exists(I_1)$	11	I	I
		$(A)+(R)$			
R	2	$-(A)$	12	0	0
$O(I_1)$	3	$(A)+1$	13	$\exists(X_1)$	$\exists(X_2)$
$O(I_2)$	4	$(A) \vee (R)$	$\exists(I_2)$		
$O(X_1)$	5	$(A_{n+1})=(A_n)$	$(A) \vee (R)$	$(A)-(R)$	$(A)-1$
$O(X_2)$	6	$(A_n)=(A_{n+1})$	$(A) \wedge (R)$		
			$(A) \oplus (R)$		

Вопросы формирования конкретных тест-программ для микропроцессорных БИС на основе массивов тестовых данных V_i , O_i и Δ_i рассматриваются в работе [3].

На основе предложенного метода разрабатывается система построения тестов для микропроцессорных БИС. Эта система находится в стадии отладки.

Преимущество предлагаемого метода автоматизированного построения тестов для микропроцессорных БИС и систем по сравнению с известными заключается в эффективном распределении работ между машиной и человеком.

Л и т е р а т у р а

1. С е л е з н е в А.В., Д о б р и ц а В.Т., У б а р Р.Р. Проектирование автоматизированных систем контроля бытового оборудования летающих аппаратов. - М.: Машиностроение, 1983. - 224 с.

2. У б а р Р.Р. Универсальный подход к автоматизации проектирования тестов для широкого класса дискретных объектов // Тр. Таллинск. политехн. ин-та. - 1986. - 626. - С.70-92.

3. А л а н г о В., К о н ь т Т., У б а р Р. Формирование тест-программ для микропроцессорных БИС при автоматическом генерировании тестов // См. наст. сб., с. 78-87.

T. Kont

An Approach to Test Generation Automation for Microprocessors

Abstract

In this paper an automatic test generation method for microprocessors is introduced. The method uses digital objects alternative graph models for automatic test generation at the functional level.

T. Kont

Testide sünteesi automatiseerimisest
mikroprotsessoritele

Kokkuvõte

Artiklis käsitletakse mikroprotsessoritele mõeldud testide automaatsünteesi meetodit. Mikroprotsessorite funktsionaalsete testide automaatne genereerimine toimub diskreetset objekti kirjeldavatel alternatiivsetel graafidel.

Р. Убар, Т. Лохуару, Б. Штраубе, Г. Элст

ПОСТРОЕНИЕ ТЕСТОВ ДЛЯ ПРОВЕРКИ ОПЕРАЦИОННЫХ
ЧАСТЕЙ ДИСКРЕТНЫХ СИСТЕМ

Быстрый рост сложности современных дискретных систем и необходимость их многоуровневого представления при решении задач тестового диагностирования вызвали появление целой совокупности различных математических моделей и методов как структурных, так и функциональных, что затрудняет создание достаточно универсальных систем автоматизированного проектирования диагностического обеспечения для изделий вычислительной техники.

В [1, 2] предложена математическая модель альтернативных графов (АГ) с целью формализации процесса синтеза тестов для широкого класса дискретных объектов, заданных на различных уровнях их представления. В данной статье рассматриваются возможности применения модели АГ для автоматизации процедур построения тестовых последовательностей для операционных частей дискретных систем.

Рассмотрим операционный автомат (операционную часть дискретной системы) в виде динамической системы

$$S = (Z, F),$$

состояние которой в каждый момент времени t определяется значениями переменных из множества

$$Z = Z_{вх} \cup Z_{вн} \cup Z_{вых},$$

где $Z_{вх}$, $Z_{вн}$ и $Z_{вых}$ — подмножества соответственно входных, внутренних и выходных переменных системы,

а значения всех переменных $z_i \in Z$ задаются на конечных множествах $A(z_i) = \{0, 1, 2, \dots\}$ и время t отсчитывается дискретно. Здесь F — множество передаточных функций объекта. В

случае дискретных объектов с памятью вводится некоторое подмножество переменных памяти $Z_n \subseteq Z_{вх}$, а множество F представляет собой систему рекуррентных временных функций

$$z_k = f(\underline{z}_{k,k}, \underline{z}'_{k,n}),$$

где $z_k \in Z_{вых} \cup Z_{вн}$, $f_k \in F$,

$\underline{z}_{k,k}$ и $\underline{z}_{k,n}$ – векторы, элементам которых однозначно сопоставлены переменные соответственно из подмножеств

$$Z_{k,k} \subset Z_{вх} \cup (Z_{вн} / Z_n)$$

и

$$Z_{k,n} \in Z_n.$$

Здесь штрих обозначает, что значение переменной относится к предыдущему моменту времени (такту).

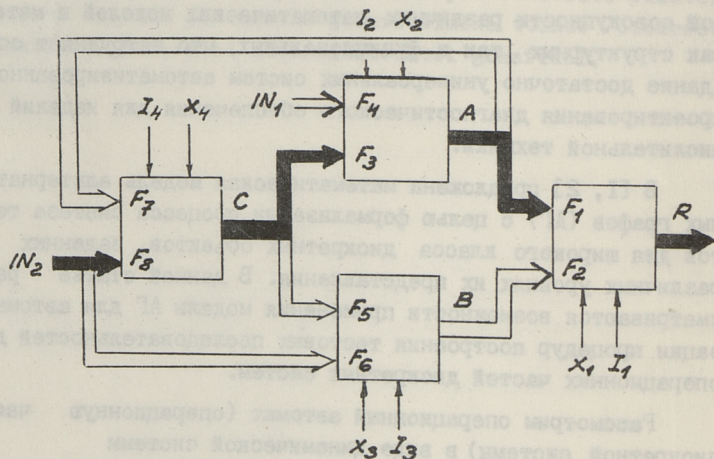


Рис. 1.

Пример такого операционного автомата (ОА) приведен на рис. 1. Здесь R, A, B, C – переменные данных (слова); F_i – функции, выполняемые компонентами ОА; x_k – логические условия (двоичные сигналы); IN_j – входные данные (слова) ОА;

Im – управляющие команды для активизации конкретных функций в ОА.

Такой дискретный объект (ДО) может быть представлен в виде модели АГ

$$S_{AG} = (Z, G),$$

где G – некоторое множество АГ, и каждой функции $f_k \in F$ сопоставлен некоторый граф (подграф) $G_k \in G$.

Представим АГ – модель компонента ДО с функцией

$$z_k = f_k(\underline{z}_k),$$

где $\underline{z}_k$ – вектор переменных множества $z_k \subset Z$ в виде ориентированного графа $G_k = (M_k, \Gamma_k, Z_k)$.

Здесь M_k – множество вершин графа;

Γ_k – отображение M_k в M_k ;

Z_k – множество переменных (аргументов функции f_k).

Каждой вершине $m \in M_k$ приписана некоторая весовая функция

$$e(m) = f_{k,m}(\underline{z}_{k,m}),$$

где $\underline{z}_{k,m} \in Z_k$.

Согласно функции f_k на графе определено отображение $\Gamma_k(m, e(m))$, по которому каждому значению $e(m)$ однозначно соответствует или некоторый последователь вершины m ,

$\Gamma_k(m, e(m)) \in \Gamma_k(m)$, или $\Gamma_k(m, e(m)) = \phi$. Любой набор $\underline{z}_k$ порождает транзитивное замыкание $\Gamma_k(m, e(m))$ в виде некоторого пути $\ell_k(\underline{z}_k) = \hat{\Gamma}_k(m_k^H, \underline{z}_k)$ из начальной вершины графа

m_k^H в некоторую терминальную вершину $m_k^T \in \ell_k(\underline{z}_k)$, где $\Gamma_k(m_k^T, e(m_k^T)) = \phi$.

Скажем, что АГ G_k представляет функцию $z_k = f_k(\underline{z}_k)$,

если для любого набора $\underline{z}_k$, порождающего на G_k путь

$\ell_k(\underline{z}_k) = \hat{\Gamma}_k(m_k^H, \underline{z}_k)$ с терминальной вершиной $m_k^T \in \ell_k(\underline{z}_k)$ имеет место $f_k(\underline{z}_k) = e(m_k^T)$.

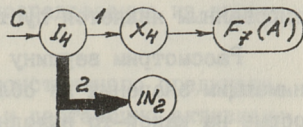
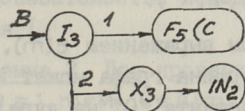
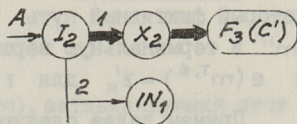
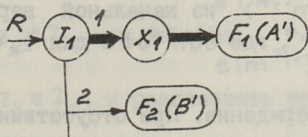


Рис. 2.

На рис. 2 рассматривается пример системы АГ, описывающий операционный автомат, представленный на рис. 1.

Синтез тестов для дискретного объекта на модели АГ формулируется как задача синтеза тестов для проверки вершин АГ. Вершины, подлежащие проверке, подбираются в процессе последовательной активизации путей, проводимом так, чтобы все вершины в графе оказались пройденными.

Рассмотрим случай однографовой модели, где ДО, реализующий функцию $z_k = f_k(z_k)$, представлен графом $G_k = (M_k, \Gamma_k, Z_k)$. Согласно определению [2] неисправности в АГ приводят или к активизации одних путей вместо других, или к активизации нескольких путей одновременно, или к неактивизации ни одного пути в графе. Одновременная активизация нескольких путей на АГ эквивалентна одновременному выполнению нескольких функций (операций) ДО. Предположим далее, что имеет место следующее утверждение.

Утверждение. При одновременном (вследствие неисправности) выполнении нескольких операций на выходе ДО образуется дизъюнкция результатов отдельных операций.

Все получаемые ниже результаты будут основываться на утверждении. В ряде случаев на практике может быть оправдана другая модель, если одновременное выполнение нескольких операций соответствует формированию конъюнкции результатов отдельных операций. Получаемые ниже результаты легко приводимы к этому случаю.

Определение 1. Пустым путем на АГ-модели называется некоторый фиктивный путь $l(m^H, m^{T,\phi})$ из начальной вершины m^H в терминальную вершину $m^{T,\phi}$, где $e(m^{T,\phi}) \equiv 0$ для $z_k \notin Z_n$ и $e(m^{T,\phi}) = z'_k$ для $z_k \in Z_n$.

Примем далее следующее утверждение. При отсутствии в АГ активизированных путей формально считается, что активизированным является пустой путь $l(m^H, m^{T,\phi})$.

Рассмотрим вершину m с весовым выражением $e(m)$, принимающим значения из области $A(m)$. Вершина графа может иметь обрыв на какой-то выходной дуге (константа "0" на дуге) или иметь постоянно "открытый" выход (константа "1"). Сопоставим с каждой выходной дугой вершины предикат

$$e(m, i), i \in A(m),$$

где

$$e(m, i) = \begin{cases} 0, & \text{если } e(m) \neq i, \\ 1, & \text{если } e(m) = i. \end{cases}$$

Определение 2. Модель неисправностей ДО, заданного в виде АГ-модели $S_{AG} = (Z, G)$, включает следующие неисправности:

1) константные предикаты

$$e(m, i) \equiv D, \quad i \in A(m), \quad D \in \{0, 1\};$$

2) условно-константные предикаты

$$e(m, j \rightarrow a) = 1, \quad j \in A(m), \quad a \in A(m),$$

где

$$e(m, j \rightarrow a) = \bigwedge_{i \in a} e(m, i) \mid e(m, j) = 1;$$

3) функциональные неисправности весовой функции $e(m)$.

В зависимости от точности отражения на модели внутренней структуры объекта предложенная модель неисправностей позволяет покрывать широкий класс структурных и функциональных неисправностей ДО. Так, например, для многоразрядных слов, взвешивающих нетерминальные вершины и представляющих коды команд (микрокоманд, состояний микропрограммных автоматов), моделью покрываются функциональные дефекты дешифраторов и управляемых мультиплексоров.

Лемма 1. Для проверки в АГ G_k константного предиката $e(m, i) \equiv 0, m \in M_k, i \in A(m)$ необходимо и достаточно активизировать в графе путь $e(m^H, m^T, i)$, проходящий вершину m при условии $e(m) = i$ так, чтобы выполнялось условие

$$e(m^T, i) \neq 0,$$

если $z_k \notin Z_n$, или условие

$$e(m^T, i) \neq z'_k,$$

если $z_k \in Z_n$, и реализовать тест $T_i(m)$, активизирующий этот путь.

Доказательство леммы следует непосредственно из определения 1.

Лемма 2. Для проверки в АГ G_k константного предиката $e(m, i) \equiv 1, m \in M_k, i \in A(m)$ необходимо и достаточно активизировать в графе два пути: путь $\ell_i(m^H, m^T, i)$, проходящий вершину m при условии $e(m) = i$, и некоторый другой путь

$\ell_j(m^H, m^{T,j})$, проходящий вершину m при условии $e(m)=j$, $j \neq 1$ так, чтобы выполнялось неравенство

$$e(m^{T,i}) \overline{e(m^{T,j})} \neq 0, \quad (1)$$

и реализовать тест $T_j(m)$, соответствующий пути $\ell_j(m^H, m^{T,j})$.

Доказательство. Действительно, для проверки указанной неисправности необходимо и достаточно, чтобы при реализации пути $\ell_j(m^H, m^{T,j})$ имело место неравенство $z_k = e(m^{T,j}) \neq z_k^H$ или $e(m^{T,j}) \oplus z_k^H \neq 0$. Здесь z_k^H — значение переменной z_k в случае неисправности. Так как, согласно утверждению I, имеет место

$$z_k^H = e(m^{T,i}) \vee e(m^{T,j}), \quad (2)$$

то получим $e(m^{T,j}) \oplus (e(m^{T,i}) \vee e(m^{T,j})) =$

$$= \overline{e(m^{T,j})} e(m^{T,i}), \quad (3)$$

что и требовалось доказать.

Лемма 3. Для проверки в АГ G_k условно-константного предиката $e(m, j \rightarrow a) = 1$, где $m \in M_k$, $a \in A(m)$, $j \in a$, необходимо и достаточно активизировать в графе оба пути: путь $\ell_j(m^H, m^{T,j})$, проходящий вершину m при условии $e(m)=j$, и некоторый другой путь $\ell_i(m^H, m^{T,i})$, где $i \in a$, $i \neq j$, проходящий вершину m при условии $e(m)=i$, так, чтобы выполнялось неравенство

$$e(m^{T,i}) \overline{e(m^{T,j})} \neq 0, \quad (4)$$

и реализовать тест $T_j(m)$, соответствующий пути $\ell_j(m^H, m^{T,j})$.

Доказательство аналогично доказательству леммы 2.

Заметим, что условие (4) леммы 3 может быть ослаблено, если $j \notin a$. Тогда вместо (4) достаточно выполнения неравенства

$$e(m^{T,i}) \neq e(m^{T,j}).$$

Лемма 4. Для проверки в АГ G_k при нетерминальной вершине $m \notin M_k^T$ функциональной неисправности r_{ij}^F

$$W_{ij}^F(z_{k,m}) = 1 \quad (5)$$

необходимо и достаточно активизировать в графе $|A(m)|$ пути

$\ell_h(m^H, m^{T,h})$, $h \in A(m)$, проходящие вершину m при условии $e(m) = h$, так, чтобы выполнялось условие

$$\forall h \in A(m), h \neq i: e(m^{T,h}) \neq e(m^{T,i}). \quad (6)$$

и реализовать тест $T_i(m)$, удовлетворяющий условию (5).

Доказательство. Предположим, что при выполнении условия (5) $e(m) = i$, а в присутствии неисправности $r_{ij}^F: e(m) = j$, $j \in A(m)$, $j \neq i$.

1. Достаточность. Пусть условие (6) выполнено. Предположим, что неисправность r_{ij}^F не обнаруживается тестом $T_i(m)$. Тогда должно иметь место $e(m^{T,i}) = e(m^{T,j})$, что противоречит условию (6).

2. Необходимость. Предположим, что $\exists i, j \in A(m)$, $i \neq j: e(m^{T,j}) = e(m^{T,i})$. Но если так, то неисправность r_{ij}^F не может быть обнаружена. Лемма доказана.

Лемма 5. Для проверки в АГ G_k при терминальной вершине $m \in M_k^T$ функциональной неисправности r_i^F

$$W_i^F(\bar{z}_{k,m}) = 1 \quad (7)$$

необходимо и достаточно реализовать тест $T_i(m)$, удовлетворяющий условию (7) и активизирующий в графе путь $\ell(m^H, m)$.

Доказательство леммы следует из определения модели АГ.

Рассмотрим ниже, не теряя общности, только такие АГ, где все нетерминальные вершины взвешены переменными $z \in Z$ (возможно, с инверсией). Легко убедиться, что любой АГ приводим к такому виду.

Из лемм I-5 следует теперь справедливость следующих теорем.

Теорема I. Для того, чтобы тест $T(m)$ проверял нетерминальную вершину $m \notin M_k^T$ в АГ G_k , достаточно, чтобы этот тест включал для каждой $i \in A(m)$ элементарную проверку $T_i(m) \in T(m)$, активизирующую путь $\ell(m^H, m^{T,i})$, проходящий вершину m через ее выход $e(m) = i$, и множество путей $\ell(m^j, m^{T,j})$, где $m_j = \Gamma_k(m, j)$, для всех $j \in A(m)$, $j \neq i$, так, чтобы выполнялись условия

$$\forall j \in A(m), j \neq i: e(m^{T,j}) e(m^{T,i}) \neq 0. \quad (8)$$

Теорема 2. Для того, чтобы тест $T(m)$ проверял терминальную вершину $m \in M_k^T$ в АГ G_k , достаточно, чтобы этот тест включал элементарные проверки $T_i(m) \in T(m)$, активизирующие путь $\ell(m^H, m)$ и удовлетворяющие условиям $W_i^F(\underline{x}_{k,m}) = 1$, $i: r_i \in R^F(m)$, где $R^F(m)$ – множество функциональных неисправностей весовой функции $e(m)$.

На основе теорем 1 и 2 базируется формальный алгоритм генерирования тестов, использующий модель АГ.

Легко заметить, что удовлетворение усиленного условия

$$\forall i, j \in A(m), i \neq j : e(m^{T,i}) \overline{e(m^{T,j})} = 0$$

позволяет организовать тест, проверяющий одновременно все выходы вершины m . Для этого необходимо повторить тест в $|A(m)|$ раз, изменяя циклически значение $e(m)$ и реализуя таким образом все ветвления из вершины m .

Итак, учитывая вышесказанное, синтез теста для некоторого выхода $e(m)=i$ вершины m состоит из следующих этапов:

- активизации некоторого пути $\ell(m^H, m)$ из начальной вершины графа до проверяемой вершины m ;
- активизации несовпадающих путей $\ell(m^j, m^{T,j})$ для каждой из вершин $m^j \in \Gamma_k(m, j)$ до терминальных вершин $m^{T,j}$;
- решения системы неравенств (8).

При проверке терминальных вершин требуется решение всего лишь первого этапа. Решение второго и третьего этапов сводится к циклическому повторению теста, найденного на первом этапе, для различных значений проверяемого весового выражения $e(m^T)$. Эти значения могут быть найдены методом полного перебора или путем генерирования случайных чисел.

Для решения системы неравенств (8) предложим алгоритм, основывающийся на решении уравнений

$$\exists_k = \overline{0, n} : F_j(k) \overline{F_i(k)} = 1,$$

где $F_j = e(m^{T,j})$ рассматривается как $(n+1)$ -разрядный двоичный вектор;

$F_j(k)$ – значение k -го разряда F_j .

Алгоритм 1.

I^0 Берем $k = 0$, $i = 0$, $j = 0$.

2° Если $j = |A(m)|$, то переходим к п. 3°, иначе $j := j+1$ и переходим к п. 5°.

3° Если $i = |A(m)|$, то конец.

4° $i := i+1$, $j := 0$.

5° Если $i \neq j$, то переходим к п. 2°.

6° Если значения $F_j(k)$ и $F_i(k)$ определены, то переходим к п. 7°, иначе переходим к п. 9°.

7° Если $F_j(k) \overline{F_i(k)} = 1$, то $k := 0$ и переходим к п. 2°.

8° $k := k+1$, то переходим к п. 6°.

9° Находим операнды в k -ом разряде так, чтобы выполнялось условие $F_j(k) \overline{F_i(k)} = 1$. Если такие операнды отсутствуют, то неисправность $e(m, i \rightarrow j)$ считается необнаруживаемой; принимаем $k := 0$ и переходим к п. 2°.

Примечание. В случае, когда в результате выполнения алгоритма I получается $k > n+1$, проверка данной группы неисправностей однократной реализацией функций $F_j, j \in A(m)$ невозможна и требуется их $\lceil (k+1)/(n+1) \rceil$ -кратная реализация с операндами, покрывающими разрядный диапазон $k+1$.

Процесс синтеза тестов для заданной дискретной системы организуется на модели АГ так, чтобы были найдены входные последовательности (программы, микропрограммы) для проверки всех вершин системы АГ. Процесс состоит из двух частей: из синтеза тестов для операционной части и синтеза тестов для управляющей части ДУ. Для активизации выбранного пути ℓ на АГ (множества путей на системе АГ) генерируется некоторая входная последовательность (программа, микропрограмма). Для контроля управляющей части путем решения неравенства (8) выберем подходящие операнды для сгенерированной программы (микропрограммы) так, чтобы выполнялись условия активизации неисправностей для всех вершин на пути ℓ . Для контроля операционной части заданной терминальной вершиной на пути ℓ найдем операнды или путем удовлетворения условия (7), или случайным образом, или по методу полного перебора.

При решении задач генерирования операндов для проверки управляющей части можно ссылаться, согласно алгоритму I, на однобитовые операнды. Каждый однобитовый опе-

ранд будет предназначен для обнаружения некоторого подмножества неисправностей. Остальные неисправности будут обнаружены другими однобитовыми операндами. Таким образом, систематически бит за битом можно конструировать целые машинные слова, состоящие из множеств однобитовых операндов. Такой метод упрощает перебор нужных данных для сгенерированной тест-программы. Для проверки отсутствия дефектов управления по всей разрядной сетке операционного автомата необходимо циклическое проведение теста на множестве сдвигаемых данных так, чтобы все однобитовые операнды были использованы для проверки всех разрядов.

В качестве примера построения тестов рассмотрим операционный автомат, представленный в виде структурной схемы на рис. 1, и АГ-модели на рис. 2. Жирными линиями представлены активизированные пути. Процедура активизации графов и построения теста иллюстрируется в виде дерева на рис. 3.

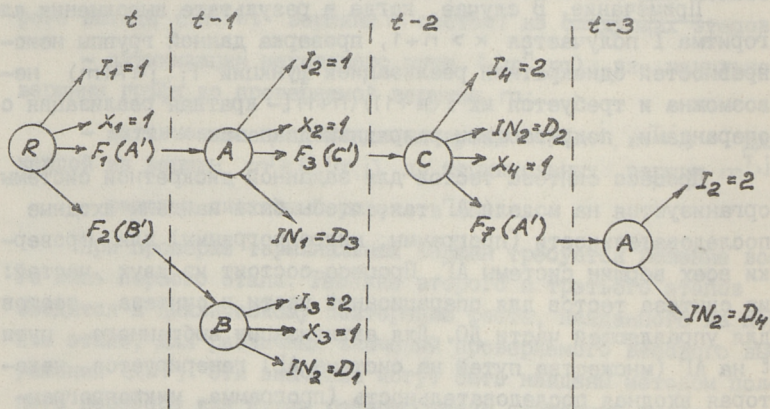


Рис. 3.

Рассматривается активизация системы АГ относительно цепи $M_u = \{F_1(A'), F_3(C'), IN_2\}$. Все вершины цепи являются терминальными. На базе этой цепи путем активизации дополнительных путей могут быть проверены вершины $I_1, X_1, F_1(A')$ (в графе R), $I_2, X_2, F_3(C')$ (в графе A), I_4, IN_2 (в графе C). В частности, для проверки неисправностей при I_1, I_2 и I_4 требуется активизировать остальные ветви из этих вершин (например, ветвь $X_4=1$ в графе C) и решить систему неравенств

$$\left. \begin{aligned} (I_1, 1 \rightarrow 2) &: F_2(D_1) \\ (I_2, 1 \rightarrow 2) &: F_1(D_3) \\ (I_4, 2 \rightarrow 1) &: F_1(F_3(D_4)) \end{aligned} \right\} \neq F_1(F_3(D_2)).$$

Здесь через $(I \cdot j \rightarrow i)$ обозначена неисправность, заключающаяся в ложном изменении значения $I=i$ (вместо управляющего сигнала $I=j$ выдается сигнал $I=i$). Тест, состоящий из 4-х тактов и полученный в соответствии с деревом на рис. 3, приведен в таблице.

Т а б л и ц а

Такт	I_1	I_2	I_3	I_4	X_1	X_2	X_3	X_4	IN_1	IN_2	R
$t-3$		2								D_4	
$t-2$				2				1		D_2	
$t-1$		1	2			1	1		D_3	D_1	
t	1				1						$F_1(F_3(D_2))$

По структуре АГ могут быть получены верхние и нижние оценки длины теста. Рассмотрим АГ $G_K = \{M_K, \Gamma_K, Z_K\}$, представляющий ДО с функцией $z_K = f_K(z_K)$. Пусть $|A(m)|$ число возможных выходов $m \in M_K$. Для $m \in M_K^T$ под $A(m)$ будем понимать число тестовых наборов, необходимых для проверки m с функцией $e(m)$. Пусть L — множество всех путей на рассматриваемом АГ. Тогда верхней оценкой длины теста для этого графа служит

$$|T|_{\max} = \sum_{m \in M_K \setminus M'} |A(m)| \cdot (|A(m)| - 1) + \sum_{m \in M'} |A(m)|, \quad (9)$$

где

$$M' = \{m \mid A(m) > 1 \ \& \ m \in M_K^T\} \subset M_K.$$

Под верхней границей имеем в виду длину теста в том наихудшем случае, когда для обнаружения каждой неисправности типа $e(m, i \rightarrow j) = 1$ в нетерминальных вершинах и для каждой функциональной неисправности в терминальных вершинах требуется отдельная элементарная проверка (проверка неисправностей типа $e(m, i) = D$ и функциональных неисправностей весовой функции $e(m)$, очевидно, легко совместима с проверкой названных).

Для нижней границы длины теста имеем

$$|T|_{\min} = |L| + \sum_{m \in M'} |A(m)| - |M'|. \quad (10)$$

Нижняя граница достигается тогда, когда каждой элементарной проверкой, активизирующей некоторый путь $l \in L$, могут быть обнаружены сразу все неисправности, приводящие к искажению данного пути. Таким образом, обе границы реально достижимы.

В качестве частного случая рассмотрим регулярный АГ, где все вершины имеют k выходов и все пути содержат n вершин. Тогда из выражений (9) и (10) следует соответственно

$$|T|_{\max} = k(k-1) \sum_{i=0}^{n-1} k^i,$$

$$|T|_{\min} = k^n.$$

Так, например, для $k = 2$ и $n = 3$ имеем $|T|_{\max} = 14$ и $|T|_{\min} = 8$. К такому графу, в частности, могут привести булевы функции с числом аргументов от трех до семи.

Заметим, что частным случаем АГ являются двоичные АГ, где $\forall z \in Z: A(z) \in \{0,1\}$, и простые двоичные АГ, где функции $e(m)$ вырождены в двоичные переменные (возможно с инверсией). В класс простых двоичных АГ входят также бинарные деревья решений (БДР) [3].

Отсюда следует, что предложенный аппарат является общим для различных классов ДО. Описанный в статье метод синтеза тестов применим одинаково как для цифровых схем, заданных на логическом уровне, так и для дискретных систем, заданных на уровне регистровых передач.

В случае, когда ДО работает по жестким микропрограммам, задаваемым соответствующим управляющим автоматом (УА), выполнение первого этапа синтеза теста видоизменяется. Активизируемые на первом этапе пути АГ определяются заданными микропрограммами УА. Каждому пути в некоторой микропрограмме соответствует некоторая система активизированных путей на альтернативных графах. Это задает исходную ситуацию для выполнения второго и третьего этапов процесса синтеза тестов. Критерий завершения построения тестов не изменяется — все вершины системы АГ должны быть проверены. Но возможность доступа ко всем вершинам (активизации соответ-

вующих путей на первом этапе) определяется заданными микропрограммами УА.

Л и т е р а т у р а

1. С е л е з н е в А.В., Д о б р и ц а Б.Т., У б а р Р.Р. Проектирование автоматизированных систем контроля бортового оборудования летательных аппаратов. - М.: Машиностроение. - 1983. - 224 с.

2. У б а р Р.Р. Универсальный подход к автоматизации проектирования тестов для широкого класса дискретных объектов // Тр. Таллинск. политехн. ин-та. - 1986. - № 626. - С. 70-92.

3. A k e r s S.B. Binary decision diagrams // IEEE Trans. on Computers. - 1978, June. - Vol. C-27. - P. 509-516.

R. Ubar, T. Lohuaru, B. Straube, G. Elst

Test Generation for Operational Parts in Digital Systems

Abstract

In this paper the apparatus of alternative graphs as an instrument of describing the wide class of digital objects is considered. A formal test generation method for operational parts of digital systems on the basis of alternative graph apparatus is proposed.

R. Ubar, T. Lohuaru, B. Straube,
G. Elst

Testide genereerimine digitaalsüsteemide operatsioonosadele

Kokkuvõte

Artiklis käsitletakse alternatiivsete graafide aparaati laia digitaalsete objektide klassi kirjeldamiseks. Esitatakse meetod testide formaalseks sünteesiks digitaalsete süsteemide operatsioonosadele. Meetod põhineb alternatiivsete graafide mudeli kasutusel.

ФОРМИРОВАНИЕ ТЕСТ-ПРОГРАММ ДЛЯ МИКРОПРОЦЕССОРНЫХ
БИС ПРИ АВТОМАТИЧЕСКОМ ГЕНЕРИРОВАНИИ ТЕСТОВ

Высокая степень интеграции и отсутствие принципиальных схем микропроцессорных БИС делают практически непригодными классические методы их тестирования, базирующиеся на знании вентиляльной структуры объекта. Поэтому в последнее время все большее развитие и практическое применение находят модульный и функциональные подходы к тестированию сложных дискретных систем, основанных на микропроцессорных БИС. Один из таких подходов основывается на применении модели альтернативных графов (АГ) для представления широкого класса дискретных объектов [1, 2].

В данной работе рассматриваются некоторые вопросы формирования тест-программ при автоматическом генерировании тестов для микропроцессорных БИС, описанных моделью АГ. При этом учитываются некоторые особенности рассматриваемого класса объектов, в частности, возможности ручного задания установочных и транспортирующих последовательностей в виду их относительной простоты. Учитываются также требования со стороны системы контроля, предназначенной для проведения синтезируемых тест-программ.

В работе [2] предложен метод синтеза групповых тестов на модели АГ, преимуществом которого является повышенная достоверность обнаружения кратных неисправностей.

Целью составления тестовых групп является не проверка отсутствия неисправностей на ветвях вершин АГ, а проверка исправности путей АГ в условиях кратных неисправностей. Путем активизации определенного пути l_i на графе составляется некоторый тестовый набор (например, команда микропроцессора, состояние микропрограммного автомата, программируемого интерфейса и т.д.), а также некоторая программа, вклю-

чающая установочную часть, тестовый набор и транспортирующую часть. Программа должна выполняться циклически так, чтобы при каждом ее выполнении изменилось значение веса при некоторой вершине $m \in \ell_i$ и чтобы в конечном итоге активизировались циклически все выходы вершин на проверяемом пути. Каждая такая модификация (вариация) программы должна выполняться в общем случае многократно с операндами, выбранными так, чтобы отклонения от проверяемого пути дали результаты, отличные от базового варианта (от выполнения проверяемого пути ℓ_i).

Таким образом, групповой тест T можно представить как совокупность трех информационных массивов $T = \{P, V, S\}$, где P — тест-программа, $V = (v_1, v_2, \dots, v_n)$ — массив вариаций программы (варьируемых значений в программе) и $S = (s_1, s_2, \dots, s_m)$ — массив операндов (см. рис. 1). Программа $P = \{P_y, P_T, P_{Tr}\}$ включает установочную часть P_y , тестовый набор P_T и транспортирующую часть P_{Tr} . В программе можно выделить еще статическую (стандартную) и динамическую (нестандартную) части. Динамическая часть программы $P_d = \{P_v, P_s\}$ (заштрихованная область на рис. 1) относится обычно к установочной части и тестовому набору, причем $P_v \subset P_T$ и $P_s \subset P_y \cup P_T$. Массив вариаций V предназначен для циклической замены некоторой части программы $P_v \subset P$ элементами $v_i \in V$. Обозначим вариацию программы P , порождаемую элементом v_i , через $P(v_i)$. Массив операндов S предназначен для циклической замены части программы $P_s \subset P(v_i)$ элементами $s_j \in S$. Под s_j подразумевается определенная совокупность данных, необходимых для сведения объекта в необходимое состояние, включая данные, являющиеся частью тестового набора P_T . Обозначим конкретную реализацию программы, порождаемую вариацией v_i и операндами s_j , через $P(v_i, s_j)$. Выполнение группового теста заключается в циклическом повторении программы P для всех сочетаний $v_i \in V$ и $s_j \in S$.

Описанный способ структурного задания тестовой информации в виде тест-программы, массива вариаций и массива операндов, к чему добавляется и массив эталонных данных, отражающих ожидаемые результаты тестовых операций, открывает возможность алгоритмической генерации тестовых наборов в самом тестере по тривиальным алгоритмам. Они сводят-

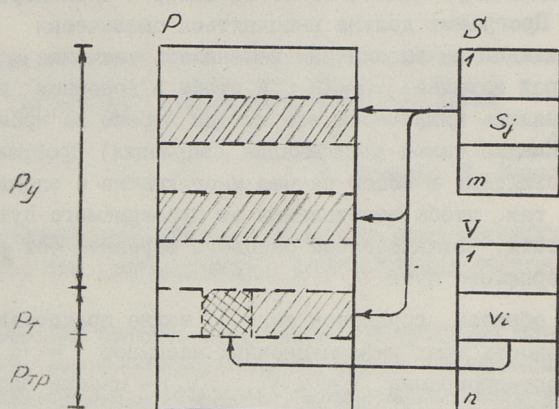


Рис. 1.

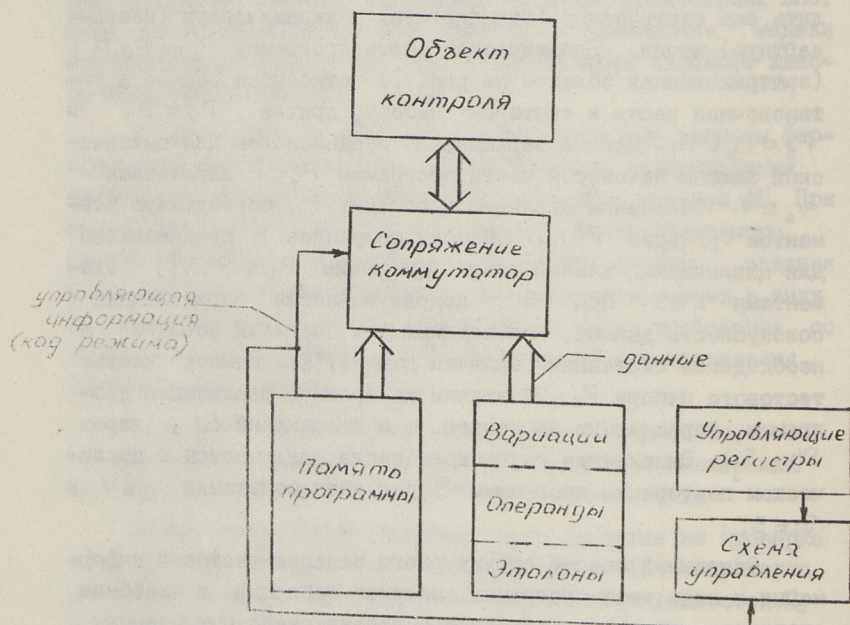


Рис. 2.

ся к управлению циклического выполнения тест-программы Р с дополнением ее соответствующими нужными данными из массивов V, S и массива эталонов.

Функциональная схема тестера, предназначенного для реализации групповых тестов, построенных на основе модели АГ, представлена на рис. 2. Тестовая информация вводится в тестер из внешнего ЗУ или ЭВМ, на которых производится автоматический синтез теста, через регистр команд. Распределение тестовой информации в памяти тестера происходит следующим образом. Тестовые наборы, состоящие из входного набора, эталона выходной реакции объекта, вектора направленности шин для рассматриваемого набора, маски выходов и кода события, определяющего момент проведения сеанса (реализации данного набора), загружаются в память программы. Память данных предназначена для хранения информации, используемой для модификации тестовых наборов при циклическом выполнении тест-программы. Память данных состоит из трех отдельно адресуемых блоков, предназначенных для хранения соответственно вариаций программы (в частности, кодов команд), операндов и эталонов (в частности, ожидаемых выходных адресов или данных). Управление коммутатором выбора источника информации производится кодом режима. В блоке управляющих регистров находятся счетчики текущих адресов для различных блоков памяти, а также счетчики и регистры, необходимые для организации циклов.

Синтез тестов производится по методике, изложенной в [2, 3]. Учитывая, однако, особенности класса рассматриваемых объектов, в методику введены следующие упрощения:

1) явное задание переменных времени заменяется приписыванием таких переменных, которые связаны с определенными моментами времени (вводимые и выводимые в микропроцессор данные), соответствующих параметризованных тестовых программ (подпрограмм ввода или "установки" для входных данных);

2) поскольку нахождение установочных и транспортирующих последовательностей для внутренних переменных объекта достаточно просто, то они подготавливаются заранее вручную в виде параметризованных подпрограмм.

Таким образом, при введенных упрощениях процесс синтеза теста можно провести на основе только проверяемого графа вместо рассмотрения всей системы графов, как это делается в общем случае [2], что в конечном итоге заметно упрощает процесс синтеза тест-программ (отпадают рекурсивные обращения из графов в графы при поиске подходящих установочных последовательностей). С другой стороны, заметно упрощается процесс обработки графа, так как в нем отсутствуют вершины, взвешенные с переменными времени [2].

Структура данных, используемых при формировании тестовой информации для тестера, изображена на рис. 3.

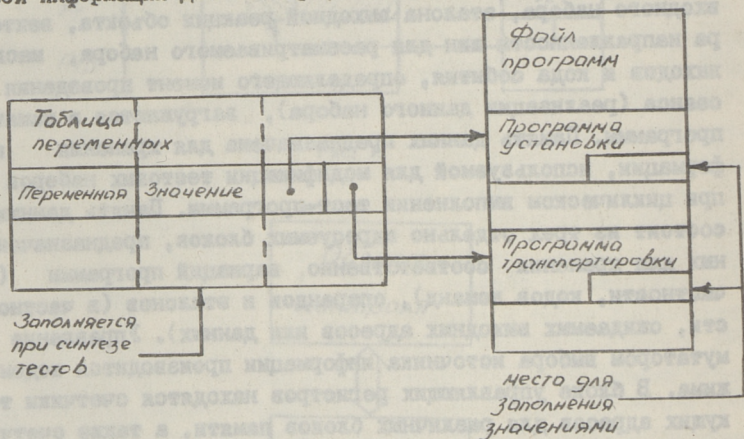


Рис. 3.

В таблице переменных переменные могут быть входными, внутренними и выходными. Их конкретные значения определяются в процессе генерации теста. Каждой входной (и внутренней) переменной приписывается некоторая программа установки, которая подготавливается заранее (вручную) и вводится в файл программ. Программа установки представляет собой последовательность неполностью определенных тест-векторов на входном языке тестера, предназначенных для ввода в объект контроля определенной информации (значения данной переменной). Каждой выходной (и внутренней) переменной приписывается некоторая программа транспортировки, которая также подготавливается заранее и вводится в файл программ. Программа установки представляет собой последовательность неполностью определенных тест-векторов на входном языке тестера, предназначенных для ввода в объект контроля определен-

ной информации (значения данной переменной). Каждой выходной (и внутренней) переменной приписывается некоторая программа транспортировки, которая также подготавливается заранее и вводится в файл программ. Программа транспортировки представляет собой последовательность неполностью определенных тест-векторов на входном языке тестера, предназначенных для вывода из объекта контроля в тестер значения рассматриваемой переменной.

Ручная подготовка программ установки для входных и программ транспортировки для выходных переменных сводится к написанию тестовых наборов, которые необходимы для организации соответствующих сеансов обмена информацией между тестером и объектом контроля. Такие наборы могут быть составлены на основе соответствующих временных диаграмм, заданных в справочниках по использованию объектом.

Ручная подготовка программ установки и транспортировки для внутренних переменных сводится к написанию программ для объекта (на языке тестера), которые необходимы для загрузки или чтения регистров, соответствующих этим переменным. При этом естественно выбирать для этих целей самые короткие программы, стремясь к тому, чтобы программы оказались независимыми, т.е. загрузка одного регистра не должна искажать содержимое другого регистра. Если последнее требование не выполняется, то должно быть проведено ранжирование переменных так, чтобы программа установки для переменной i -го ранга не могла искажать значения переменных j -го ранга, где $j < i$. Если для некоторой группы переменных такое ранжирование невозможно, необходимо их объединить в одну сложную переменную и подготовить для них одну единую программу установки, имея в виду, что взаимное искажение этих переменных детерминированно и может быть учтено при вычислении необходимых значений для них.

Процесс генерирования определенного сегмента теста (для проверки некоторой вершины АГ) окончится определением значений для некоторого подмножества переменных. После этого компилируется сегмент теста в следующем порядке:

- 1) выбираются программы установки для всех внутренних переменных в порядке увеличения рангов переменных;
- 2) выбираются программы установки для всех входных переменных;

3) выбирается программа транспортировки для проверяемой выходной или внутренней переменной.

Переменные объекта контроля могут в процессе генерирования тестов получить как константные, так и символьные (VAR, OP, ET) значения. В случае константы, последняя загружается непосредственно в компилируемую программу. В случае символов VAR, OP или ET значение переменной вводится соответственно в массив вариаций, массив операндов или массив эталонов, а соответствующий тест-вектор, который должен содержать это значение, снабжается управляющей информацией (кодом режима), которая указывает из какого источника данное значение берется (из памяти вариаций, операндов или эталонов).

Рассмотрим методику синтеза тест-программ для микропроцессорных БИС при помощи следующего упрощенного примера

Пусть объект контроля (гипотетический микропроцессор) задан в виде следующего множества предикатных выражений $P_i: F_i [3]$, отражающих систему команд (здесь $I1$ — код команды, Ii — входные данные, вводимые на i -ом машинном цикле, $R1$ и $R2$ — переменные регистров, Oj — выходные данные, выводимые в j -ом машинном цикле):

$(I1 = 1): R1 := I2;$

$(I1 = 2): R2 := I3;$

$(I1 = 3): O2 := R1, O3 := R2;$

$(I1 = 4): R1 := R1 + R2.$

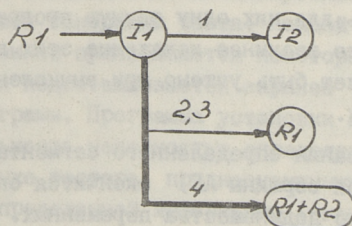


Рис. 4.

Соответствующий этому описанию АГ для регистра $R1$, синтезированный по методике [4], приведен на рис. 4. Общее множество переменных включает следующие подмножества:

$$Z_{Bx} = \{11, 12, 13\}, Z_{Bn} = \{R1, R2\}, Z_{Bых} = \{02, 03\},$$

где Z_{Bx} , Z_{Bn} и $Z_{Bых}$ множества соответственно входных, внутренних и выходных переменных. Переменным $z \in Z_{Bx} \cup Z_{Bn}$ приписаны соответствующие подпрограммы установки $Y_z(z^x)$, где z^x — параметр, заменяемый конкретным значением. Аналогично переменным $z \in Z_{Bn} \cup Z_{Bых}$ приписаны подпрограммы транспортировки $T_z(z^x)$, где z^x — параметр, заменяемый конкретным транспортируемым значением.

Иллюстрируем процесс синтеза для проверки поведения регистра R1 при выполнении команды I1=4 (проверки вершины I1 графа R1 при значении I1=4). Согласно методике [2] активизируется путь через вершины I1 и R1'+R2' и решается неравенство

$$I2 \neq R1' + R2' \neq R1'.$$

При этом получим: $I1=4$, $R1=x_1$, где $x_1 = R1' + R2'$ для текущего момента (момента тестирования) и $R1'=x_2$, $R2'=x_3$ для предыдущего момента (момента установления состояния объекта до тестирования). Тест-программа реализуется в виде следующей последовательности

$$ТП = Y_{R1}(x_2) \cdot Y_{R2}(x_3) \cdot Y_{I1}(4) \cdot T_{R1}(x_3).$$

Так, например, если программирование тестера происходит по событийному принципу [2], то окончательная тест-программа принимает следующий вид:

$$\left. \begin{array}{l} T1: F=1, DB=1; \\ T2: RD=1, DB=x_2; \end{array} \right\} \text{установка R1;} \\ \left. \begin{array}{l} T3: F=1, DB=2; \\ T4: RD=1; \\ T5: RD=1, DB=x_3; \end{array} \right\} \text{установка R2;} \\ \left. \begin{array}{l} T6: F=1, DB=4; \\ T7: F=1, DB=3; \end{array} \right\} \text{тест-вектор;} \\ \left. \begin{array}{l} T8: WR=1, DB=x_1; \end{array} \right\} \text{транспортировка R1,}$$

где T_i — метка тест-вектора;

$F=1$ — событие начала приема команды;

RD — событие начала приема данных;

$WR=1$ — событие начала вывода данных;

DB -- переменная шины данных (при $RD=1$ она служит входом, а при $WR=1$ -- выходом и ее значение соответствует ожидаемой реакции объекта на данном тесте).

Преимущество предлагаемого подхода автоматизированного синтеза тест-программ для микропроцессорных БИС и систем, по сравнению с известными, заключается в эффективном распределении работ между машиной и человеком. Инженеру остается возможность выбора разумных и эффективных вспомогательных процедур (для установки и транспортировки), машинный поиск которых может привести к длительному перебору вариантов и необязательно к нахождению эффективного решения. В то же время для машины остается рутинная работа по синтезу общих структур тестовой программы, синтезу и решению неравенств для подбора необходимых тестовых данных и, наконец, по компиляции последовательностей конкретных тест-векторов.

Благодаря структурированному представлению тестовой информации резко сокращается объем памяти, необходимый для хранения теста, а также время проверки контролируемого объекта за счет минимизации объема данных, вводимых в память тестера. Структурированная тестовая информация представляется в сжатом виде, а генерируемые по ней в реальном масштабе времени тестовые последовательности могут быть очень длинными, причем не влияющими на производительность устройства.

Получаемый эффект от введенных возможностей представления и реализации тестов, заданных в описанном структурированном виде, можно определить формулой

$$K_3 = n P V O / (n P + m(V + O)),$$

где n -- длина слова памяти программы;
 m -- длина слова памяти данных;
 P -- длина программного массива;
 V -- длина массива вариаций;
 O -- длина массива операндов.

Л и т е р а т у р а

И. У б а р Р.Р. Тестовая диагностика цифровых устройств. Часть I-II. --Таллин: ТПИ, 1981. -- 226 с.

2. У б а р Р.Р. Универсальный подход к автоматизации проектирования тестов для широкого класса дискретных объектов // Тр. Таллинск. политехн. ин-та. - 1986. - № 626. - С. 70-92.

3. К о н ь т Т. Подход к автоматизации построения тестов для микропроцессорных БИС // См. наст. сб., с. 55-64.

4. А л а н г о В.Р. Функциональный подход к созданию математических моделей дискретных объектов на обобщенных альтернативных графах // См. наст. сб., с. 43-54.

V. Alango, T. Kont, R. Ubar

Test-Program Generating for Microprocessor VLSI

Abstract

Some problems of test-program generating for microprocessor VLSI on the alternative-graph-model are discussed. Special features of this class of digital objects, like the possibility of the straightforward solving of the controllability and observability tasks, are exploited to find the suitable correlation of manual and automatic parts in test generation processes.

V. Alango, T. Kont, R. Ubar

Testprogramide formeerimine mikroprotsessorilülituste testide automaatsel genereerimisel

Kokkuvõte

Käsitletakse testprogramide formeerimise probleeme testide genereerimise automatiseerimisel mikroprotsessorilülitustele alternatiivsete graafide mudelil. Näidatakse võimalust kasutada vaadeldava klassi objektide eriomadusi (asetus- ja transportjadade etteandmise lihtsust) optimaalse suhte leidmisel käsitsitöö ja automatiseerimise vahel testide genereerimise protsessis.

Б. Бомбик, В. Вийес

ПАКЕТ ГРАФИЧЕСКИХ ПРОГРАММ ДЛЯ КУРСА МОДЕЛИРОВАНИЯ ЦИФРОВЫХ СХЕМ

Характерной чертой настоящего времени является интенсивное использование графических пакетов прикладных программ для персональных ЭВМ (ПЭВМ). На кафедре ЭВМ Таллинского политехнического института имеется класс отечественных ПЭВМ ДВК-2М. В связи с тем, что специального программного обеспечения, поддерживающего спецкурс моделирования кафедры, для них нет — только было положено начало его создания. Одной из первых создана ниже описываемая система программ графического отображения и моделирования дискретных объектов. Надо заметить, что наши ДВК-2М снабжены платой графического интерфейса, позволяющей подключить в конфигурацию периферийное устройство — цветной телевизор, и использовать для работы программ два экрана.

Теоретически алгоритм программ основывается на функциональном описании дискретных объектов. Причем предполагается, что компоненты цифровой схемы задаются в виде множества булевых функций с двоичным алфавитом для значений сигнала. Такой подход позволяет решать простые задачи моделирования и тестирования цифровых схем.

В дальнейшем так разработанные схемы могут стать учебными объектами для изучения аппаратуры диагностики.

Технически пакет программ основан на использовании алгоритмического языка Паскаль с библиотекой внешних процедур управления графическим экраном.

Данными процедурами являются:

Point (C,X,Y) — занесение точки на экран;
Vectr (C,X,Y) — черчение прямой от места курсора до заданной точки;

Vecti (C, X₁, Y₁, X₂, Y₂) – соединение прямой двух точек;
 Dchar (C, S) – занесение на экран цветного изображения
 символа по коду ASC11.

При графической реализации логических функций было принято следующее ограничение – схема должна поместиться на экран. Величина экрана состоит из 255x255 точек, и на экране возможно использовать четыре основных цвета (0 – черный, 1 – красный, 2 – голубой, 3 – зеленый). Из этого ограничением на число аргументов функции $Y = f(X_1 \dots X_n)$ стало $N=7$, а на ее сложность – (число составных компонентов) $N = 5$.

Создание графического отображения функции пояснено на примере $Y = (A \vee B) \& C$, где A, B, C, D – входные сигналы, а Y – выходная функция D. На первом этапе элементная база графических схем ограничена логическими элементами нижнего уровня (логическое И, ИЛИ, НЕТ), но созданы уже подпрограммы для Т-Т триггера и сумматора. При графическом построении функции Y используется матрица с размерностью 300 членов по глубине, для каждого возможного элемента на экране. Такая матрица представлена на рис. 1, где в зависимости от очередности входов в логические элементы цифровой схемы на выбранные для них места записываются соответствующие коды.

Y X	0	1	2	3	4	5	6	7	8	9
(0	0	50	0	0	0	0	0	0	0	0)
10	5	20	0	0	0	0	0	0	0	30
20	3	301	302	0	0	0	0	0	0	10
30	2	10	40	0	0	0	0	0	0	50
40	1	303	0	0	0	0	0	0	0	30
50	3	30	304	0	0	0	0	0	0	0
60	0	0	0	0	0	0	0	0	0	0
:	:	:	:	:	:	:	:	:	:	:
290	0	0	0	0	0	0	0	0	0	0

Рис. 1. Исходная матрица графической схемы.

Далее, исходя из выбранной топографии элементов, которая задана в виде входного меню при начале работы, увеличивается шаг дискретности на экране и происходит переход к матрице рисунка 2.

Y X	0	1	2	3	4	5
1	0	0	0	0	0	0
2	0	0	0	0	0	0
3	301(A)	0	0	0	0	0
4	302(B)	0	20(V)	0	0	0
5	303(C)	0	40(^)	30(&)	50(V)	0
6	304(D)	0	0	0	0	0
7	0	0	0	0	0	0

Рис. 2. Конечная матрица графической схемы.

В результате последнего преобразования матрицы получается изображение, представленное на рис. 3 (конечно графическое и цветное).

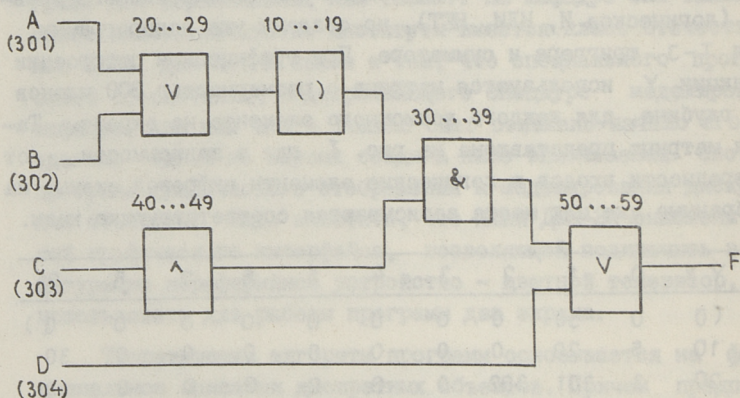


Рис. 3. Эскиз графического отображения функции.

После получения логической схемы может произойти ее тестирование в соответствии с заданием путем ввода значений входных переменных. Все значения входов и выходов также отображаются на экране.

Полученное изображение записывается агрегатно на гибкий диск для дальнейшей обработки. Последняя может включать локализацию конкретных путей на схеме и их выделения.

Описанная система программ состоит из отдельных программных модулей, в реализации которых участвовали кроме студентов ТПИ стажер кафедры из ГДР и члены кружка ЭВМ.

Л и т е р а т у р а

1. М а м и к о н о в А.Г. Проектирование АСЦ. - М.: Высшая школа. - 1987. - 303 с.

2. Ф л о р е с И. Структура и управление данными. - М.: Финансы и статистика, 1982. - 319 с.

B. Bombik, V. Viies

Program with Graphic Screen-output for Digital Schemes Model Course

Abstract

In this paper the idea to use the personal computer DVK-2M with two screens for digital schemes model course is described. One of the screens is a color-tv and we also have a special subprogram library to work with it. At the first stage we can describe only logical functions with seven input parameters.

The program is written in the Pascal language.

B. Bombik, V. Viies

Graafiliste programme pakett numbriliste skeemide modelleerimise kursusele

Kokkuvõte

Antud artiklis on kirjeldatud rakendusprogrammide paketti modelleerimise kursuse toetuseks, kasutades kodumaist personaalarvutit DVK-2M. Antud arvutid on meil varustatud graafikaplaadiga, mis võimaldab kasutada välisseadmena värviteleviisorit. Numbriliste skeemide funktsioonid kujutatakse graafiliselt televiisori ekraanil. Samuti vaartustatakse loodud loogikaskeemide sisendsuurused ja kontrollitakse tulemuse vastavust ülesandele.

Programm on kirjutatud algoritmilises keeles Pascal.

КОМПАРАТОРЫ С ОБРАТНОЙ СВЯЗЬЮ

В настоящей работе исследуются компараторы с обратной связью. Особое внимание уделяется передаточным характеристикам компаратора. Представляется новая методика исследования этих схем, которая намного упрощает анализ.

Компаратор представляет собой схему, предназначенную для сравнения двух сигналов (φ_a и φ_b). Один из них является опорным, а другой исследуемым сигналом. Выходной сигнал может принимать два значения ($\varphi_{вын}$ и $\varphi_{выл}$) в зависимости от того, какое из соотношений сигналов φ_a и φ_b будет выполняться. Указанное наглядно представлено на рис. 1.

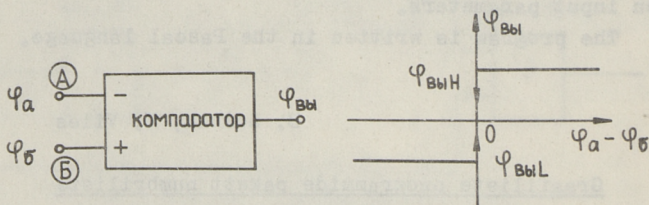


Рис. 1. Схема и передаточная характеристика компаратора.

Соответственно передаточная функция идеального компаратора аналитически выражается в виде формулы:

$$\varphi_{вы} = \varphi_{вын} \cdot I(\varphi_a - \varphi_b) + \varphi_{выл} \cdot I(\varphi_b - \varphi_a), \quad (I)$$

где $I(\varphi_a - \varphi_b)$ — представляет собой функцию единичного скачка.

По-видимому, опрокидывание выходного знака происходит при нулевом переходе значения $\varphi_a - \varphi_b$.

Один из двух входов называется прямым, а другой инвертирующим. Обозначаются они знаками + и -. Для их идентификации по желанию на один из них подается нуль, а на другой сигнал с каким-нибудь конечным значением. Поскольку положительное значение этого сигнала соответствует $U_{\text{выл}} = U_{\text{вын}}$ (отрицательное $U_{\text{выл}} = U_{\text{выл}}$), то данный вход является неинвертирующим, а другой инвертирующим. Следовательно, вход А является прямым, а вход Б инвертирующим.

Для сравнения напряжений применяются операционные усилители ($A > 80$ дБ). Идеальная характеристика компаратора (рис. 1) может быть реализована лишь с помощью операционного усилителя неограниченным разомкнутым циклом. Следствием ограниченного усиления является, то что при $U_{\text{д}} = U_{\text{с}}$ компаратор будет функционировать линейным усилителем, поэтому скорость изменения выходного сигнала зависит от формы входного сигнала и степени управления.

Усиление передаточной системы с помощью обратной связи выражается формулой:

$$A_{\text{ув}} = \frac{A_{\text{уо}}}{1 + H}, \quad (2)$$

где

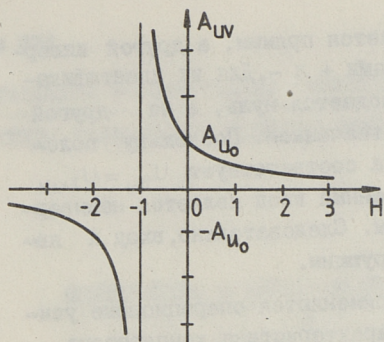
$A_{\text{уо}}$ — усиление напряжения системы без обратной связи;
 H — усиление разомкнутого контура со знаком.

На рисунке 2 показаны зависимость усиления разомкнутого контура от коэффициента передачи усилителя с обратной связью и влияние усиления разомкнутого контура на передаточную характеристику. При получении передаточной характеристики предполагалось, что выходное напряжение изменится в пределах от $U_{\text{вын}}$ до $U_{\text{выл}}$ определяемых схемы.

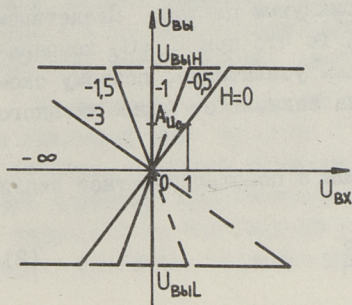
В качестве компаратора в дальнейшем рассматривается усилитель с положительной обратной связью, показанный на рис. 3.

Опрокидывание выходного уровня у компараторов без обратной связи происходит при превышении напряжения $U_{\text{выс}}$. Это действительно и в нашем случае. В качестве условия опрокидывания на вход Б подается сигнал, имеющий вид:

$$U_{\text{с}} = U_{\text{д}} \frac{aR}{(1+a)R} + U_{\text{выл}} \frac{R}{(1+a)R}, \quad (3)$$



а)



б)

Рис. 2. Зависимость усиления разомкнутого контура от коэффициента передачи усилителя с обратной связью и влияние усиления разомкнутого контура на передаточную характеристику.

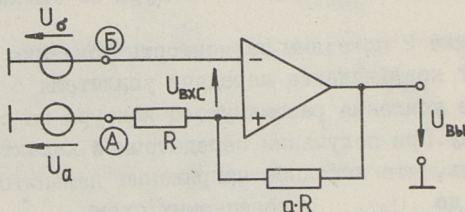


Рис. 3. Компаратор с обратной связью.

где a постоянная, определяемая сопротивлением обратной связи.

Выражение $U_{Вы}$ имеет вид

$$U_{Вы} = (1 + a)U_{\delta} - aU_a. \quad (4)$$

Это выражение справедливо лишь между напряжениями $U_{ВыН} \geq$

$U_{ВыЛ} \leq U_{Вы}$. Ограничивание выходного напряжения выполняется с помощью операционного усилителя.

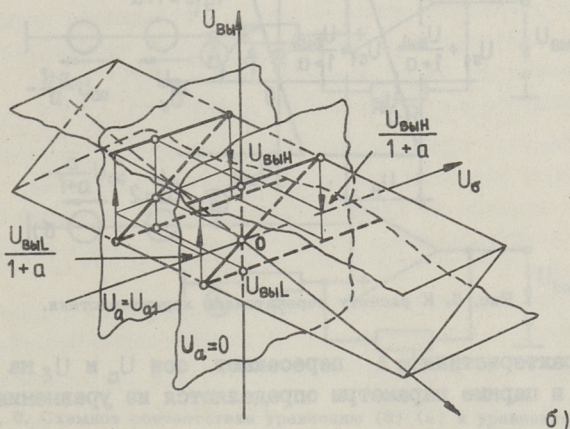
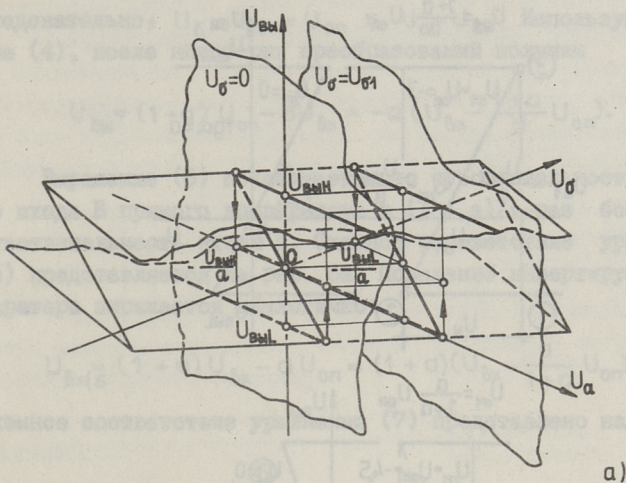
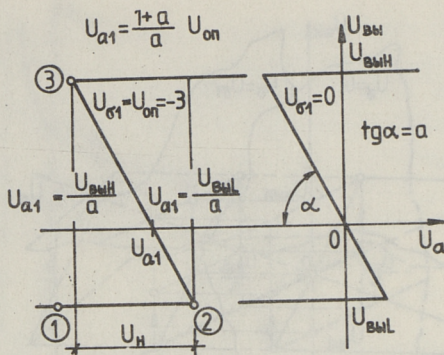
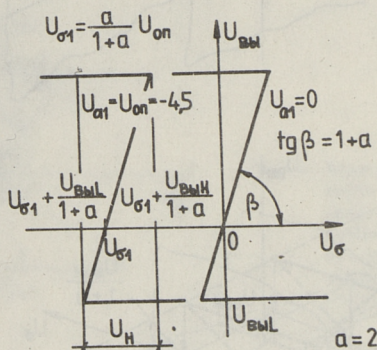


Рис. 4. Аксонометрическое изображение функции двух переменных для уравнения (4).

На рис. 4 представляется аксонометрическое изображение функции двух переменных (4). Рис. 4а иллюстрирует ситуацию для постоянной U_{σ} , а на рисунке 4б показана ситуация, где постоянная U_{α} . По форме кривой функциональную связь опрокидывания в дальнейшем будем называть характеристикой $\bar{z}$. Одна или другая переменная в формуле (4) называется параметром. Характеристика опрокидывания представляется на плоскости, как показано на рис. 5.



а)



$a=2$

б)

Рис. 5. К расчету передаточной характеристики.

Характеристики z пересекают оси U_a и U_δ на месте $U_{вх} = 0$, и парные параметры определяются из уравнения:

$$(1+a) U_{\delta 1} = a U_{a 1}, \quad (5)$$

используя и выражение (4). На рисунках видно, что передаточная функция перемещается при изменении параметра параллельно себе. На рис. 5а $U_a = U_{вх}$ и $U_\delta = U_{он}$ — схема прямая, а $U_a = U_{он}$ и $U_\delta = U_{вх}$ — схема инвертирующая.

Рассмотрим физическое-схемное содержание параметров с помощью рис. 5а.

Как выше установлено, на вход Б прямого компаратора поступает опорный сигнал, а на А — исследуемый (см. рис. 3).

Следовательно, $U_{\delta} = U_{\delta 1} = U_{оп}$ и $U_a = U_{bx}$. Используя выражение (4), после некоторых преобразований получим

$$U_{bbl} = (1+a) U_{оп} - a U_{bx} = -a \left(U_{bx} - \frac{1+a}{a} U_{оп} \right). \quad (6)$$

Выражение (6) показывает, что чувствительность опорного входа В прямого компаратора в $(1+a)/a$ -раз больше, чем чувствительность входа А. Схемное соответствие уравнению (5) представлено на рис. 6а. Поведение инвертируемого компаратора выражается аналогично:

$$U_{bx} = (1+a) U_{bx} - a U_{оп} = (1+a) \left(U_{bx} - \frac{a}{1+a} U_{оп} \right). \quad (7)$$

Схемное соответствие уравнению (7) представлено на рис. 6б.

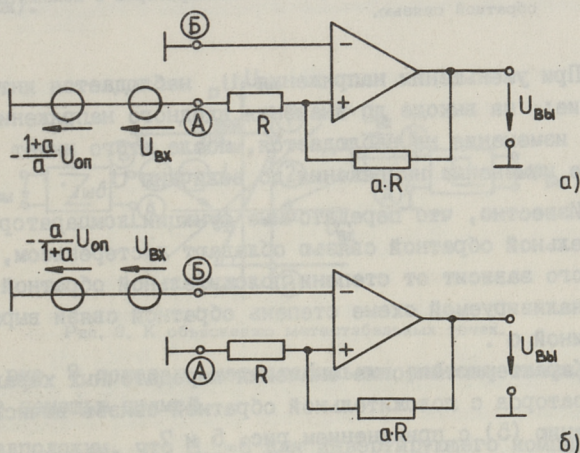


Рис. 6. Схемное соответствие уравнению (6) (а) и уравнению (7) (б).

До сих пор еще не говорилось о том, что рабочие точки, находящиеся на наклонной линии передаточных характеристик, представляют метастабильные состояния схемы компаратора. Это значит, что при увеличении входного напряжения от точки 1 до точки 2 (рис. 5) выходное напряжение имеет постоянную величину U_{bbl} . При дальнейшем увеличении напряжения U_a выходное напряжение достигает величины U_{bblH} , после чего выходное напряжение опять остается неизменным. Подобным образом ведет себя и инвертирующий компаратор. На рис. 7

показаны передаточные характеристики, которые соответствуют стабильным значениям рабочих точек компаратора.

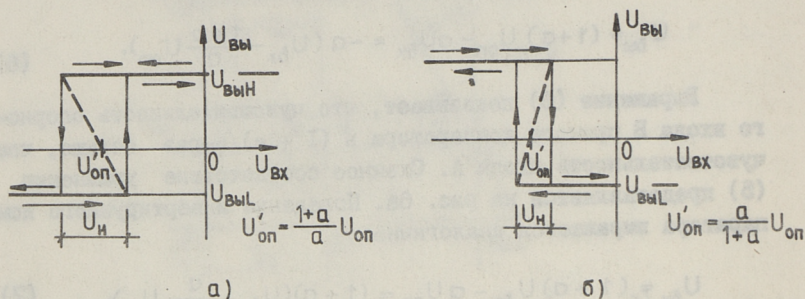


Рис. 7. Передаточные характеристики компараторов с положительной обратной связью.

При уменьшении напряжения U_a наблюдается интересное явление: на выходе до значения входного напряжения (точка 3) изменения не наблюдается, после этого имеет место резкое изменение напряжения до величины $U_{\text{выЛ}}$.

Известно, что передаточные функции компараторов с положительной обратной связью обладают гистерезисом, ширина которого зависит от степени положительной обратной связи. При анализируемой схеме степень обратной связи выражается величиной a .

Характеристические значения передаточных характеристик компараторов с положительной обратной связью вычисляются по выражению (5) с применением рис. 5 и 7.

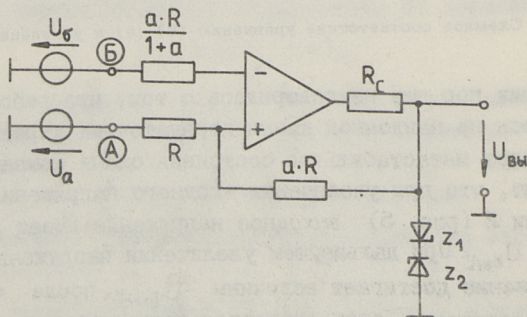


Рис. 8. Компаратор с дополнительными элементами.

На рис. 8 представляются дополнительные элементы схемы. Сопротивление, соединенное с инвертирующим входом операционного усилителя последовательно, служит для компенсации тока входной рабочей точки. На выходе схемы имеется стабилизатор уровня. Он увеличивает скорость переключения.

Компараторы с положительной обратной связью имеют следующие преимущества перед компараторами без обратной связи:

1) опрокидывание происходит всегда с максимальной скоростью изменения выходного сигнала независимо от скорости изменения входного сигнала;

2) гистерезис обеспечивает защиту от шума (ε характеристика).

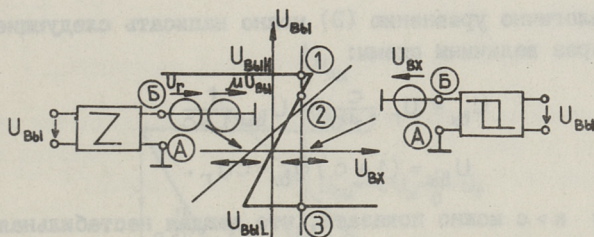


Рис. 9. К объяснению метастабильных точек.

На рис. 9 показаны метастабильные рабочие точки, связанные с помощью прямой.

Предположим, что $U_{оп} = 0$ для инвертирующего компаратора. Прямая $U_{вх} = \text{const}$ пересекает ε характеристику в трех точках, из которых точка 2 нестабильная, а точки 1 и 3 стабильные.

Дальше предположим, что на вход компаратора попадает входной сигнал следующего вида (например, электромоторная сила датчика):

$$U_{вх} = U_{\Gamma} + \mu U_{вЫ}, \quad (8)$$

где μ постоянная.

ε характеристика может представляться, например, на экране осциллографа.

Один из вариантов практической реализации управления с помощью уравнения (8) показан на рис. 10.

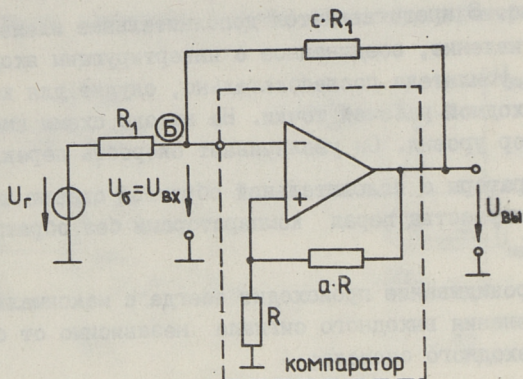


Рис. 10. Пример реализации управления с компаратора.

Аналогично уравнению (3) можно написать следующие уравнения через величины схемы:

$$U_{вх} = U_r \frac{c}{1+c} + U_{вы} \frac{1}{1+c}, \quad (9)$$

откуда

$$U_{вы} = (1+c) U_{вх} - c U_r. \quad (10)$$

При $a > c$ можно показать, что каждая неустойчивая точка является стабильной, а при $a = c$ передача не достигает гистерезиса.

Характеристика прямого компаратора с положительной обратной связью определяется аналогично вышеприведенному примеру. Проблема усложняется тогда, когда из-за фазовых соотношений в управляющую ветвь включается аналоговый инвертор. Практическое решение показано на рис. 11.

Выходное напряжение инвертирующего сумматора, осуществляемое с помощью АИ, достигает напряжения на входе А прямого компаратора по закону

$$U_a = U_{вх} = -\frac{U_r}{\delta} - \frac{U_{вы}}{c}, \quad (11)$$

откуда

$$U_{вы} = -c U_a - \frac{c}{\delta} U_r. \quad (12)$$

Результаты показаны на рис. 12, где крутость прямой равняется c , а ось U_r пересекается в точке $-U_r/\delta$.

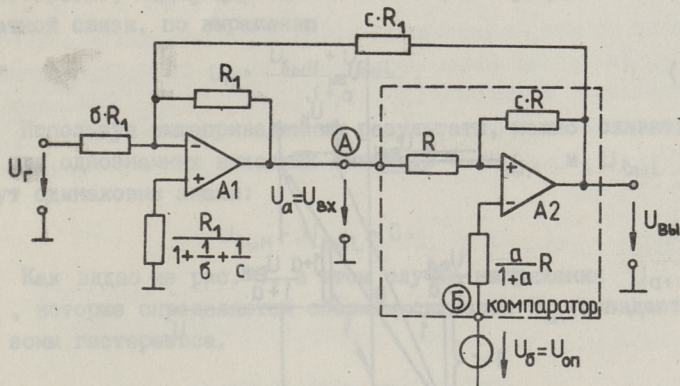


Рис. 11. Пример реализации управления в случае аналогового инвертора на входе.

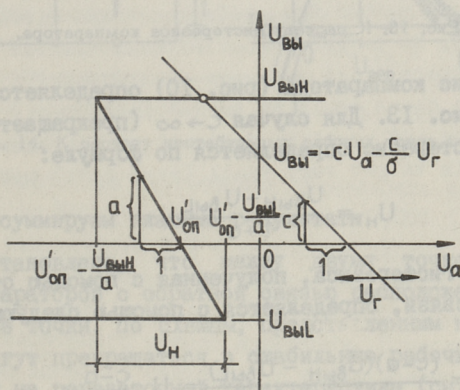


Рис. 12. К расчету передаточной характеристики.

Аналогично вышеустановленному, каждая рабочая точка z характеристики является стабильной, если прямая, определенная передаточной функцией (4) и выражением (12), имеет только одну точку сечения. В данном случае это реализуется тогда, когда $a \geq c$. Получается новый и интересный результат в случае, если управляющим сигналом считать U_r . Для $a \leq c$ величина гистерезиса зависит только от c (a является данным). Следовательно, для $a = c$ компаратор не обладает гистерезисом.

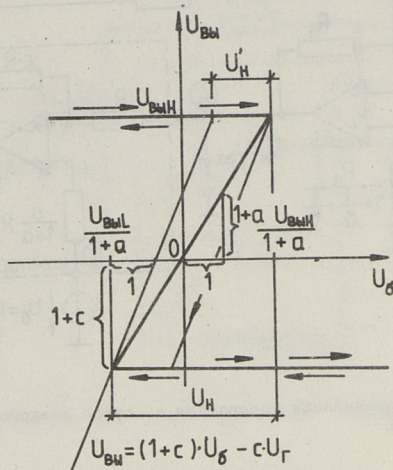


Рис. 13. К расчету гистерезиса компаратора.

Гистерезис компаратора (рис. 10) определяется, как это показано на рис. 13. Для случая $c \rightarrow \infty$ (прекращается обратная связь) гистерезис определяется по формуле:

$$U_H = \frac{U_{ВнH} - U_{ВнL}}{1+a}. \quad (I3)$$

Величина гистерезиса, полученная с помощью отрицательной обратной связи, определяется с помощью следующего уравнения:

$$U'_H = \frac{(c-a)(U_{ВнH} - U_{ВнL})}{(1+a)(1+c)} = U_H \frac{c-a}{1+c}. \quad (I4)$$

Из выражения (I4) при $c=a$ $U'_H = 0$.

Выражения, аналогичные предыдущему, могут представляться и для компаратора, превращающего фазу, показаны на рис. II.

Учитывая условие $c \geq a$ и используя методику, показанную на рис. 12, величина гистерезиса напряжения определяется для $c \rightarrow \infty$ с помощью следующей формулы (аналогично формуле (I3)):

$$U_H = \frac{U_{ВнH} - U_{ВнL}}{a}, \quad (I5)$$

а гистерезис, модифицированный с помощью отрицательной обратной связи, по выражению

$$U_H = \frac{U_{выН} + U_{выЛ}}{1 + \alpha} \quad (I6)$$

Используя вышеприведенные результаты, можно сказать, что для однозначных выходных сигналов ($U_{выН}$ и $U_{выЛ}$) будут одинаковые знаки:

$$U_{выН} \cdot U_{выЛ} > 0. \quad (I7)$$

Как видно из рис. 5, в этом случае напряжения $U_{\alpha 1}$ и $U_{\delta 1}$, которые определяются опорным сигналом $U_{оп}$, попадают вне зоны гистерезиса.

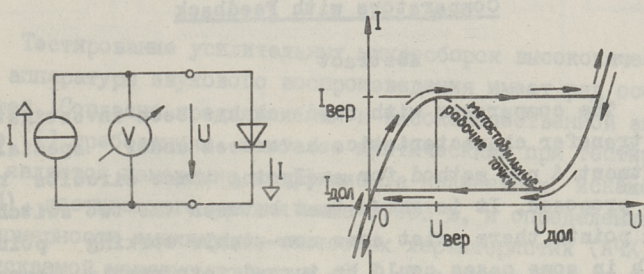


Рис. 14. К расчету нестабильных рабочих точек компаратора.

Итак, суммируем главные результаты.

1. Установлено, что между двумя точками опрокидывания компараторов с обратной связью расположены нестабильные рабочие точки. По схемам, представленным на рис. I0 и II, они могут превращаться в стабильные рабочие точки и находиться на передаточной характеристике (см. также схему измерения на рис. I4).

2. При совместном применении положительной и отрицательной обратной связи можно получить как угодно малые (даже равные нулю) гистерезисы так, что при опрокидывании скорость изменения выходного сигнала, аналогично триггеру Шмитта, зависит только от параметра "slow-rate" операционного усилителя.

3. Величины напряжений $U_{\alpha 1}$ и $U_{\delta 1}$, определяемые опорным напряжением $U_{оп}$ (5), представляют собой точки пересечения

чения прямой, соединяющей точки опрокидывания с осью U_d и U_{δ} . Указанная точка пересечения имеет гистерезис, пропорциональный отношению значений $U_{\text{вын}}$ и $U_{\text{выл}}$. При однозначных выходных напряжениях (при выполнении условия $U_{\text{вын}} \cdot U_{\text{выл}} > 0$) значения U_{d1} и $U_{\delta 1}$ попадают вне гистерезиса.

S. Gallai, T. Rang

Comparators with Feedback

Abstract

The comparator with feedback has been investigated. The transfer characteristics have been under special treatment. A new method for analysing these circuits has been proposed. It is shown that between the two switch-over points there exist some non-stable working points which in some cases could be turned into stable ones. Using negative and positive feedback at the same time it is possible to get very small (even zero) hysteresis.

S. Gallai, T. Rang

Tagasisidestatud komparaatorid

Kokkuvõte

Uuritakse tagasisidestatud komparaatori ülekandekarakteristikut. Esitatakse uudne analüüsimeetod ülekandekarakteristiku saamiseks. Näidatakse, et kahe ümberlülitamispunkti vahel eksisteerib veel kolmas - mittestabiilne. Positiivse ja negatiivse tagasiside üheaegsel kasutamisel on võimalik ümberlülitamispunktide vahelise piirkonna ulatust suvaliselt muuta.

А. Лаансоо, В. Мяннама, О. Пикков

ИЗМЕРЕНИЕ ПАРАМЕТРОВ АНАЛОГОВЫХ МИКРОСБОРОК
ПРИМЕНЕНИЕМ КОМПЕНСАЦИОННОГО МЕТОДА

Тестирование усилительных микросборок высококачественной аппаратуры звукового воспроизведения имеет ряд особенностей. Согласно предъявляемым к высококачественной аппаратуре [1] требованиям наиболее критическими при тестировании являются измерение коэффициента нелинейных искажений (КНИ), достигающего уровня ниже 0,003 %, и определение неравномерности амплитудно-частотных характеристик (АЧХ) с допускаемой неравномерностью $\pm 0,1$ дБ в диапазоне частот 20 Гц ... 25 кГц.

Соответственно разрешающая способность измерительной аппаратуры по КНИ должна быть не ниже 0,001 %, а изменение показаний измерителей АЧХ в диапазоне частот не должно быть выше $\pm 0,3$ %. Как следует из материалов работ [2, 3] для достижения этой цели наиболее рациональным является применение компенсационного метода измерения. В настоящей работе рассмотрены вопросы, связанные с использованием этого метода в составе автоматизированного измерительного комплекса (АИК) для тестирования микросборок.

Сущность компенсационного метода иллюстрирована на рис. 1. Сигнал от генератора Г подается через делитель напряжения Д на тестируемую микросборку МС, на один из входов сумматора С и на вход синхронизации ВВ. На второй вход С подается выходной сигнал от МС. Схему С и коэффициент передачи Д выбирают так, чтобы на выходе С получить достаточную компенсацию основной гармоник сигнала (если МС неинвертирующий, то один из входов С должен быть инвертирующим). Сигнал на выходе С при этом будет пропорционален высшим

гармоникам, возникающим в тракте Д-МС-С, и не будет зависеть от содержания высших гармоник испытательного сигнала на выходе Г, так как при малых частотных искажениях тракта Д-МС будет компенсироваться и их влияние. Тем самым разрешающая способность компенсационной схемы при измерении КНИ будет определяться линейностью Д и С, остаточным содержанием гармоник на выходе Г и параметрами ВВ, которые в данном случае используются в режиме анализатора гармоник [4].

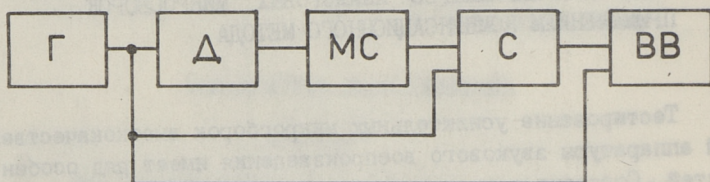


Рис. 1. Компенсационная схема измерения параметров:
Г - генератор; Д - делитель напряжения;
МС - микросборка; С - сумматор; ВВ - вектор-
ный вольтметр.

Дополнительно фиксируя коэффициент деления Д на средней частоте (1 кГц), можно в дальнейшем определить частотные искажения измерением уровня основной гармоники сигнала частоты f на выходе С, которая в данном случае прямо пропорциональна неравномерности АЧХ МС. Применение ВВ в качестве измерителя оправдано и в этом случае, так как дополнительная информация о фазе позволяет полностью восстановить величину выходного напряжения МС.

Преимуществом применения компенсационной схемы измерения является существенное понижение требований к точности поддержания уровня выходного напряжения Г.

Особенность применения компенсационной схемы в составе АИК для тестирования МС заключается в том, что с целью упрощения аппаратуры в [1] используются делители Д с фиксированными коэффициентами деления для каждого типа МС. Коэффициенты деления Д выбраны по номинальному входному сигналу МС данного типа. Из-за разброса коэффициентов передачи отдельных МС (для усилительных МС не более $\pm 5\%$)

$$|\delta_k| \leq 0,05 \quad (I)$$

Измерительная схема не позволяет получить полную компенсацию основной гармоники сигнала на выходе сумматора. Однако полученный выигрыш даже в 20 раз вполне достаточен и позволяет применить в качестве Г и ВВ стандартной аппаратуры для измерения сверхмалых значений КНИ.

Наличие в составе АИК контроллера (микроЭВМ) позволяет учитывать меру неполной компенсации основной гармоники сигнала программным способом. При определении неравномерности АЧХ конкретной ИС сначала запоминается погрешность коэффициента передачи МС на средней частоте δ_k , а затем неравномерность АЧХ определяется вычислением коэффициента частотной погрешности δ_f .

Погрешность коэффициента передачи МС:

$$\delta_k = \frac{U_0}{U_n} - 1 = \sqrt{\left(1 + \frac{\Delta U_{OR}}{U_n}\right)^2 + \left(\frac{\Delta U_{OI}}{U_n}\right)^2} - 1, \quad (2)$$

где U_n - номинальное (идеальное, заданное) выходное напряжение МС данного типа;

U_0 - реальное выходное напряжение данного экземпляра МС на средней частоте;

ΔU_{OR} , ΔU_{OI} - вещественная и мнимая составляющие выходного напряжения С на средней частоте.

В машинном расчете δ_k фиксируется как постоянная для каждого МС. Отклонение АЧХ на частоте f от фиксированного на средней частоте значения:

$$\begin{aligned} \delta_f = \frac{U_f}{U_0} - 1 &= \frac{\sqrt{\left(1 + \frac{\Delta U_{fR}}{U_n}\right)^2 + \left(\frac{\Delta U_{fI}}{U_n}\right)^2}}{\delta_k + 1} - 1 \approx \\ &\approx \sqrt{\left(1 + \frac{\Delta U_{fR}}{U_n}\right)^2 + \left(\frac{\Delta U_{fI}}{U_n}\right)^2} - 1 - \delta_k, \end{aligned} \quad (3)$$

где ΔU_{fR} , ΔU_{fI} - вещественная и мнимая составляющие выходного напряжения С на частоте f .

Векторная диаграмма напряжений приведена на рис. 2.

Весьма существенным при реализации АИК являлся выбор схемы сумматора С. Сравнению подвергались схемы с параллельной и последовательной обратными связями, построенные на базе ОУ типов К140УД7, К140УД8А, КР544УД2А, 153УД2, 153УД6 и

К574УД1А. Экспериментальное исследование схем С проводилось согласно рис. 3, т.е. в схеме на рис. 1 тракт Д-МС был заменен коротким замыканием. Частотные искажения определялись по уровню основной гармоники выходного сигнала С согласно (2), нелинейные искажения – по уровню высших гармоник.

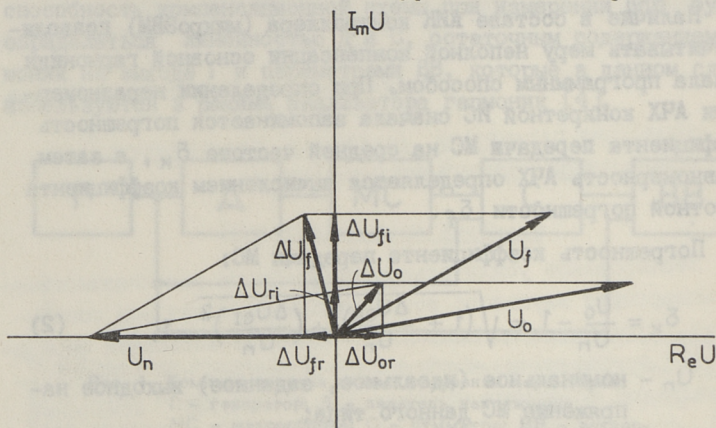


Рис. 2. Векторная диаграмма напряжений:

- U_n – номинальное выходное напряжение МС;
- U_o – реальное выходное напряжение МС на средней частоте;
- U_f – выходное напряжение МС на частоте f ;
- ΔU_o – выходное напряжение сумматора на средней частоте;
- ΔU_f – выходное напряжение сумматора на частоте f .

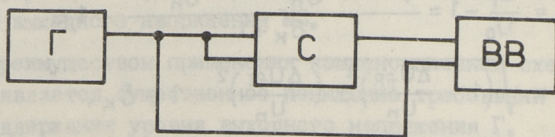


Рис. 3. Экспериментальное исследование сумматора.

Наилучшие результаты для С были получены схемой с параллельной обратной связью, собранной на базе двух ОУ КР544УД2А (рис. 4). Исследование схемы С дало следующие результаты: частотные искажения были меньше 0,1 % в диапазоне частот от 20 Гц до 50 кГц, а высшие гармоники находились

на уровне не более 0,0008 % от входного сигнала U для частот от 40 Гц до 16 кГц и при входных напряжениях от 0,5 до 5 В.

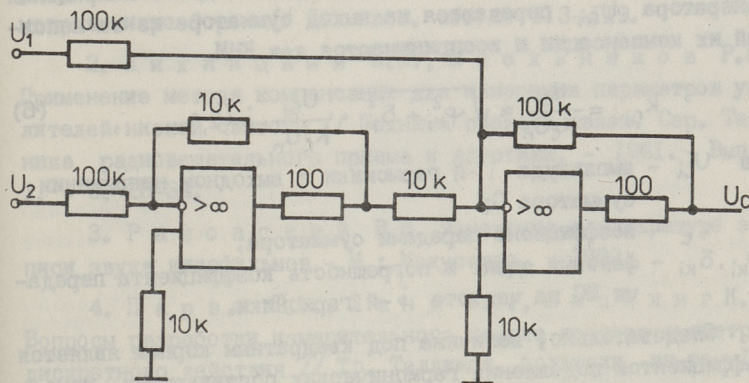


Рис. 4. Сумматор.

Напряжение U_c пропорционально разности $(U_2 - U_1)$.

Отметим, что указанный метод – замена исследуемой МС и делителя D коротким замыканием – может быть применен и при проверке нуля всего тракта АИК. Погрешности, вызванные параметрами генератора Γ и векторного вольтметра ВВ, могут быть оценены анализом схемы на рис. 1. Приведем здесь лишь окончательные результаты.

Нестабильность амплитуды выходного напряжения генератора вызывает относительную погрешность определения коэффициента передачи МС величиной

$$\delta_{kG} \approx \delta_k \cdot \delta_G, \quad (4)$$

где δ_G – относительная нестабильность U_G .

Погрешность по амплитуде векторного вольтметра ВВ вызывает относительную погрешность определения δ_k величиной

$$\delta_{kVV} \approx \delta_k \cdot \delta_{VV}, \quad (5)$$

где δ_{VV} – относительная погрешность измерения выходного напряжения сумматора U_c векторным вольтметром ВВ.

Фазовая погрешность ВВ влияет незначительно, так как согласно (2) квадратичная составляющая ΔU_{01} сама является погрешностью.

Высшие гармонические составляющие выходного напряжения генератора U_{Gi} передаются на выход сумматора из-за неполной их компенсации и воспринимаются как КНИ.

$$k_{Ghi} = \frac{U_{Ci}}{K_C U_n} \approx \sqrt{\varphi_{ki}^2 + \delta_{ki}^2} \cdot \frac{U_{Gi}}{K_C U_n}, i=2,3,\dots, \quad (6)$$

где U_{Ci} - амплитуда i -й гармоники в выходном напряжении сумматора C ;

K_C - коэффициент передачи сумматора;

$\varphi_{ki}, \delta_{ki}$ - фазовый сдвиг и погрешность коэффициента передачи МС на частоте i -й гармоники.

Следовательно, величина под квадратным корнем является коэффициентом подавления гармонических составляющих генератора. С другой стороны, погрешность нуля ВВ при использовании его в режиме анализатора гармоник воспринимается как КНИ МС величиной

$$k_{vvhi} = \frac{U_{vvi}}{K_C U_n}, \quad (7)$$

где U_{vvi} - показание ВВ при отсутствии i -й гармоники на его входе.

Используя оценку

$$\varphi_{ki} \leq 0,05, \quad (8)$$

получим на основе (1) и (4) - (7) следующие требования к свойствам Г и ВВ, обеспечивающие выполнение приведенных в начале работы условий к параметрам АИК:

- генератор должен обладать нестабильностью выходного напряжения не хуже $\pm 3\%$ и собственным КНИ выходного напряжения не выше $0,01\%$;

- векторный вольтметр должен иметь погрешность по амплитуде основной гармоники сигнала не хуже $\pm 3\%$ и погрешность нуля в режиме анализатора гармоник не более $0,01\%$ от уровня основной гармоники сигнала на его входе. Как следует из [1] и из паспортных данных приборов, подходящим генератором является ГЗ-118, а векторный вольтметр ТВ5-80А соответствует требованиям выходного измерительного прибора АИК.

Л и т е р а т у р а

1. Лаансоо А.А., Мяннамаа В.Р., Пикков О.М. Вопросы разработки коммутационно-согласующего устройства автоматизированного комплекса измерения параметров микросборок // См. наст. сб., с. II3-II9.

2. Лихницкий А.М., Школьников Р.М. Применение метода компенсации для измерения параметров усилителей низкой частоты // Техника средств связи. Сер. Техника радиовещательного приема и акустики. - 1981. - Вып. I. - С. 25-34.

3. Раковский В.В. Измерения в аппаратуре записи звука кинофильмов. - М.: Искусство, 1962.

4. Парве Т.Э., Ланд Р.Э., Липпинг К.К. Вопросы разработки измерительного канала векторвольтметров дискретного действия // Тр. Таллинск. политехн. ин-та. - 1984. - № 583. - С. 17-28.

A. Laansoo, V. Männama, O. Pikkov

Measuring Parameters of the Analog Microcircuits Using the Compensational Method

Abstract

In this paper the problems of computer-aided test and measurement of the very low harmonic distortion and frequency response of the high-quality audio amplifier microcircuits have been discussed.

A distortion-rejection method combined with the use of vector analyzer TB5-80A has been presented.

As a result of this study new practical solution has been recommended for use in the automatic testing system.

A. Laansoo, V. Männama,
O. Pikkov

Analoog-mikroskeemide parameetrite mõõtmine
kompensatsioonmeetodiga

Kokkuvõte

Käsitletakse väga väikeste mittelineaar- ja sagedus-moonutuste automaatse mõõtmise ning testimisega seoses olevaid probleeme.

Analüüsitakse kompensatsioonmeetodi ja vektoranalüsaatori TB5-80A kasutamise võimalusi.

Saadud tulemusi soovitatakse praktiliselt kasutada automaatse testsüsteemi kavandamisel.

А. Лаансоо, В. Мяннама, О. Пикков

ВОПРОСЫ РАЗРАБОТКИ КОММУТАЦИОННО-СОГЛАСУЮЩЕГО УСТРОЙСТВА АВТОМАТИЗИРОВАННОГО КОМПЛЕКСА ИЗМЕРЕНИЯ ПАРАМЕТРОВ МИКРОСБОРОК

Производство микросхем и микросборок (МС), а также их применение в высококачественной электронной аппаратуре предполагает всестороннее тестирование таких элементов для установления соответствия их характеристик и параметров к предъявляемым требованиям. В условиях массового производства и применения МС их тестирование с приемлемой точностью и скоростью может быть осуществлено лишь применением высокопроизводительных автоматизированных измерительных комплексов (АИК), которые обеспечивают также обработку и оформление данных тестирования [1 — 3].

Блок-схема АИК для тестирования МС представлена на рис. 1. Он включает в себя измерительные приборы (ИП) и генераторы испытательных сигналов (ГС), снабженные интерфейсами канала общего пользования (КОП), микроЭВМ для управления устройствами АИК и обработки данных тестирования, и коммутационно-согласующее устройство (КСУ), которое обеспечивает требуемые режимы измерения тестируемых МС, переключение этих режимов при тестировании разных характеристик и параметров, а также коммутацию внешних источников сигнала и измерительных приборов.

АИК для тестирования некоторого класса МС комплектуется, как правило, из серийных (стандартных) ГС, ИП, микроЭВМ, а для учета конструктивной и схемотехнической специфики конкретных МС, их характеристик, параметров и режимов тестирования разрабатываются специальное КСУ и программное обеспечение ЭВМ.

Ниже рассмотрены некоторые вопросы разработки КСУ

автоматизированного измерительного комплекса, предусмотренного для тестирования микросборок полного усилителя звуковой частоты высшего класса.

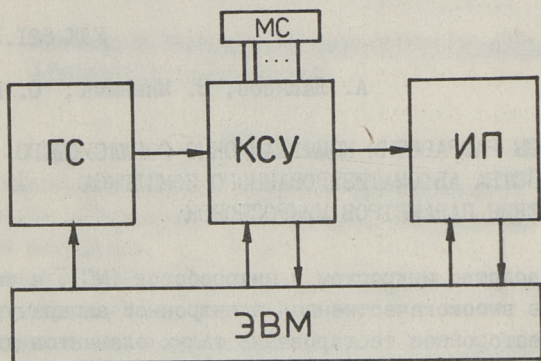


Рис. 1. Блок-схема автоматизированного измерительного комплекса:
КСУ - коммутационно-согласующее устройство;
ГС - генераторы сигналов; ИП - измерительные приборы;
МС - микросхема (микросборка).

Комплект тестируемых МС состоит из следующих элементов:

- эмиттерного повторителя;
- усилителя-корректора;
- блока фильтров;
- усилителя напряжения;
- усилителя-ограничителя;
- усилителя темброблока;
- узла индикации;
- модуля защиты;
- стабилизатора напряжения +40 В;
- стабилизатора напряжения -40 В.

Перечисленные МС делятся на усилительные МС (первые шесть элементов) и на МС специального назначения (остальные). Усилительные МС тестируются по следующим показателям:

- току потребления;
- напряжению смещения на выходе МС;
- коэффициенту передачи на частоте 1 кГц;
- амплитудно-частотной характеристике (АЧХ);
- входному сопротивлению;

- переходному затуханию между каналами;
- уровню шума на выходе микросборок;
- коэффициенту нелинейных искажений выходного сигнала.

Для остальных МС определяют:

- ток потребления;
- весовые напряжения МС индикации;
- пороговые напряжения срабатывания модуля защиты;
- выходное напряжение и сопротивление, коэффициент стабилизации и напряжение помех стабилизаторов напряжения.

Наиболее серьезные измерительные проблемы связаны с определением сверхмалых уровней коэффициента нелинейных искажений усилительных МС (порядка $3 \dots 5 \cdot 10^{-3} \%$), при котором требуется измерение микровольтных напряжений высших гармонических составляющих выходного сигнала. Относительно точные измерения требуются также при определении амплитудно-частотных характеристик МС.

Из совокупности тестируемых показателей представленного выше комплекса МС, а также электрических и конструктивных особенностей МС непосредственно вытекают требования к КСУ АИС, которое должно обеспечить:

- заданный статический и динамический режим работы тестируемой МС, осуществляемый путем подключения к выводам МС предусмотренных предписанием тестирования пассивных и активных элементов, источников питания и т.д.;
- коммутацию режимных элементов МС при проведении конкретных измерений;
- коммутацию внешних измерительных приборов и источников испытательного сигнала;
- необходимое масштабирование испытательных сигналов и результатов измерений, а также первичную обработку измеряемого сигнала для достижения заданной точности и чувствительности измерений;
- помехозащищенность измерительных цепей;
- необходимую универсальность АИС для тестирования всех МС данного комплекта;
- высокую производительность тестирования МС.

Исходя из отмеченных требований для КСУ разработаны следующие основные схмотехнические и конструктивные решения [4, 5].

1. Измерение коэффициента нелинейных искажений проводится компенсационной схемой [6], причем в качестве внешнего ИП используется векторный вольтметр, а для подавления основной гармонической сигнала в цепь измерения включены активные фильтры. Компенсационная схема предусмотрена также для определения отклонения АЧХ от номинальной,

2. Специфические для каждой МС режимы тестирования, а также коммутация таких режимов обеспечиваются специальной коммутационно-согласующей схемой, собранной на плате смежного адаптера (СА), к которой прикреплено и измерительное гнездо для подключения данной МС к АИС.

3. С помощью разъема СА соединен с блоком коммутации и обработки сигнала (БКО), который обеспечивает все соединения СА с внешними устройствами (ГС, ИП, ЭВМ) и блоком питания КСУ, коммутацию внешних устройств, первичную обработку измерительного сигнала с помощью компенсационной схемы и режекторных фильтров, масштабирование испытательных и измерительных сигналов.

4. Управление переключателями (реле) происходит через адаптер канала общего пользования (КОП) или через собранную на плате БКО логическую схему, соединенную с шиной управления ЭВМ.

Блок схема КСУ представлена на рис. 2.

Все узлы КСУ смонтированы в общий металлический корпус, к которому прикреплены разъемы и гнезда шин и кабелей для соединения внешних устройств и который обеспечивает надежную защиту от внешних помех.

В комплект СА входит 10 адаптерных плат с разными измерительными гнездами для соответствующих МС. Расположение платы СА в КСУ и конструктивное решение измерительного гнезда обеспечивают универсальность разработанного устройства, быструю замену тестируемых МС и их экранировку, а также несложную замену СА в целом.

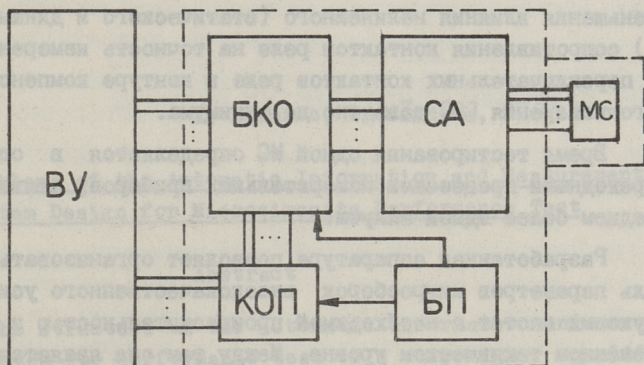


Рис. 2. Блок-схема КСУ.

СА - схемный адаптер;
 БКО - блок коммутации и обработки сигнала;
 КОП - интерфейс канала общего пользования;
 БП - блок питания;
 ВУ - внешнее устройство.

На плате БКО располагаются реле для коммутации четырех внешних генераторов сигнала, двух выходных измерительных приборов и внутренних цепей БКО, пять масштабирующих усилителей испытательных и измерительных сигналов, цепь обработки измерительных сигналов, которая включает компенсационную схему на высококачественных ОУ и активные режекторные фильтры. К плате БКО подключены интерфейс КОП и блок питания (БП) КСУ. Последний имеет двухполярные выходные напряжения для питания МС и ОУ, а также однополярные напряжения для питания логических вентилей и реле.

В схемах СА, БКО и БП имеются защитные элементы и узлы, предотвращающие выход из строя КСУ в случае неисправных МС и внешних устройств.

Помехозащищенность внутренней измерительной схемы достигнута тщательной экранировкой измерительных цепей, отсутствием общих проводов питания логических и измерительных цепей и низким уровнем пульсации напряжений БП.

Для измерения сверхмалых коэффициентов нелинейных искажений МС требуется высокая линейность всей измерительной

цепи, включая и элементы переключения (реле) цепи. С целью уменьшения влияния нелинейного (статического и динамического) сопротивления контактов реле на точность измерений число переключаемых контактов реле в контуре компенсационного измерения [6] доведено до минимума.

Время тестирования одной ИС определяется в основном переходными процессами измерительных приборов, составляя в среднем более одной минуты.

Разработанная аппаратура позволяет организовать контроль параметров микросборок высококачественного усилителя звуковых частот с необходимой производительностью и на современном техническом уровне. Между тем она является достаточно универсальной, ее можно применять также для контроля и измерения параметров узлов радиоэлектронной аппаратуры другого назначения.

Л и т е р а т у р а

1. Страх ов А.Ф. Автоматизированные измерительные комплексы. - М.: Энергоиздат, 1982. - 216 с.

2. Бай да Н.П., Ку з ь м и н И.В., Ш п и л е - в о й В.Т. Микропроцессорные системы элементарного диагностирования РЭА. - М.: Радио и связь, 1987. - 255 с.

3. B a n d l e r J.W., S a l a m a A.E. Fault diagnosis of analog circuits // Proceedings of the IEEE. August 1985. - P. 1279-1325.

4. Автоматизация контрольно-регулирующих процессов производства бытовой радиоэлектронной аппаратуры. Разработка макета контрольной системы "Автомат Сакала-87У": Отчет промежуточный / Таллинск. политехн. ин-т; Рук. А. Лаансоо. № ГР 01860115847. - Таллин, 1987. - 90 с.

5. Автоматизация контрольно-регулирующих процессов производства бытовой радиоэлектронной аппаратуры. Создание экспериментального образца установки "Автомат Сакала-87У": Отчет / Таллинск. политехн. ин-т; Рук. А. Лаансоо. № ГР 01860115847. - Таллин, 1987. - 97 с.

6. Лаансоо А., Мяннама В., Пякков О. Измерение параметров аналоговых микросборок применением компенсационного метода // См. наст. сб., с.105-112.

A. Laansoo, V. Männama, O. Pikkov

Problems of the Automatic Information and Measurement
System Design for Microcircuits Performance Test

Abstract

The structure of the automatic information and measurement system for performance test of a very high quality frequency power amplifier's microcircuits has been presented.

The main problems of designing this system have been described and some recommendations of their possible solution have been given.

A. Laansoo, V. Männama,
O. Pikkov

Mikroskeemide parameetrite automatiseeritud
mõõtekompleksi kommutatsiooni-sobitusseadme
väljatöötamisega seoses olevaid küsimusi

Kokkuvõte

Kirjeldatakse kõrge kvaliteetse helisagedusvõimendi mikroskeemide automatiseeritud mõõtekompleksi struktuuri ning koostist.

Selgitatakse selle mõõtekompleksi keskse seadme - kommutatsiooni-sobitusseadme - ülesannet ning ülesehituspõhimõtet.

Seade on suhteliselt universaalne ning kasutatav ka teistsuguse funktsionaalse suunitlusega mikroskeemide kontrolliks.

ПРОЕКТИРОВАНИЕ ЭЛЕКТРОННЫХ УСТРОЙСТВ ДЛЯ МЕДИЦИНЫ

Введение

Общим для радиоэлектронных преобразователей, используемых для медицинских целей, например, в физиотерапии, является создание ими электромагнитных полей, имеющих различную форму, частоту и мощность.

Для грамотного проектирования медицинской аппаратуры необходимо учитывать ряд ограничений, накладываемых биотелом на те или иные характеристики устройств [1].

Использование уравнений электродинамики позволяет как эффективно решать проблемы анализа и синтеза различной аппаратуры, так и находить объяснения первичных физических механизмов воздействия преобразователей на биотело [2, 3].

На кафедре электроники ТПИ в последние годы исследованы некоторые устройства, служащие для физиотерапии, которые можно отнести к моделям II-У уравнений электродинамики [1].

В работе последовательно рассмотрены:

1. Формирователь токов для биологически активных точек - БАТ.

2. Магнитное поле спиральной катушки для БАТ.

3. Анализ устройств для местной дарсонвализации и вопросы их модернизации.

1. Формирователь токов для БАТ

При создании устройств (цифровых), формирующих и обрабатывающих токи (сигналы), воздействующие на биологически активные точки (БАТ), необходимо учитывать законы воздейст-

вия токов на ткани организма. Воздействующие токи могут быть постоянными, импульсными и переменными.

Действие постоянного тока на ткани обусловлено главным образом поляризационными явлениями [1, 4].

Раздражение импульсным током по закону Дюбуа-Реймона вызывается изменением силы тока и зависит от скорости этого изменения. При этом должно учитываться также свойство ткани к аккомодации [4]. Раздражающее действие прямоугольного импульса зависит также от его длительности, обуславливающей наибольшее смещение заряженных частиц за время действия импульса, и описывается уравнением Вейса-Лапика [4]: $i_n = a/\tau + b$, где i_n - пороговый ток, т.е. минимальный ток, вызывающий реакцию возбудимой ткани, τ - длительность импульса, a и b - постоянные коэффициенты, зависящие от ткани. Действие на ткани ритмично-повторяющихся импульсов называется частотным раздражителем. При этом ткань проявляет свойство лабильности, т.е. способности давать оптимальную реакцию только в определенных пределах частоты.

Действие переменного тока не отличается от действия импульсного, однако при этом отсутствуют поляризационные явления, связанные с постоянной составляющей импульсного тока. С повышением частоты уменьшается длительность положительных и отрицательных импульсов и соответственно снижается раздражающее действие.

По Нерсту [4] в пределах частот 100-3000 Гц пороговый ток увеличивается пропорционально корню квадратному из частоты: $i_n = k_1 \sqrt{f}$. В пределах частот 50-300 кГц пороговый ток увеличивается пропорционально первой степени частоты: $i_n = k_2 f$. При частотах порядка сотен килогерц и выше, когда тепловое смещение заряженных частиц соизмеримо со смещением их, обусловленных переменным током, ток вызывает легкое раздражение [5] и далее оказывает только тепловое воздействие [4].

Величина тока, воздействующего на БАТ, находится в пределах $(1 \dots 10^2) \cdot 10^{-5}$ А при напряжении порядка 1-10 В. Форма тока может быть весьма разная. Чаше других используется прямоугольная форма. Частота воздействующего тока на-

ходится в пределах от 0,001 Гц до 10 кГц. Кроме того, возможен режим модуляции высокочастотных токов более низкочастотными, например, модуляция токовых импульсов, следующих с частотой 2-1000 Гц импульсами с частотой 0,001-0,5 Гц. При этом длительность модулирующих импульсов может находиться в пределах от 10 мкс до сотен миллисекунд.

Из вышеизложенного следует, что элементная база формирователя токов для БАТ должна быть работоспособна до частот 1 МГц при максимальной амплитуде 10 В и нагрузке 10 кОм. Таким требованиям в основном удовлетворяют микросхемы на дополняющих МДП-транзисторах. Один вариант реализации формирователя тока для БАТ представлен на рис. 1.

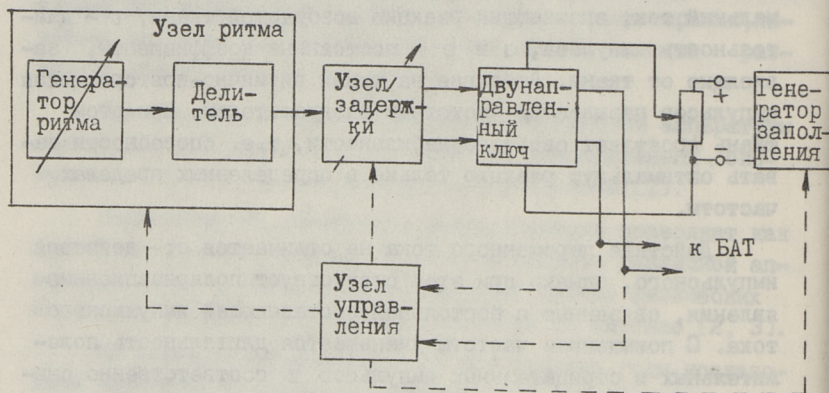


Рис. 1. Формирователь токов для БАТ.

Узел управления может включать в себя устройство выборки и хранения (УВХ) [9], запоминающие устройства, а также АЦП и ЦАП [10].

2. Магнитное поле спиральной катушки и о механизмах воздействия магнитного поля на БАТ

В [3] указано, что для эффективного лечения ряда заболеваний используется узкоконцентрированный поток переменного магнитного поля частотой 50 Гц с расхождением потока в $3-10^0$, напряженностью 100-300 Э, градиентом 50-70 Э/см и длительностью 15-120 секунд.

Несомненный интерес представляют собой вопросы расчета простого устройства, способного обеспечивать заданные выше пространственные характеристики магнитного поля. Такое устройство может быть выполнено в виде катушки [3], условный вид которой показан на рис. 2. На оси Z и при $Z_2 = L$, где радиус катушки наибольший и равен R_1 , требуется обеспечить заданные выше пространственные характеристики, т.е. $H_z = 100-300$ Э,

$|\vec{H}_1| = \left(\frac{3}{90} - \frac{10}{90}\right) H_z$, $\Delta H_z = 50-70$ Э/см. Все величины даны в эффективных значениях, так как в катушке создается переменное по величине и направлению магнитное поле с частотой 50 Гц. Вопросы расчета поля такой катушки автору неизвестны.

Пусть через катушку протекает эффективный по значению ток I в направлении, показанном на рис. 2. Для точки 2 с координатами x_2, y_2, z_2 напряженность магнитного поля $d\vec{H}$ от элемента проводника точки 1 можно рассчитать по закону Био-Савара-Лапласа (закон Ампера)

$$d\vec{H} = \frac{I}{4\pi} \frac{d\vec{l} \times \vec{R}}{|\vec{R}|^3} = i dH_x + j dH_y + k dH_z = d\vec{H}_1 + k dH_z, \quad (I)$$

где $d\vec{l}$ — вектор элементарной длины катушки для точки 1;
 $\vec{R}$ — радиус-вектор, проведенный из точки 1 в точку 2;
 $|\vec{R}|$ — модуль радиус-вектора $\vec{R}$;
 $d\vec{l} \times \vec{R}$ — векторное произведение $d\vec{l}$ на $\vec{R}$;
 i, j, k — единичные векторы, совпадающие с положительными направлениями осей x, y, z .

Остальные обозначения ясны на рис. 2. Катушка, изображенная на этом рисунке, является пространственной кривой, радиус которой, при его линейном увеличении, изменяется в соответствии с выражением (2):

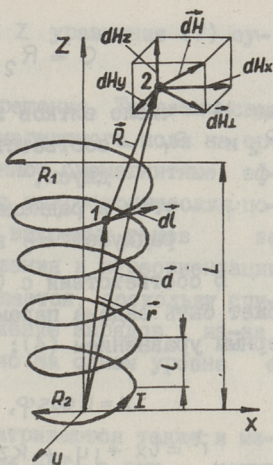


Рис. 2. Спиральная катушка.

$$C = R_2 + \frac{R_1 - R_2}{2\pi N} \varphi = a + b\varphi, \quad (2)$$

где N — число витков катушки;

R_2 и R_1 — соответственно ее наименьший и наибольший радиусы;

φ — угол в радианах, отсчитываемый от начала катушки с радиусом R_2 в плоскости XY .

В соответствии с (2) указанная пространственная кривая может быть задана параметрическими уравнениями (3) или векторным уравнением (4):

$$x = C \cos \varphi, \quad y = C \sin \varphi, \quad z = m\varphi \quad (3)$$

$$\vec{r} = ix + jy + kz = iC \cos \varphi + jC \sin \varphi + km\varphi, \quad (4)$$

где

$$m = \frac{l}{2\pi} = \frac{L}{2\pi N};$$

$\vec{r}$ — это также и радиус-вектор от начала координат до точек на катушке.

Радиус-вектор $\vec{R}$, исходя из геометрических соображений рис. 2, находим в виде:

$$\vec{R} = \vec{d} - \vec{r} = i(x_2 - C \cos \varphi) + j(y_2 - C \sin \varphi) + k(z_2 - m\varphi), \quad (5)$$

где $\vec{d} = ix_2 + jy_2 + kz_2$ и является радиус-вектором точки 2.

Наконец, вектор $d\vec{l}$ уравнения (I) можно математически трактовать как дифференциал радиус-вектора $\vec{r}$:

$$\begin{aligned} d\vec{l} = d\vec{r} &= \frac{d\vec{r}}{d\varphi} d\varphi = \left(i \frac{dx}{d\varphi} + j \frac{dy}{d\varphi} + k \frac{dz}{d\varphi} \right) d\varphi = \\ &= \{ i [-(a+b\varphi)\sin\varphi + b\cos\varphi] + j [(a+b\varphi)\cos\varphi + b\sin\varphi] + \\ &+ km \} d\varphi. \end{aligned} \quad (6)$$

На основании выражений (I), (5) и (6) вектор напряженности $\vec{H}$ катушки при изменении φ от 0 до $2\pi N$ можно найти из уравнения

$$\vec{H} = \frac{I}{4\pi} \int_0^{2\pi N} \frac{d\vec{r} \times \vec{R}}{|\vec{R}|^3}. \quad (7)$$

Следует отметить, что для общего случая уравнение (7) часто не решается в элементарных функциях. При нахождении

напряженности магнитного поля на оси Z уравнение (7) существенно упрощается.

В [3] реализовано именно такое решение. Теоретические объяснения механизмов воздействия магнитного поля на организмы, как правило, сложны и объясняют биоманнитные эффекты либо с микроскопических, либо с макроскопических позиций [6]. При этом эффект наведения вихревых токов за счет индукции Фарадея и эффект разделения и переориентации зарядов за счет сил Лоренца не учитывается, поскольку считается, что при $H = (0,2-300)$ Э движение зарядов из-за названных эффектов меньше или примерно на одном уровне с их хаотическим движением.

В известной литературе не рассматриваются также и механизмы воздействия магнитного поля на БАТ, а они могут представлять значительный интерес. Ниже в теоретическом плане рассмотрены возможные первичные механизмы воздействия магнитного поля на БАТ.

Известно, что БАТ имеют сопротивление R по постоянному току от 40 до 500 кОм, что в несколько раз меньше сопротивления окружающей их ткани, т.е. это физически проводящие "колодцы" от поверхности кожи до ближайших ответвлений [7]. Существует также разность потенциалов U между БАТ и окружающей тканью 2-300 мВ, которая поддерживается изменяющимся по величине и по знаку электрическим током i_r . Ток i_r меняется от долей микроампера до 30 мкА и является отражением регуляторных процессов организма [7]. С радиоэлектронной точки зрения БАТ являются активными датчиками для системы жизнеобеспечения организма.

Для ориентировочной оценки первичного физического воздействия магнитного поля на БАТ примем, что датчик, представляющий собой "колодец активного сопротивления", выглядит как цилиндр с площадью основания $\sim 1 \text{ мм}^2$ и длиной l см (см. рис. 3). В дно цилиндра a входят силовые линии в общем случае переменного неоднородного магнитного поля катушки, а из b выходят. Положим, что в датчике содержатся свободно заряженные частицы (ионы Na^+, Cl^- - внеклеточная среда), а также и движущиеся заряды тока i_r , эмиттируемые чувствительной сетью клеток C (нервными ответвлениями).

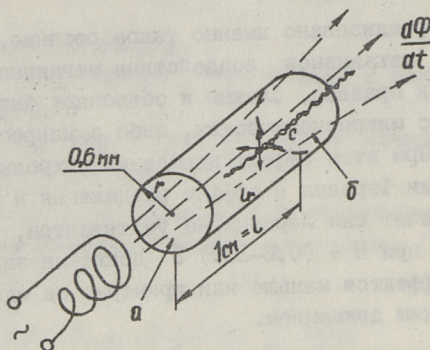


Рис. 3. Условный вид БАТ.

Оценим два возможных механизма влияния переменного магнитного поля на датчик. При этом можно рассматривать независимо свободно заряженные и движущиеся частицы.

I. Переменное магнитное поле вызывает в датчике вихревые токи, плоскость которых перпендикулярна к силовым магнитным линиям и которые возрастают от оси датчика к его периферии. Последнее связано с тем, что наводимая электродвижущая сила $\mathcal{E}$, создающая вихревые токи, пропорциональна площади S' контура, пронизываемого магнитными линиями, а сопротивление контура пропорционально его линейным размерам.

Свободно заряженные частицы, участвующие в создании вихревых токов, сталкиваясь с чувствительной сетью клеток, могут вызвать переменный по величине и по знаку вращающий момент для сети клеток, который также возрастает от центра датчика к периферии и ослабевает по оси с приближением к плоскости "б" (поскольку здесь поток магнитной индукции $\Phi = \mu_0 H S' \cos \omega t$ меньше, чем в плоскости "а"). Таким образом, чувствительная сеть клеток испытывает как бы переменное скручивание и раскручивание как поперек, так и вдоль своей оси с частотой 50 Гц. Однако подобное механическое истолкование влияния поля на БАТ может быть правомерным при $H > 10^6$ А/м (80 А/м ≈ 1 Э). При названных же выше полях движение заряженных частиц вихревого тока су-

щественно меньше их теплового хаотического движения и, следовательно, указанный выше эффект маскируется.

Действительно, оценим, например, какое максимальное смещение ион натрия может приобрести под действием вихревого поля электродвижущей силы $\mathcal{E} = -\frac{d\Phi}{dt} = -\omega \mu_0 H S' \sin \omega t$.

Считая $\omega = 300$ рад/с, $\mu_0 = 4\pi \cdot 10^{-7}$ Гн/м, $H = 10^4$ А/м, $S' = 10^{-6}$ м² получим $\mathcal{E}_{\text{макс}} \approx 3$ мВ. Учитывая, что средние скорости $V = b_{\text{Na}} \cdot E$, где E — внутреннее поле, а $b_{\text{Na}} = 0,45 \times 10^{-7}$ м²/с. b — подвижность ионов натрия, для наружного контура датчика в первом приближении можно оценить скорости движения ионов натрия $V = b_{\text{Na}} \frac{\mathcal{E}}{2\pi r}$ и их смещение $S = \int V dt = \frac{b_{\text{Na}}}{2\pi r} \mu_0 H S' \cos \omega t$.

Имеем $V_{\text{макс}} \sim 0,4 \cdot 10^{-10}$ м/с и $S_{\text{макс}} \sim 0,13 \cdot 10^{-12}$ м, что существенно меньше, чем смещение молекул электролита (жидкости) из-за их теплового движения: $S \sim 3 \cdot 10^{-10}$ м.

2. Рассмотрим теперь наиболее реальный механизм воздействия магнитного поля на движущиеся заряды тока $\vec{j}_r$ биологически активных точек. Пусть ток $\vec{j}_r$ образуется зарядами Na^+ , K^+ и Cl^- . На заряды действует электрическое поле с напряженностью $E \approx \frac{U}{r}$ (внутреннее поле между БАТ и окружающей тканью) и магнитное поле $\vec{H}$ (внешнее, возмущающее поле), причем сила $\vec{f}$, приложенная к зарядам q , равна векторной сумме силы $\vec{f}_e = q\vec{E}$, действующей на заряд со стороны электрического поля, и силы Лоренца $q\vec{V} \times \mu_0 \vec{H}$ магнитного поля, т.е.

$$\vec{f} = \vec{f}_e + \vec{f}_{\text{Лор}} = q\vec{E} + q\vec{V} \times \mu_0 \vec{H}, \quad (8)$$

где $\vec{V} \times \vec{H}$ — векторное произведение скорости ионов $\vec{V}$ и поля $\vec{H}$.

Мы пришли к задаче о движении заряженных частиц, направленных под углом друг к другу в электрическом и магнитном полях. Решение такой задачи для случая однородных и взаимоперпендикулярных и постоянных $\vec{E}$ и $\vec{H}$ дано в [2]. Для случая переменного магнитного поля решение задачи существенно усложняется. Пусть в прямоугольной системе координат вектор $\vec{H}$ направлен вдоль ее Z оси, которая совпадает с осью датчика, вектор $\vec{E} (E_x, E_y, E_z)$ направлен под

произвольным углом к $\vec{H}$. Тогда уравнения движения рассматриваемых ионов можно записать в виде системы дифференциальных уравнений

$$\begin{aligned} m\dot{V}_x &= F_x + \mu_0 q V_y H \cos \omega t \\ m\dot{V}_y &= F_y - \mu_0 q V_x H \cos \omega t; \quad m\dot{V}_z = F_z, \end{aligned} \quad (9)$$

где $\dot{V}_i$, V_i и F_i — проекции ускорения, скорости и силы, действующие на заряды со стороны электрического поля на координату i .

Решая (9), можно точно определить траектории движения ионов. Наша задача существенно упростится, если мы, во-первых, попытаемся сделать только оценку отклонения траектории движения ионов, например, по координате X из-за постороннего поля $\vec{H}$ и, во-вторых, примем, что скорости ионов до введения поля $\vec{H}$ установились и вектор их скорости $\vec{V}_0 = iV_{0x} + jV_{0y} + kV_{0z}$. При этом уравнения (9) существенно упростятся, поскольку $F_x, F_y, F_z = 0$ (скорости постоянны), а $\vec{V}_0 = b\vec{E}$.

Тогда по координате X последовательным интегрированием можно получить проекции ускорения $\dot{V}_x$, скорости V_x и смещения S_x по оси X ($S_0 = 0$):

$$\begin{aligned} \dot{V}_x &= \frac{\mu_0}{m} q b E_y H \cos \omega t \\ V_x &= \frac{\mu_0}{m} q b E_y H \frac{\sin \omega t}{\omega} + V_{0x} \\ S_x &= -\frac{\mu_0}{m} q b \frac{E_y H}{\omega^2} \cos \omega t + V_{0x} t + S_0 = S_H + S_E. \end{aligned}$$

Пусть $t = 2 \cdot 10^{-3}$ с (потенциал действия клеток длится около 2 мс [8]). Тогда для ионов, например, натрия, принимая $E_y \approx E_x \approx 4-600$ В/м и $H = 10^4$ А/м, получим смещения из-за H : $S_H \approx 10^{-5}-10^{-7}$ м, из-за E : $S_{E_{\max}} \approx 5 \cdot 10^{-8}$ м. Видно, что отклонения траекторий ионов, вынуждаемые внешним полем $\vec{H}$ от заданных траекторий, определяемых полем в БАТ $\vec{E}$, могут существенно превышать размеры самих атомов.

Для того, чтобы поверить, что столь малые отклонения в движении ионов могут воздействовать на БАТ, сошлемся на известную работу Б. Катца [8], в которой описываются механизмы возбуждения и передачи возбуждения по нервным ответв-

лениям на клеточном уровне при помощи калий-натриевого насоса. Поскольку физические процессы происходят при этом на атомном уровне (движение ионов натрия и калия через мембрану в клетку и из клетки), то прямые методы наблюдения таких процессов в настоящее время отсутствуют. Однако многочисленные косвенные наблюдения подтверждают эти процессы. В [8], в частности, предполагается, что затруднительное проникновение, например, ионов натрия в клетку в неподходящих условиях обусловлено несоответствием размеров ионов натрия и пор в мембране. Очевидно, если размеры станут соизмеримыми, то из-за возмущения полем $\vec{H}$, отклоняющего траекторию движения ионов на расстояние даже меньшее, чем размеры ионов, ион Na^+ не попадет в отверстие мембраны. С другой стороны, известно, что для поддержания тонуса мышц организма, контроля его отдельных параметров и т.д. из центральной нервной системы даже в состоянии покоя посылаются на периферию запросы в виде серии импульсов от 10 до 100 и более в секунду. По-видимому, и разность потенциалов между БАТ и окружающей тканью и сам ток i_r обусловлены такими же сериями импульсов. Вероятно, существует и механизм обратной связи для оценки такого запроса, например, через сопряженные точки [7]. Каждый импульс сопровождается потенциалом действия, при котором происходит обмен ионов Na^+ и K^+ внутрь и наружу нервных клеток [8]. Учитывая, что потенциал действия $\Delta u \sim 0,1$ В, а емкость мембраны клеток $C \sim 1$ мкФ/см² заряд, переносимый за волну возбуждения ($\Delta t = 2$ мс), плотность тока j и эффективное значение плотности тока $j_{\text{эфф}}$ будут: $j = \Delta q / \Delta t = 0,05$ мкА/см², $j_{\text{эфф макс}} = \Delta q \cdot f = 100$ мкА/см². Рассчитанные макроскопические величины на основании вышесказанного должны меняться из-за поля $\vec{H}$ и за 15–120 секунд оказывать влияние на организм [6].

В заключение отметим, что рассмотренный выше механизм позволяет с единых позиций объяснить воздействие на БАТ и других сил, в том числе оценить максимально возможную энергию электромагнитных полей, воздействие которых на БАТ еще будет иметь эффект [6].

3. Устройство для местной дарсонвализации

Устройства для местной дарсонвализации находят широкое применение в физиотерапии [5].

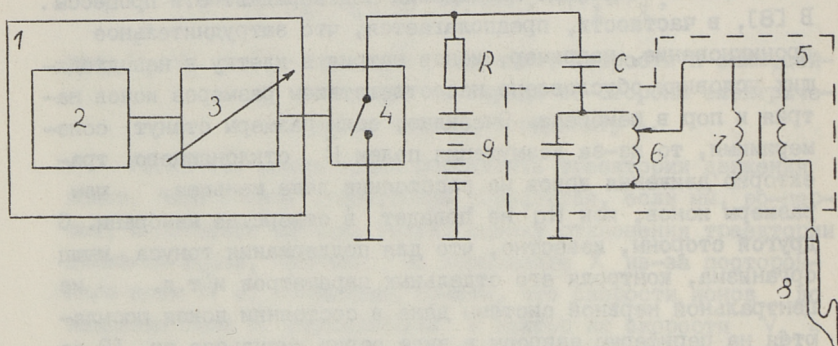


Рис. 4. Синтезированная схема дарсонвализатора.

На рис. 4 представлена структурная схема, обобщающая [12, 13, 14] и содержащая формирователь длительности импульсов I, электронный ключ 4, контур ударного возбуждения 5.

Индуктивный делитель 6 вместе с автотрансформатором 7 и лечебным электродом 8 представляет собой сложную многоконтурную систему [15], настраиваемую на частоту колебаний устройства для местной дарсонвализации, причем все, что справа от индуктивного делителя 6 – это два связанных между собой магнитным полем контура, и коэффициент полезного действия η всей системы зависит только от двух параметров – фактора связи A и обобщенной расстройки второго контура ξ_2 [15].

Введение индуктивного делителя позволяет существенно увеличить η как за счет исключения активных потерь, так и за счет резонансных настроек сложной контурной системы.

При этом следует учитывать, что индуктивный делитель связан с автотрансформатором и электродом лишь электрически, через кабель.

Выигрыш η блока 5 предложенного варианта с индуктивным делителем по сравнению с делителем на активном сопро-

тивлении [5] будет от 1,5 до 10 раз (в системе правее ключа).

Конденсатор С и сопротивление R позволяют не только свести к минимуму причины, приводящие к перегрузке электронного ключа от случайных факторов, но и упростить требования, предъявляемые к формирователю импульсов.

Физически соединения С и R на схеме строго и надежно ограничивают располагаемую мощность источника питания $\mathcal{G}$, и в любой момент времени ограничивают мгновенную мощность на ключе до безопасного уровня.

Известно, что на практике используется вакуумированный стеклянный электрод. Однако возможен и электрод, выполненный из диэлектрика, имеющий металлический колпак и наполненный проводящей жидкостью, например, водой.

Такое решение позволяет при производстве электродов для дарсонвализации отказаться от вакуумной и высокотемпературной технологии, т.е. резко упростить технологию их производства.

Работоспособность предложенного электрода подтверждена практическими испытаниями. Расчеты подтверждают, что жидкости с удельным сопротивлением 0,5–1000 Ом·м вполне заменяют графит или разреженный газ в электродах.

Отметим также, что сопротивление внутренней полости электрода длиной 15 см и диаметром 1 см, наполненной разреженным газом, графитом или жидкостью, будет 10^3 – $2 \cdot 10^6$ Ом.

Внедрение данного жидкостно-наполненного электрода при сохранении стабильности его характеристик позволит упростить технологию их производства. Кроме того, использование других диэлектрических (нехрупких) материалов, таких, как полиэтилен, полиуретан и в особенности резина или поливинилхлорид (катетер), позволяет как исключить травматизм при лечении пациентов, так и использовать дарсонвализацию для лечения ранее недоступных другими электродами мест (пищевод, желудок и т.д.).

На выше рассмотренных примерах были показаны проблемы, возникающие при разработке аппаратуры для медицинских целей.

Приведены некоторые приемы их разрешения, которые требуют знания элементной базы, математических и изобретательских приемов.

Л и т е р а т у р а

1. Г у р ь я н о в Б.Г. Модели уравнений электродинамики для учета реакций биотел // Тр. Таллинск. политехн. ин-та. - 1982. - № 540. - С. 91-98.

2. К и т т е л ь Я. и др. Механика. - М.: Наука, 1983. - С. 133-137.

3. Г у р ь я н о в Б.Г. Низкочастотное магнитное поле для БАТ. Материалы доклада и тезисы докладов. XXXIV Всесоюзная научная сессия, посвященная Дню радио. - Ч. I. - М.: Радио и связь. - 1984. - С. 77-78; доклад 9 с.

4. Л и в е н ц е в Н.М. Курс физики. - М., 1978. - С. 146-170.

5. Г у р ь я н о в Б.Г. Расчет дарсонвализатора // Тр. Таллинск. политехн. ин-та. - 1982. - № 538. - С. 73-81.

6. П р е с м а н А.С. Электромагнитные поля и живая природа. - М.: Наука, 1986. - С. 48-79.

7. Т а б е е в а Д.М. Руководство по иглорефлексотерапии. - М.: Медицина, 1982. - С. 77-81.

8. К а т ц В. Нерв, мышца и синапс. - М.: Мир, 1968. - С. 54-114.

9. Г у р ь я н о в Б.Г. и др. Устройство выборки и хранения для аналого-цифрового преобразователя видеосигналов // Проблемы цифрового кодирования и преобразования изображений: Тез. Всесоюз. симпозиума. - Тбилиси, 1980. - С. 27-28.

10. Г у р ь я н о в Б.Г. Телевизионные аналого-цифровые преобразователи. - Таллин // Side. Raadio. Televisioon. - 1980. - № 2. - С. 10-12.

11. Г у р ь я н о в Б.Г., Б е л к и н а И.И. Энергетика устройств для местной дарсонвализации // XXXIV Всесоюз. науч. сессия, посвященная Дню радио: Материалы доклада и тезисы докладов. - Ч. I. - М.: Радио и связь, 1984. - С. 82.

12. А.с. II53I29 СССР. Устройство для местной дарсонвализации / Б.Г. Гурьянов. Оpubл. 1985, Б.И., № 17.

13. А.с. II026I2 СССР. Устройство для местной дарсонвализации / Б.Г. Гурьянов. - Оpubл. 1984, Б.И., № 26.

14. А.с. II43427 СССР. Устройство для местной дарсонвализации / Б.Г. Гурьянов. - Оpubл. 1985, Б.И., № 9.

15. Г у р ь я н о в Б.Г. Машинный расчет автотрансформатора для дарсонвализации // Тр. Таллинск. политехн. ин-та.-1986.- № 6I7. - С. 57-6I.

B. Guryanov

Electronics in Medicine

Abstract

This paper studies some problems concerning the design of medical equipment.

An apparatus for acupuncture points is presented. The axial magnetic field strength of the coiled coil device is calculated.

The possible physical aspect of the effect of magnetic field on the acupuncture points is considered.

Some ways for improving darsonvalizators are also given.

B. Gurjanov

Meditsiinielektroonika

Kokkuvõte

Käsitletakse meditsiinielektroonikas kasutatavaid raadioelektroonseid muundureid ning nende loomisel esinevaid üldisi probleeme.

Näidatakse, et elektrodünaamika võrrandid lubavad võtta arvesse biovooludest tingitud piiranguid ja selgitada muundurite mõju bioloogilisele organismile.

Üksikasjalikult analüüsitakse kohaliku darsonvalisatsiooni seadet, samuti ka selle täiustamisega seoses olevaid küsimusi.

ВРЕМЕНА ЖИЗНИ НЕОСНОВНЫХ НОСИТЕЛЕЙ ЗАРЯДА В
ПОЛУПРОВОДНИКАХ С УЧЕТОМ ОЖЕ-ПРОЦЕССОВ
НА ЛОКАЛИЗОВАННЫХ СОСТОЯНИЯХ

Введение

В настоящее время при численном моделировании полупроводниковых приборов рассматриваются, как правило, два механизма рекомбинации носителей, обуславливающие затухание их плотности:

- 1) механизм Шокли-Рида-Холла (ШРХ) — рекомбинации через локализованные состояния;
- 2) межзонная Оже-рекомбинация, не зависящая от энергетической структуры запрещенной зоны.

С другой стороны, при анализе различных явлений в ППП, в том числе и времен жизни неосновных носителей заряда (ННЗ), привлекаются представления об Оже-процессах, идущих с участием локализованных состояний [1]. Изучению этих процессов посвящено значительное количество работ (например, обзоры [1, 2] и литература, указанная там). Однако имеются довольно большие неточности в определении соответствующих коэффициентов. Интересно проследить влияние различных Оже-процессов на локализованных состояниях на время жизни ННЗ, чему и посвящена данная статья.

I. Учет процессов низшего порядка

На рис. I изображена зонная диаграмма полупроводника с двумя энергетическими уровнями E_1 и E_2 в запрещенной зоне и показаны различные Оже-процессы, идущие через эти уровни. На диаграмме указаны процессы четырех различных ти-

пов. Процессы Ia, б есть межзонные Оже-процессы, в которых состояния в запрещенной зоне не участвуют. Имея в виду количество локализованных носителей, участвующих в акте Оже-рекомбинации, эти процессы естественно определить как процессы нулевого порядка. Тогда 1в-3 процессы первого порядка (аналогичные процессы для уровня E_2 на рис. I не показаны).

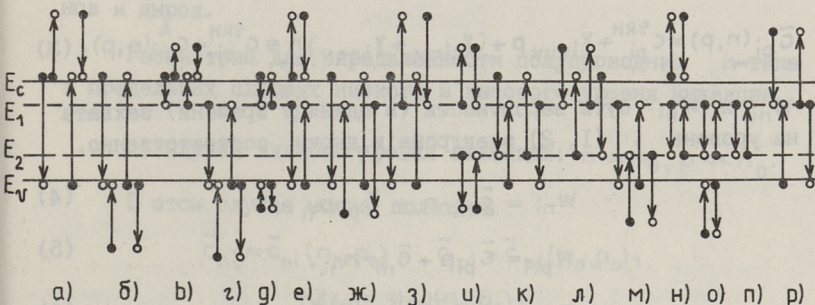


Рис. 1. Иллюстрация Оже-процессов различных порядков. Процессы рекомбинации и обратные к ним процессы ударной ионизации: (а, б) - прямые межзонные процессы; (в-з) - Оже-процессы с участием уровня E_1 ; (и-о) - Оже-процессы с участием двух уровней E_1 и E_2 ; п - "ШРХ-взаимодействие" уровней E_1 и E_2 ; Р - нелинейный Оже-процесс на уровне E_1 . (●) - электроны; (○) - дырки.

Остальные процессы, 1и-р суть процессы второго порядка, причем 1п соответствует "ШРХ-взаимодействию" уровней E_1 и E_2 ; 1р есть процесс второго порядка при наличии даже и одного уровня в запрещенной зоне. Процессы 1и-о соответствуют "Оже-взаимодействию" уровней E_1 и E_2 . Для облегчения идентификации всех процессов условимся обозначать соответствующие им коэффициенты четверным индексом. Например, коэффициенты для прямого и обратного процессов 1К суть γ_{cv21} и γ_{21cv} . Первые два индекса соответствуют начальным состояниям электронов, вторые - конечным. Для коэффициентов γ_{ccvc} , γ_{vccc} и γ_{cvvv} , γ_{vvcv} используем часто употребляемые обозначения γ_n, a_n и γ_p, a_p соответственно.

С учетом Оже-процессов нулевого и первого порядков, а также ШРХ-рекомбинации через локальные уровни кинетические уравнения для плотностей n_1 и n_2 носителей, локализованных на уровнях E_1 и E_2 , имеют вид:

$$\dot{n}_i = \tilde{c}_{ni}(n, p)(np_i - n_{ii}n_i) - \tilde{c}_{pi}(n, p)(pn_i - p_{ii}p_i) \equiv W_{ni}p_i - W_{pi}n_i, \quad (I)$$

где $p_i = N_i - n_i$;

$$\tilde{c}_{ni}(n, p) = c_{ni}^{SRH} + \gamma_{ccic}n + (\gamma_{cvi v} + \gamma_{cvvi})p \equiv c_{ni}^{SRH} + c_{ni}^A(n, p); \quad (2)$$

$$\tilde{c}_{pi}(n, p) = c_{pi}^{SRH} + \gamma_{ivvv}p + (\gamma_{cicv} + \gamma_{iccv})n \equiv c_{pi}^{SRH} + c_{pi}^A(n, p). \quad (3)$$

W_{ni} и W_{pi} суть вероятности (в единицу времени) захвата на уровень i (1, 2) электрона и дырки соответственно.

$$W_{ni} = \tilde{c}_{ni} \cdot n + \tilde{c}_{pi} \cdot p_{ii}. \quad (4)$$

$$W_{pi} = \tilde{c}_{pi}p + \tilde{c}_{ni} \cdot n_{ii}, \quad (5)$$

$$n_{ii} = N_c g_i^{-1} e^{(E_i - E_c)/kT}, \quad (6)$$

$$p_{ii} = N_v g_i e^{(E_v - E_i)/kT}. \quad (7)$$

Здесь N_c и N_v эффективные плотности состояний для носителей в c - и v -зонах соответственно; g_i - кратность вырождения уровня i .

Тогда кинетические уравнения генерации-рекомбинации для свободных носителей имеют вид (с учетом принципа детального равновесия):

$$\dot{n} = - \sum_{i=1}^2 \tilde{c}_{ni}(np_i - n_{ii}n_i) - \gamma_{vv}(n, p)(np - n_0p_0), \quad (8)$$

$$\dot{p} = - \sum_{i=1}^2 \tilde{c}_{pi}(pn_i - p_{ii}p_i) - \gamma_{vv}(n, p)(np - n_0p_0), \quad (9)$$

где

$$\gamma_{vv}(n, p) = \gamma_n n + \gamma_p p. \quad (10)$$

В режиме стационарной рекомбинации, т.е. при $\dot{n}_i = 0$ из уравнений (I), (8) и (9) получим:

$$n_i = \frac{N_i W_{ni}}{W_{ni} + W_{pi}} \equiv \frac{N_i W_{ni}}{W_i}, \quad (II)$$

$$\dot{n} = \dot{p} = -R_n = -R_p = -\left[\sum_i \frac{N_i \tilde{c}_{ni} \tilde{c}_{pi}}{W_i} + \gamma_{BB}\right](np - n_0 p_0), \quad (I2)$$

где n_0 и p_0 суть плотности равновесных свободных электронов и дырок;

R_n и R_p — скорости результирующей рекомбинации электронов и дырок.

Рассмотрим для определенности полупроводник n -типа в предельных случаях низкого и высокого уровня инжекции.

А. Случай низкого уровня инжекции: $n \approx n_0, p \ll n_0$.

В этом случае можно положить

$$\begin{aligned} \tilde{c}_{ni} &\approx \tilde{c}_{ni}(n_0, p_0), \quad \tilde{c}_{pi} \approx \tilde{c}_{pi}(n_0, p_0), \\ \gamma_{BB} &\approx \gamma_{BB}(n_0, p_0). \end{aligned} \quad (I3)$$

$$np - n_0 p_0 \approx n_0 \Delta p, \quad \Delta p = p - p_0 \approx p.$$

В уравнении (I3) учтено, что $p_0 \ll n_0$ и $\Delta n = \Delta p$ (квазинейтральная область). Тогда из (I2) получим

$$\tau_p \equiv \Delta p / R_p = \frac{1}{n_0 F_{\text{нуй}}}, \quad (I4)$$

где

$$F_{\text{нуй}} \equiv \left\{ \sum_i \frac{N_i \tilde{c}_{ni} \tilde{c}_{pi}}{W_i} + \gamma_{BB} \right\}_{\text{нуй}} \approx \sum_i \frac{N_i c_{ni}^{\text{SRH}} c_{pi}^{\text{SRH}}}{(c_{ni}^{\text{SRH}} n_{i1} + c_{pi}^{\text{SRH}} p_{i1})} \cdot \frac{\varphi_{1i}^{\text{нуй}}}{\varphi_{2i}^{\text{нуй}}} + \gamma_n \cdot n_0, \quad (I5)$$

$$\varphi_{1i}^{\text{нуй}} \equiv 1 + \left[\frac{(\gamma_{cicv} + \gamma_{iccv})}{c_{pi}^{\text{SRH}}} + \frac{\gamma_{ccic}}{c_{ni}^{\text{SRH}}} \right] \cdot n_0 + \frac{\gamma_{ccic}(\gamma_{cicv} + \gamma_{iccv})}{c_{ni}^{\text{SRH}} c_{pi}^{\text{SRH}}} \cdot n_0^2, \quad (I6)$$

$$\varphi_{2i}^{\text{нуй}} \equiv 1 + \left[\frac{c_{ni}^{\text{SRH}} + \gamma_{ccic} n_{i1} + (\gamma_{cicv} + \gamma_{iccv}) p_{i1}}{c_{ni}^{\text{SRH}} n_{i1} + c_{pi}^{\text{SRH}} p_{i1}} \right] \cdot n_0 + \frac{\gamma_{ccic}}{c_{ni}^{\text{SRH}} n_{i1} + c_{pi}^{\text{SRH}} p_{i1}} \cdot n_0^2. \quad (I7)$$

В. Случай высокого уровня инжекции: $p \approx n \gg n_0$.

$$\tau_p = \frac{1}{n F_{\text{вуй}}}, \quad (I8)$$

$$F_{\text{вуй}} \equiv \left\{ \sum_i \frac{N_i \tilde{c}_{ni} \tilde{c}_{pi}}{W_i} + \gamma_{BB} \right\}_{\text{вуй}} \approx \sum_i \frac{N_i c_{ni}^{\text{SRH}} c_{pi}^{\text{SRH}}}{(c_{ni}^{\text{SRH}} n_{i1} + c_{pi}^{\text{SRH}} p_{i1})} \cdot \frac{\varphi_{1i}^{\text{вуй}}}{\varphi_{2i}^{\text{вуй}}} + (\gamma_n + \gamma_p) \cdot n, \quad (I9)$$

$$\varphi_{1i}^{вун} = 1 + \left[\frac{(\gamma_{ivvv} + \gamma_{cicv} + \gamma_{iccv})}{C_{pi}^{SRH}} + \frac{(\gamma_{ccci} + \gamma_{cviv} + \gamma_{cvvi})}{C_{ni}^{SRH}} \right] \cdot n +$$

$$+ \frac{(\gamma_{ivvv} + \gamma_{cicv} + \gamma_{iccv})(\gamma_{ccci} + \gamma_{cviv} + \gamma_{cvvi})}{C_{ni}^{SRH} \cdot C_{pi}^{SRH}} \cdot n^2, \quad (20)$$

$$\varphi_{2i}^{вун} = 1 + \left[\frac{(\gamma_{ccci} + \gamma_{cviv} + \gamma_{cvvi}) \cdot n_{1i} + (\gamma_{ivvv} + \gamma_{cicv} + \gamma_{iccv}) \cdot p_{1i} + (C_{ni}^{SRH} + C_{pi}^{SRH})}{C_{ni}^{SRH} \cdot n_{1i} + C_{pi}^{SRH} \cdot p_{1i}} \right] \cdot n +$$

$$+ \frac{(\gamma_{ccci} + \gamma_{cviv} + \gamma_{cvvi} + \gamma_{ivvv} + \gamma_{cicv} + \gamma_{iccv})}{C_{ni}^{SRH} \cdot n_{1i} + C_{pi}^{SRH} \cdot p_{1i}} \cdot n^2. \quad (21)$$

Как видно из сравнения (I6), (I7) и (20), (21) величины φ_{1i} и φ_{2i} при высоком и низком уровнях инжекции отличаются лишь коэффициентами при степенях плотности основных носителей.

С. Количественные оценки

Сечение каскадного захвата свободных носителей на притягивающие центры (мелкие и глубокие) согласно [3] имеет вид:

$$\sigma = \frac{4 \cdot \pi}{3 \epsilon_0} \cdot \frac{e^2 Z}{\kappa kT} \left[\frac{e^2 Z}{\kappa (kT + 2,74 m s^2)} \right]^2, \quad (22)$$

где m — эффективная масса носителя заряда;

s — скорость звука в кристалле;

κ — диэлектрическая проницаемость кристалла;

eZ — заряд центра;

$\epsilon_0 = v_T \cdot \tau_\epsilon$;

v_T — средняя тепловая скорость носителей заряда;

τ_ϵ — время их энергетической релаксации.

Применительно, например, к глубоким центрам золота в кремнии при комнатной температуре эта формула дает $\sigma_{Au}^+ \approx 6 \cdot 10^{16} \text{ см}^2$ и $\sigma_{Au}^- \approx 1,6 \cdot 10^{15} \text{ см}^2$. Экспериментальное значение σ_{Au}^+ почти на порядок выше, т.е. $\sigma_{Au}^+ = 6 \cdot 10^{15} \text{ см}^2$ [3]. Для мелких кулоновских центров область применимости (22) ограничена температурой $T \lesssim 100 \text{ К}$ [3].

Теория многофононного захвата дает при комнатной температуре сечения захвата (по порядку величины) [4, с. 212]:

$\sigma_{att} \sim 10^{-13} \text{ см}^2$ — для захвата носителя на центр с зарядом противоположного знака;

$\sigma_{neut} \sim 10^{-15} \text{ см}^2$ — для захвата на нейтральный центр;

$\sigma_{rep} \sim 10^{-17} \text{ см}^2$ — для захвата на одноименно заряженный центр.

Таким образом, для оценок можно принять

$$\sigma_{att} \sim 10^{-14} \text{ см}^2, \quad \sigma_{neut} \sim 10^{-15} \text{ см}^2 \quad \text{и} \quad \sigma_{rep} \sim 10^{-17} \text{ см}^2.$$

Для коэффициентов захвата C_n^{SRH} и C_p^{SRH} это дает:

$$C_{n,p}^{SRH} \equiv \sigma \cdot v_{n,p} \sim 10^{-7} \text{ см}^3/\text{с}, \quad 10^{-8} \text{ см}^3/\text{с} \quad \text{и} \quad 10^{-10} \text{ см}^3/\text{с}$$

соответственно.

Другими словами, для донора (глубокого или мелкого)

$$C_{nd}^{SRH} \sim 10^{-7} \text{ см}^3/\text{с}, \quad C_{pd}^{SRH} \sim 10^{-8} \text{ см}^3/\text{с}; \quad (23)$$

для акцептора (глубокого или мелкого)

$$C_{na}^{SRH} \sim 10^{-8} \text{ см}^3/\text{с}, \quad C_{pa}^{SRH} \sim 10^{-7} \text{ см}^3/\text{с}. \quad (24)$$

Зависимости коэффициентов γ_{ccic} , γ_{civv} , γ_{cicv} , γ_{ivvv} от энергии E_i приведены в [1]. Результаты различных теоретических вычислений γ_{ccic} приведены также в [5]. Разброс в оценках составляет примерно порядок величины. В работе [6] показано, что учет электрон-фононного взаимодействия в Оже-процессах на локализованных состояниях приводит к увеличению коэффициента γ_{ccic} примерно в четыре раза (то же, по-видимому, будет иметь место и для других коэффициентов для процессов первого порядка). Наконец, в [7] указаны дополнительные причины увеличения Оже-коэффициентов (количественные оценки, однако, не сделаны).

Для приближенных оценок можно принять энергетические зависимости коэффициентов γ_{ccic} и γ_{cicv} , приведенные в [1, 5].

Рассмотрим, например, пару энергетических уровней в кремнии, расположенных симметрично относительно середины запрещенной зоны $E_{in} = \frac{E_c + E_v}{2}$.

Пусть $E_1 = E_c - 0,2$ эВ, $E_2 = E_v + 0,2$ эВ. Значения коэффициентов γ и параметров n_{1i}, p_{1i} приведены в таблице I.

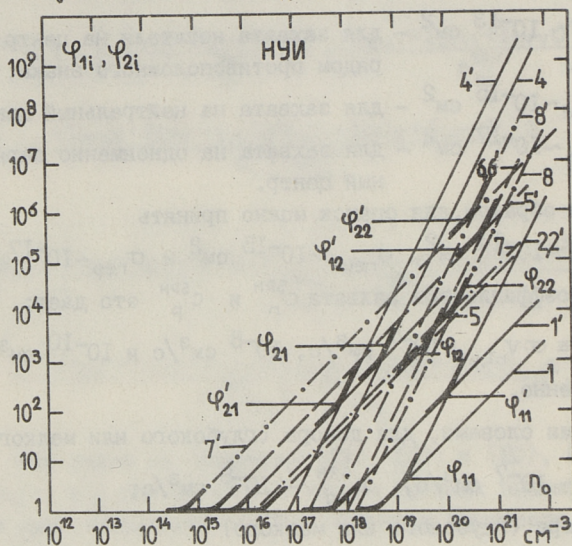


Рис. 2. Функции (16а) и (17а) при НУИ:

1 - $a_{11} \cdot n_0$; 2 - $b_{11} \cdot n_0^2$; 3 - $a_{21} \cdot n_0$; 4 - $b_{21} \cdot n_0^2$; 5 - $a_{12} \cdot n_0$; 6 - $b_{12} \cdot n_0^2$; 7 - $a_{22} \cdot n_0$; 8 - $b_{22} \cdot n_0^2$. (1-8) - для случая, когда E_1 - донор, E_2 - акцептор. (1' - 8') - аналогичные линии для случая, когда E_1 - акцептор, E_2 - донор. Жирные линии отмечают переходы к главным по величине членам в $\varphi_{1i}, \varphi_{2i}$.

Рассчитанные по формулам (I6) и (I7) с данными таблицы I функции $\varphi_{1i}^{НУИ}$ и $\varphi_{2i}^{НУИ}$ приведены на рис. 2. Они представлены в виде совокупности двух пересекающихся семейств прямых с линейным и квадратичным по n_0 наклоном, соответственно общему виду функций φ :

$$\varphi_{1i} = 1 + a_{1i} n_0 + b_{1i} n_0^2, \quad (I6a)$$

$$\varphi_{2i} = 1 + a_{2i} n_0 + b_{2i} n_0^2. \quad (I7a)$$

Рис. 2 отчетливо демонстрирует различные области, обуславливающие различную степенную зависимость τ_p . Функции $\tau_p(n_0)$, полученные с учетом рис. 2, приведены на рис. 3. Для использованных параметров и типов центров получено, что

Таблица I

Тип центра	$n_{1i}, \text{см}^{-3}$	$p_{1i}, \text{см}^{-3}$	$\text{SRH}_{\text{ни}}, \text{см}^3/\text{с}$	$\text{SRH}_{\text{дп}}, \text{см}^3/\text{с}$	$\chi_{\text{сссд}}, \text{см}^6/\text{с}$	$\chi_{\text{свев}}, \text{см}^6/\text{с}$	$\chi_{\text{сдсв}}, \text{см}^6/\text{с}$	$\chi_{\text{свдв}}, \text{см}^6/\text{с}$	$\chi_{\text{сссв}}, \text{см}^6/\text{с}$
E_1 (донор)	$4,8 \cdot 10^{15}$	$6,5 \cdot 10^4$	10^{-7}	10^{-8}	$2 \cdot 10^{-26}$	$4 \cdot 10^{-27}$	$9 \cdot 10^{-28}$	$1,4 \cdot 10^{-28}$	$\leq \chi_{\text{сдсв}}^{1)}$
E_2 (акцептор)	$5,5 \cdot 10^4$	$5,8 \cdot 10^{15}$	10^{-8}	10^{-7}	$3 \cdot 10^{-28}$	$6 \cdot 10^{-29}$	$5 \cdot 10^{-24}$	$4 \cdot 10^{-24}$	$\leq \chi_{\text{свдв}}^{1)}$
E_1 (акцептор)	$4,8 \cdot 10^{15}$	$6,5 \cdot 10^4$	10^{-8}	10^{-7}	$2)$	$2)$	$2)$	$2)$	$\leq \chi_{\text{сдсв}}^{1)}$
E_2 (донор)	$5,5 \cdot 10^4$	$5,8 \cdot 10^{15}$	10^{-7}	10^{-8}	$-$	$-$	$-$	$-$	$-$

1) Значения отсутствуют.

2) Значения получены экстраполяцией кривых рис. 7 из [1].

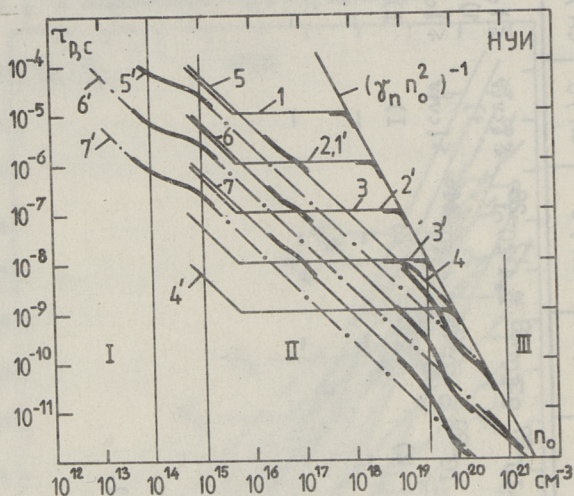


Рис. 3. Время жизни дырок $\tau_{p,c}$ в n-Si при НУИ. Учет одного уровня.

E_1 - донор: 1 - $N_1 = 10^{13} \text{ см}^{-3}$; 2 - $N_1 = 10^{14} \text{ см}^{-3}$; 3 - $N_1 = 10^{15} \text{ см}^{-3}$; 4 - $N_1 = 10^{16} \text{ см}^{-3}$.

E_1 - акцептор: 1' - 4' аналогичны 1-4.

E_2 - донор: 5 - 7 аналогичны 1-4. $N_2 = 10^{18} \text{ см}^{-3}$; 6 - $N_2 = 10^{14} \text{ см}^{-3}$; 7 - $N_2 = 10^{15} \text{ см}^{-3}$.

E_2 - акцептор: 5' - 7' аналогичны 5-7. $\gamma_n = 10^{-31} \text{ см}^6/\text{с}$.

Примечание: жирными линиями выделены переходные участки.

Оже-рекомбинация через уровень E_1 , лежащий вблизи с-зоны, (мелкий донор или глубокий акцептор) не влияет на τ_p (исключение составляет случай концентрации мелких доноров $N_{1d} = 10^{16} \text{ см}^{-3}$). Это, очевидно, является следствием малости соответствующих Оже коэффициентов. В то же время для глубокого донора или мелкого акцептора (уровень E_2), для которого некоторые Оже-коэффициенты достаточно велики, Оже-рекомбинация через этот уровень играет решающую роль. В частности, полностью отсутствует плато, характерное для ШРХ-рекомбинации. На рис. 3 для кривых 5' - 7' отмечены три области концентраций ОНЗ с различным вкладом Оже-процессов в τ_p :

$$\text{Область I: } \tau_p^I = \frac{c_{n2}^{SRH} n_{12} + c_{p2}^{SRH} p_{12}}{n_0 N_2 c_{n2}^{SRH} c_{p2}^{SRH}}, \quad \text{ШРХ-процессы};$$

$$\text{Область II: } \tau_p^{II} = \tau_p^I \frac{a_{22}}{a_{12}} = \frac{c_{n2}^{SRH} + \gamma_{cc2c} n_{12} + (\gamma_{c2cv} + \gamma_{2ccv}) p_{12}}{n_0 N_2 [c_{n2}^{SRH} (\gamma_{c2cv} + \gamma_{2ccv}) + c_{p2}^{SRH} \gamma_{cc2c}]},$$

ШРХ- и Оже-процессы;

$$\text{Область III: } \tau_p^{III} = \tau_p^I \frac{b_{22}}{b_{12}} = \frac{1}{n_0 N_2 (\gamma_{c2cv} + \gamma_{2ccv})}, \text{Оже-процессы.}$$

Аналогичные области можно указать и для кривых 5 - 7. Как видно из рис. 3, Оже-процессы через локализованные состояния при определенных условиях могут сильно уменьшать время жизни носителей заряда уже в области промежуточного легирования. Аналогичные результаты, очевидно, будут справедливы и для случая ВУИ при любом уровне легирования.

При наличии конечной плотности состояний в запрещенной зоне время жизни, очевидно, может быть записано как

$$\tau_p^{-1} = \int_{E_{v\mu}}^{E_{c\mu}} \tau_p^{-1}(E) dE + (\gamma_{BB} n^2)^{-1}, \quad (25)$$

где

$$\tau_p^{-1}(E) = \frac{n N(E) c_{nE}^{SRH} c_{pE}^{SRH}}{c_{nE}^{SRH} n_{1E} + c_{pE}^{SRH} p_{1E}} \cdot \frac{\varphi_{1E}}{\varphi_{2E}}. \quad (26)$$

Здесь $N(E)$ есть плотность состояний в запрещенной зоне при энергии E ; $E_{c\mu}$ и $E_{v\mu}$ пороги подвижности для электронов и дырок соответственно. Смысл остальных обозначений очевиден. В общем случае интеграл (25), конечно, формально не сводится к некоторой результирующей двухуровневой системе, а дает более сложную функциональную зависимость τ_p от n (или n_0).

2. Учет процессов второго порядка

При учете процессов II-р кинетические уравнения для плотностей локализованных и делокализованных носителей записываются как

$$\begin{aligned} \dot{n}_i = & W_{ni} p_i - W_{pi} n_i + 2\gamma_{cvi}(n_i^2 - \alpha_i^2 p_i^2) + \gamma_{ij}(n_i p_j - \alpha_{ij}^2 p_i n_j) + \\ & + c_{ij} p_i n_j - c_{ji} n_i p_j, \quad i, j = 1, 2, \end{aligned} \quad (27)$$

где

$$\gamma_{ij} = \gamma_{ji} \equiv \gamma_{cvij} + \gamma_{cvji}; \quad (28)$$

$$\alpha_{ij}^2 \equiv \frac{n_{ii}}{p_{ij}} = \frac{n_{ij}}{p_{ii}} = \frac{1}{g_i g_j} \frac{N_c}{N_v} \exp\{2(E_{ij} - E_{iv})/kT\}, \quad E_{ij} \equiv \frac{E_i + E_j}{2}, \quad (29)$$

$$\alpha_i^2 \equiv n_{ii}/p_{ii} = \alpha_{ij}^2 |_{i=j}; \quad (29a)$$

$$c_{ij}(n, p) \equiv c_{ij}^{SRH} + c_{ij}^A(n, p); \quad (30)$$

$$c_{ij}^A(n, p) \equiv (\gamma_{cjit} + \gamma_{cjci})n + (\gamma_{vjiv} + \gamma_{vjvi})p; \quad (31)$$

$$c_{ij}/c_{ji} \equiv \alpha'_{ij} = \frac{g_j}{g_i} \exp\{(E_i - E_j)/kT\}; \quad (32)$$

$$\dot{n} = -\sum_{i=1}^2 \tilde{c}_{ni}(np_i - n_{ii}p_i) - \sum_{i=1}^2 \gamma_{cvii}(np_i^2 - \alpha_i^2 p n_i^2) - \gamma_{ij}(np_i p_j - \alpha_{ij}^2 p n_i n_j); \quad (33)$$

$$\dot{p} = -\sum_{i=1}^2 \tilde{c}_{pi}(pn_i - p_{ii}p_i) - \sum_{i=1}^2 \gamma_{cvii}(np_i^2 - \alpha_i^2 p n_i^2) + \gamma_{ij}(np_i p_j - \alpha_{ij}^2 p n_i n_j). \quad (34)$$

В режиме стационарной рекомбинации $\dot{n}_i = 0$ наличие взаимодействия уровней (т.е. переходов между ними) сводится лишь к изменению их стационарных функций распределения $f_i \equiv n_i/N_i$, которые определяются совместным решением системы двух алгебраических уравнений (27). Полагая $n_i = n_i^{(0)} + \Delta n_i$, где $n_i^{(0)}$ определяется формулой (II), из (33) получим

$$R_n = R_n^{(0)} + \Delta R_n, \quad (35)$$

где $R_n^{(0)}$ дается формулой (I2) (с учетом межзонной Оже-рекомбинации);

$$\begin{aligned} \Delta R_n = & -\sum_i \tilde{c}_{ni}(n + n_{ii})N_i \delta f_i + \sum_i \gamma_{cvii}N_i^2 [n(1-f_i)^2 - \alpha_i^2 p f_i^2] + \\ & + \gamma_{ij}N_i N_j [n(1-f_i)(1-f_j) - \alpha_{ij}^2 p f_i f_j], \end{aligned} \quad (36)$$

$$\delta f_i = f_i - f_i^{(0)}.$$

Рассмотрим полупроводник с одним энергетическим уровнем в запрещенной зоне и с $\gamma_{cvii} \neq 0$. Тогда из (27) при $\dot{n}_i = 0$ получим стационарную функцию распределения f_i . Полагая, что квадратичные члены ($\sim \gamma_{cvii}$) малы, с линейной по γ_{cvii} точностью получим

$$f_i \approx f_i^{(0)} + \frac{2\gamma_{cvii} N_i}{W_i} \{ n(1-f_i^{(0)})^2 - \alpha_i^2 p f_i^{(0)2} \}. \quad (37)$$

Это выражение справедливо при $\delta f_i \ll f_i^{(0)}$. Подстановка в (37) данных из табл. I показывает, что, например, при низком уровне инжекции в полупроводнике n -типа населенность уровня E_1 очень сильно зависит от уровня инжекции $z \equiv p/n_0$ уже при $\gamma_{cvii} N_i \sim 10^{-17}$ см³/с, тогда как для уровня E_2 влияние нелинейности мало даже и при $\gamma_{cvii} N_i \sim 10^{-10}$ см³/с и не зависит от z . Таким образом, верхние уровни сильно опустошаются при повышении z за счет квадратичного по N_i процесса, если $\gamma_{cvii} \sim 10^{-32} - 10^{-33}$ см⁶/с.

Количественные оценки ΔR_n по формуле (36), например, при НУИ, показывают, что при определенных условиях поправка к $R_n^{(0)}$ может быть велика и зависит от z . Подробный количественный анализ влияния Оже-процессов второго порядка на времена жизни носителей будет дан в другом месте. Некоторые качественные черты этих процессов заключаются в следующем. Как процессы $\sim \gamma_{cvii}$, так и процессы, соответствующие взаимодействию уровней, приводят в стационарном режиме к некоторому изменению населенностей уровней, причем населенность верхних уровней уменьшается, а нижних увеличивается. Первый член в (36) является весьма чувствительным к изменению функций распределения, которые сами, в свою очередь, весьма чувствительны к уровню инжекции. Расчет, проведенный для модельной системы двух уровней с параметрами $E_1 = E_c - 0,55$ эВ, $E_2 = E_v + 0,35$ эВ, $C_{n1}^{SRH} = 2 \cdot 10^{-8}$, $C_{p1}^{SRH} = 10^{-7}$, $C_{n2}^{SRH} = 4 \cdot 10^{-8}$, $C_{p2}^{SRH} = 3 \cdot 10^{-8}$ см³/с, $C_{12}^{SRH} = 10^{-8}$ см³/с, все $\gamma \equiv 10^{-25}$ см⁶/с (кроме $\gamma_{cvii}: \gamma_{cvii} \equiv 0$) — для n - и p -Si при $T = 300$ К — показывает, что величины δf_i составляют в среднем несколько процентов (рис. 4). Принято также $N_1 = N_2 = 10^{16}$ см⁻³. Для времен жизни ННЗ в данном примере получено $\delta\tau \sim \delta f$, $\delta\tau > 0$.

3. Выводы и заключение

Приведенное рассмотрение вклада Оже-процессов с участием локализованных носителей в темп стационарной рекомбинации с определенностью показывают, что, в принципе, возможны условия, при которых Оже-процессы первого порядка являются глав-

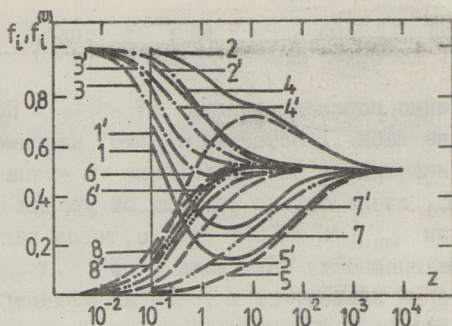


Рис. 4. Стационарные функции распределения для модели двух уровней:

$N_d = 10^{18} \text{ см}^{-3}$: 1 - f_1 , 2 - f_2 ; $N_d = 10^{18} \text{ см}^{-3}$;

$N_a = 10^{18} \text{ см}^{-3}$: 3 - f_1 , 4 - f_2 .

$N_a = 10^{18} \text{ см}^{-3}$: 5 - f_1 , 6 - f_2 .

$N_a = 10^{18} \text{ см}^{-3}$: 7 - f_1 , 8 - f_2 .

Примечание: линии со штрихованными индексами соответствуют функциям $f_i^{(0)}$.

ными, определяя времена жизни ННЗ (фактически сильно уменьшая их по сравнению со случаем ШРХ-рекомбинации). Что касается влияния Оже-процессов второго порядка на времена жизни ННЗ, то уже краткое рассмотрение простейшей двухуровневой системы показывает, что определить его значительно сложнее. Отметим, что теоретические значения коэффициентов для процессов первого порядка, полученные различными авторами, имеют значительный разброс, коэффициенты же для процессов второго порядка фактически неопределены. Значения всех коэффициентов, по-видимому, сильно зависят от природы центров, их концентраций и статистических распределений и, таким образом, от энергетической структуры запрещенной зоны полупроводника. В сильнолегированных полупроводниках со сложным энергетическим спектром запрещенной зоны эти процессы, по-видимому, должны играть заметную роль.

Л и т е р а т у р а

1. L a n d s b e r g P.T. Non-radiative transitions in semiconductors // Phys. Stat. Sol. - 1970. - Vol. 41. - P. 457-489.

2. S t o n e h a m A.M. Non-radiative transitions in semiconductors // Rep. Progr. Phys. - 1981. - Vol. 44. - P. 1251-1995.

3. Абакумов В.Н., Перель В.И., Ясевич И.Н. Захват носителей заряда на притягивающие центры в полупроводниках // ФТП, 1978. - Т. 12. - Вып. 1.- С. 3-32.

4. Бургуэн Ж., Ланно М. Точечные дефекты в полупроводниках. Экспериментальные аспекты / Пер. с англ. под ред. В.Л. Гуревича. - М.: Мир, 1985.

5. Robbins D.J., Landsberg P.T. Impact ionization and Auger recombination involving traps in semiconductors // J. Phys. C: Solid St. Phys. - 1980. - Vol. 13. - P. 2425-2439.

6. Rebsch J.-T. A combination of Auger and many-phonon processes in nonradiative recombination // Solid St. Comm. - 1979. - Vol. 31. - P. 377-381.

7. Опцкий А.П. К вопросу об Оже-рекомбинации в полупроводниках // Тр. Таллинск. политехн. ин-та. - 1987. - № 651. - С. 101-113.

A. Opataki

Lifetime of the Minority Charge Carriers in Semiconductors with Calculation of the Auger-processes on Localized States

Abstract

The influence of Auger-recombination involving localized states on the lifetime of the minority charge carriers is considered. It is shown that under certain conditions such processes may reduce considerably the lifetime of carriers in comparison with Shockley-Read-Hall-recombination.

Vähemus-laengukandjate eluiga pooljuhise
Auger'-protsesside arvestamisega lokaalsetes olekutes

Kokkuvõte

Kirjeldatakse Auger'-protsesside mõju vähemus-laengukandjate elueale, kui rekombinatsioon toimub läbi lokaalsete olekute.

Näidatakse, et teatud tingimustel võivad sellised protsessid oluliselt lühendada vähemus-laengukandjate eluiga Shockley-Read-Halli-elueaga võrreldes.

КОМПЛЕКСНЫЙ ПОДХОД К СИСТЕМАМ ПРОЕКТИРОВАНИЯ ЭЛЕМЕНТОВ ПОЛУПРОВОДНИКОВОЙ И ИНТЕГРАЛЬНОЙ ЭЛЕКТРОНИКИ

Введение. В настоящее время единственной возможностью уменьшить затраты на проектирование и первичное испытание является точное моделирование каждого этапа производства и испытания полупроводниковых приборов и интегральных схем. Как известно [1] в настоящее время самым перспективным направлением эффективного и дружественного человеку и природе является производственная система, которая базируется на концепции CIM (Computer Integrated Manufacturing). Эта концепция сильно связана с машинным проектированием (CAD). В данной работе делается попытка дать обобщенный подход к системе проектирования, исходя из концепции CIM. Для этого материал статьи разделяется по следующим условным названиям: производство – проектирование – система – модель – физическое явление. Первые два относятся к т.н. системному уровню, третье к переходному и последние два к объектному уровню (рис. 1).

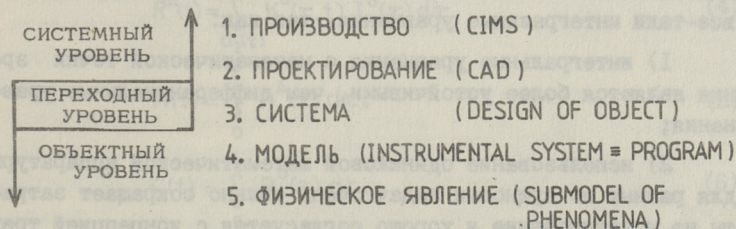


Рис. 1. Условные названия комплексного подхода к системам проектирования.

Принципы проектирования на системном уровне. В литературе (см., например, [2]) доказано, что на базе концепции CIM созданная производственная система CIMS представляет собой большую систему (LSS). Таким образом, общее проектирование разделяется на два больших подраздела — системное проектирование и объектное проектирование. Системное проектирование охватывает два главных аспекта дизайна системы — проектирование (дизайн) системы и проектирование функционирования системы. Можно сказать, что большие системы отвечают следующим требованиям — они многоуровневые, многоемкие и многоцелевые; входные величины интенсивного, а выходные экстенсивного характера.

Для того, чтобы получить идеологическую совокупность системного и объектного проектирования надо найти подходящую математическую аппаратуру, которую можно было бы использовать во многоцелевых направлениях, т.е. для моделирования производственной системы (CIMS), с одной стороны, а с другой стороны, была бы годной и для описания базы управления производственной системы (CIMS) — сеть ЭВМ. Кроме этого целесообразно найти такую математическую аппаратуру, с помощью которой можно описать систему проектирования (CAD) и ее физическую базу — локальную сеть ЭВМ.

Опыт показал, что моделирование вышеуказанных проблем приводит нас к дифференциальным уравнениям и вместе с этим к разным иерархическим моделям с соответствующей задачей оптимизации или к интегральным уравнениям и вместе с этим к теории развивающихся систем. Как и отмечается в литературе [3], использование одного или другого метода принципиально не отличается, но на наш взгляд можно предпочесть все-таки интегральные уравнения, так как:

1) интегральные уравнения с математической точки зрения являются более устойчивыми, чем дифференциальные уравнения;

2) использование одинаковой математической аппаратуры для разных по сущности задач очень сильно сокращает затраты на моделирование и хорошо согласуется с концепцией трехуровневого моделирования, разработанного нами.

Как известно [4], развивающиеся системы могут быть двух- или многопродуктные. Сама система, которая моделиру-

ется, должна соответствовать следующим двум очень важным требованиям:

1. Общую систему можно разделить на две зависимые одна от другой подсистемы.

2. Система развивается, исходя из внутренних действий системы.

Таким образом, для одной подсистемы двухпродуктного случая можно написать следующую модель:

$$R(t) = \int_{-\infty}^t K(\tau, t) I(\tau) d\tau \quad (1)$$

и для n -продуктного случая

$$R_j(t) = \sum_{j=1}^n \int_{-\infty}^t K_j(\tau, t) I_j(\tau) d\tau, \quad (2)$$

где R - вектор реакций системы;

I - вектор входных воздействий;

K - матрица переходных функций системы.

Данная математическая база обеспечивает нам возможность описать проблемы САД на обоих системных уровнях.

Предположим, что общая система разделяется на две подсистемы: А и В. Учитывая уравнение (1) для двухпродуктной модели, можно написать следующую математическую модель:

$$R^A(t) = \int_{a_1(t)}^t K^A(\tau, t) I^A(\tau) d\tau, \quad (3)$$

$$R^B(t) = \int_{a_2(t)}^t K^B(\tau, t) I^B(\tau) d\tau, \quad (4)$$

$$M(t) = \int_0^t R^A(\tau) d\tau, \quad (5)$$

$$C(t) = \int_0^t R^B(\tau) d\tau, \quad (6)$$

$$P(t) = \int_{a_1(t)}^t I^A(\tau) d\tau + \int_{a_2(t)}^t I^B(\tau) d\tau, \quad (7)$$

$$F(t) = k^A R^A(t) + k^B R^B(t). \quad (8)$$

Для p -продуктной модели уравнений (3)–(7) дополняются знаком суммирования. Теперь рассмотрим, например, как интерпретировать величины модели (3)–(8) на уровне "производство".

Пусть подсистема А описывает систему управления и система В соответствует системе производства. В этом случае производственная система является дискретной и состоит еще из других подсистем, точное моделирование которых не является нашей целью.

$K^A(\tau, t)$ показатель эффективности функционирования подсистемы А: доля $R^A(t)$, создаваемая в расчете на одно WP момента τ в А и WP , обозначает место выполнения определяемой операции (во многих случаях WP можно сравнивать с понятием рабочего места). $K^B(\tau, t)$ показатель эффективности подсистемы В: доля $R^B(t)$, создаваемая в расчете на одно WP момента τ в В; K^A и K^B можно интерпретировать как передаточную матрицу. $I^A(\tau)$ и $I^B(\tau)$ интерпретируется как количество новых WP в единицу времени момента τ , направляемого в подсистему А и В соответственно. Явно, что $I^B(\tau) = 1 - I^A(\tau)$, $q_1(t)$ и $q_2(t)$ являются временными границами ликвидации в момент t устаревших WP в подсистеме А и В соответственно. $P(t)$ общее количество всех функционирующих WP в системе (А+В) в момент t . $F(t)$ количество новых ресурсов на единицу времени t , поступающих в систему (А+В), а k^A и k^B эмпирические коэффициенты модели. Для такой системы обычно определяется задача оптимизации, которая имеет следующий вид:

$$\max = \int_{t_0}^T C(t) dt, \quad (9)$$

где $[t_0, T]$ период моделирования.

Теперь сделаем попытку интерпретировать параметры моделей (3)–(8) для уровня проектирования. На рис. 2 показана связь данного уровня через специализированную базу данных с общей системой, точнее специализированная база данных связывается с общей системой (СІМ) через минимальный банк данных системы. Система состоит из двух подсистем – подси-

стема А соответствует системе проектирования РС, а подсистема В системе проектирования DD.

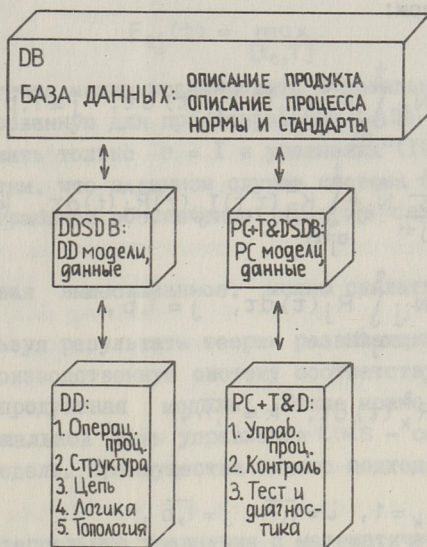


Рис. 2. Уровень проектирования (CAD) и связь через DB.

$K^A(\tau, t)$, $K^B(\tau, t)$ – показатели эффективности функционирования подсистемы А и В соответственно, т.е. передаточные матрицы. Например, эти показатели характеризуют быстроту работы и объема памяти используемых ЭВМ. $I^A(\tau)$ и $I^B(\tau)$ интерпретируются как число новых программ в единицу времени в подсистемах А и В соответственно, что по существу соответствует разработке новых более эффективных пакетов прикладных программ (ППП). $\alpha_1(t)$, $\alpha_2(t)$ и $P(t)$ имеют такую же интерпретацию как выше. Под $F(t)$ можно подразумевать и новые ЭВМ, более мощные операционные системы и т.д. Задача оптимизации также совпадает с вышеуказанными, но можно предложить и более сложные задачи.

Материальной базой систем обоих уровней являются сети ЭВМ, где в первом случае говорят о так называемой глобальной системной сети и во втором случае локальной сети. Работы обеих сетей можно моделировать с помощью общей модели (2), т.е. с помощью п-продуктной модели.

Целесообразно ввести в модель некоторые дополнения. В этом случае математическую модель сети ЭВМ можно описать следующим образом:

$$R_j^A(t) = \sum_{i=1}^r N_{ij} \int_{a_j(t)}^t K_{ij}^A(\tau, t) I_{ij}^A(\tau) d\tau, \quad j = \overline{1, p}, \quad (I0)$$

$$R_k^B(t) = \sum_{i=1}^r \sum_{j=1}^r N_{ij} \int_{a_j(t)}^t K_{jk}^B(\tau, t) I_{jk}(\tau) R_j^A(t) d\tau, \quad k = \overline{1, n}, \quad (II)$$

$$F_{ij}^A(t) = \sum_{i=1}^r N_{ij} \int_{a_j(t)}^t R_j^A(\tau) d\tau, \quad j = \overline{1, p}, \quad (I2)$$

$$F_{2k}^B(t) = \int_{t_0}^t R_k^B(\tau) d\tau, \quad k = \overline{1, n}, \quad (I3)$$

$$Y_{ij} + \sum_{k=1}^n I_{jk} = 1, \quad i = \overline{1, r}, \quad j = \overline{1, p}. \quad (I4)$$

где r — число центральных управляющих ЭВМ, т.е. число подразделений в системе (CIMS), где имеется самостоятельная вычислительная мощность.

В каждом этом подразделении имеется p -тип разных типов ЭВМ. Величина N_{kl} это число ЭВМ l -го типа в k -ом подразделении ($l = \overline{1, p}, k = \overline{1, r}$). Сеть ЭВМ подразделений производит k видов полезных продуктов, таких как научно-технические расчеты, календарное планирование и т.д. Программное обеспечение (ПО) разделим на две подсистемы. Подсистема А ПО общего вида и подсистема В ПО прикладного характера. $R_j^A(t)$ интерпретируется как скорость производства ПО одной ЭВМ j -го типа, $I_{ij}^A(t)$ относительная доля общего ПО в объеме всего ПО, производимого в момент t на одной ЭВМ j -го типа в i -ом подразделении. $I_{jk}^B(t)$ доля прикладного ПО, предназначенного для производства k -го полезного продукта ($k = \overline{1, n}$); $R_k^B(t)$ скорость производства k -го продукта во всей сети. Величина $a_j(t)$ обозначает временную границу ликвидации устаревшего ПО для ЭВМ j -го типа, $K_{ij}^A(t, \tau)$ и $K_{jk}^B(t, \tau)$ производительность общего и прикладного ПО для

ЭВМ j -го типа в i -ом подразделении на одной ЭВМ (j -го типа). $F_{ij}(t)$ работающее ПО на j -ом ЭВМ и $F_{2j}(t)$ объем расчетов. Задача оптимизации может быть определена следующим образом:

$$F_{2j}(t) = \max_{[t_0, T]}. \quad (I5)$$

Аналогично можно моделировать и локальную сеть ЭВМ, специализированную для проектирования (CAD). В этом случае надо принимать только $r = I$ и уравнения (I0) – (I4) упрощаются. Отметим, что в данном случае система В производит заказное программное обеспечение (custom design software или ППП).

Учитывая вышесказанное, можно сделать следующий вывод.

Используя результаты теории развивающихся систем, можно описать производственную систему соответствующей концепции СИМ (двухпродуктная модель). Также можно реализовать модель материальной базы управления СИМС – сеть ЭВМ на p -продуктной модели. Преимущества такого подхода состоят в следующем:

а) интегральные уравнения с математической точки зрения являются более устойчивыми, чем дифференциальные уравнения, и по мнению автора описывают проблему компактнее, чем дифференциальные уравнения;

б) используя одинаковую математическую аппаратуру, можно моделировать очень разные проблемы по существу, что намного сокращает затраты на моделирование.

Принципы проектирования на переходном уровне. Как указывает уже само название, на этом уровне действительно соединяются проблемы, вытекающие из системного подхода и из объектного подхода. Системы проектирования элементов полупроводниковой и интегральной электроники обычно рассматриваются как иерархические. Но, к сожалению, по нашему мнению такой подход является неправильным, т.е. неправильной является интерпретация использования понятия иерархического моделирования (см. например, [I3]).

Предлагаем новую, более подходящую, а также более точную по существу иерархическую структуру для описания системы проектирования. Таким образом, системы проектирования можно

разделить на следующие иерархические уровни: алгоритмический, модельный и цифровой. Такой подход разрешает построить системы проектирования с более эффективным построением. Имеется возможность построить системы с так называемыми возможностями, т.е. пользователь имеет возможность выбирать для своей задачи самый подходящий алгоритм, самые подходящие модели, а также по вкусу экспериментальные данные.

Из вышесказанного следует, что конкретная система проектирования (в нашем случае CADSEDIC) получает модульный характер. Как известно [5], модули конкретного пакета определяются на основе модульного анализа рассматриваемой задачи, т.е. в результате разбиения сложной задачи на ряд более простых. Единственным требованием является то, что и при более простых задачах (по существу это отдельные ППП) должна сохраняться трехуровневая иерархическая структура.

Пусть имеется множество ППП:

$$P = \{P_1, P_2, \dots, P_k\}, \quad (I6)$$

где P_i — наименование конкретного пакета.

В системе CADSEDIC $P_1 = EPIM^*$, $P_2 = DIOXM$, $P_3 = ETCHM$, $P_4 = IMPLM$, $P_5 = NI$, $P_6 = INTER1$, $P_7 = DYNAMIT$, $P_8 = SCHOTSIM^*$, $P_9 = FAAS^*$, $P_{10} = INTER2$, $P_{11} = TRANZ-TRAN$, $P_{13} = NAP2^*$, $P_{14} = INTER3$ (* — отсутствуют в рабочей версии в Будапеште).

Известно, что ППП по заданному ей набору данных (входные данные):

$$I_i = \{I_{i1}, I_{i2}, \dots, I_{i\ell_i}\} \quad (I7)$$

формирует заданным образом набор данных (результаты)

$$O_i = \{O_{i1}, O_{i2}, \dots, O_{iN_i}\}. \quad (I8)$$

Множества наборных данных имеют вид:

$$I = \{I_1, I_2, \dots, I_k\} \text{ и } O = \{O_1, O_2, \dots, O_\ell\}. \quad (I9)$$

Для решения задачи надо выполнить цепочку последовательностей следующего вида:

$$R_j = P_i \rightarrow P_{i+M_1} \rightarrow \dots \rightarrow P_{i+M_j}, \quad (20)$$

где $i + M_j \leq k$.

Следует организовать вычисления по цепочке типа (20) так, чтобы обеспечить ее экономическое многократное функционирование с минимальными затратами на организацию и проведение вычислений с сохранением указанных I_{ij} и O_{ij} . В данном случае должна быть известна схема взаимодействия модулей и установлена дисциплина работы с пакетом и его компонентами. С этой целью создается граф-модель системы проектирования (на рис. 3 показана главная граф-модель, где на правом краю показана иерархия в обыкновенном смысле).

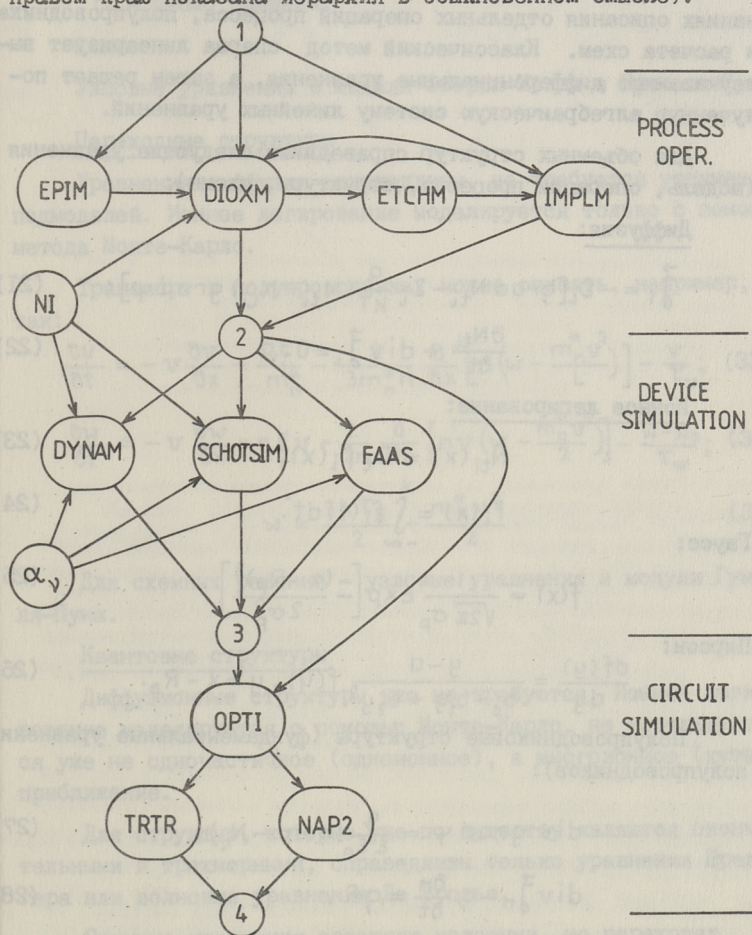


Рис. 3. Главная граф-модель системы проектирования (связи системы CADSEDIC).

Принципы проектирования на объектном уровне. Новый универсальный трехуровневый подход реализуется и на объектном уровне. Кроме этого мы должны соблюдать размерность моделируемых структур. В работе [II] представлена новая, более полная классификация структур по размерам: объемные, переходные и квантовые структуры. Для объемных приборов подходит классический подход и классические модели.

Классическая модель базируется на фундаментальных уравнениях описания отдельных операций процесса, полупроводника и расчета схем. Классический метод сперва линейризует вышеуказанные дифференциальные уравнения, а затем решает полученную алгебраическую систему линейных уравнений.

Для объемных структур справедливы следующие уравнения (модель, операции процесса, структуры, схемы):

Диффузия:

$$\vec{j}_i = -D_i [\text{grad } N_{ti} - z_i \frac{q}{kT} (N_{ti} - N_{ci}) \text{grad } \psi]; \quad (21)$$

$$\frac{\partial N_{ti}}{\partial t} + \text{div } \vec{j}_i = 0. \quad (22)$$

Ионное легирование:

$$N_{ti}(x) = D_t f_i(x); \quad (23)$$

$$F_i(x) = \int_{-\infty}^x f_i(t) dt. \quad (24)$$

Гаусс:

$$f(x) = \frac{1}{\sqrt{2\pi} \sigma_p} \exp \left[-\frac{(x - R_p)^2}{2\sigma_p^2} \right]. \quad (25)$$

Пирсон:

$$\frac{df(y)}{dy} = \frac{y - a}{b_0 + b_1 y + b_2 y^2} f(y), \quad y = x - R_p. \quad (26)$$

Полупроводниковые структуры (фундаментальные уравнения полупроводников):

$$\text{div grad } \psi = \frac{q}{\epsilon \epsilon_0} (n - p - N_t), \quad (27)$$

$$\text{div } \vec{j}_n - q \frac{\partial n}{\partial t} = q R, \quad (28)$$

$$\text{div } \vec{j}_p + q \frac{\partial p}{\partial t} = -q R, \quad (29)$$

$$\vec{j}_n = q n \mu_n \vec{E}_n + q D_n \text{grad } n, \quad (30)$$

$$\vec{j}_p = q p \mu_p \vec{E}_p - q D_p \text{grad } p, \quad (31)$$

где

$$N_t = \sum_{j=1}^n N_{ti};$$

n — число примеси.

Схемы:

Узловые уравнения и модели Эберса-Молла и Гуммеля-Пуна:

Переходные структуры

Уравнения диффузии справедливы, но требуется уточнение подмоделей. Ионное легирование моделируется только с помощью метода Монте-Карло.

Транспорт в полупроводниках можно описать, например, так:

$$\frac{\partial v}{\partial t} = -v \frac{\partial v}{\partial x} + \frac{qE}{m_n^*} - \frac{2}{3m_n^* n} \frac{\partial}{\partial x} \left[n \left(w - \frac{m_n^* v^2}{L} \right) \right] - \frac{v}{\tau_v}; \quad (33)$$

$$\frac{\partial w}{\partial t} = -v \frac{\partial w}{\partial x} + qE v - \frac{2}{2n} \frac{\partial}{\partial x} \left[n v \left(w - \frac{m_n^* v^2}{2} \right) \right] - \frac{w - w_0}{\tau_w}; \quad (34)$$

$$w = \frac{3kT_n}{2} + \frac{m_n^* v^2}{2}. \quad (35)$$

Для схемных решений — узловые уравнения и модули Гуммеля-Пуна.

Квантовые структуры

Диффузионные структуры уже не требуются. Ионное легирование моделируется с помощью Монте-Карло, но используется уже не одночастичное (одноионное), а многоионное (кучек) приближение.

Для структур, которые уже по существу являются окончательными и трехмерными, справедливы только уравнения Шредингера или волновые уравнения Де Бройля.

Схемные уравнения остаются узловыми, но переходят к функциональным моделям.

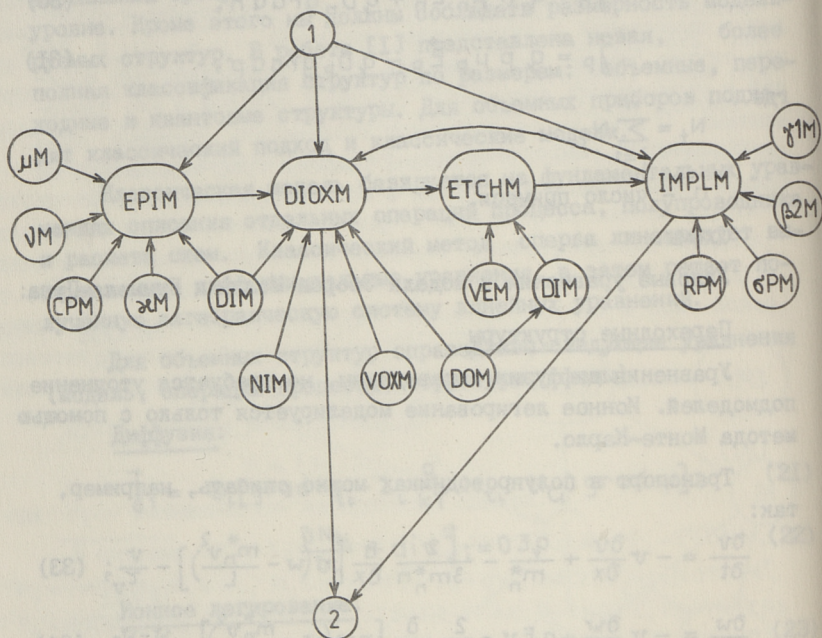


Рис. 4. Полная граф-модель для моделирования разных технологических операций процесса (связи системы STEP):

μM – модель кинетического коэффициента; νM – модель коэффициента вязкости; αM – модель коэффициента теплопроводности; CPM – модель удельной теплоемкости; DIM – модель полного коэффициента диффузии; NIM – модель собственной конденсации носителей; $VOXM$ – модель скорости роста толщины оксидного слоя; DOM – модель коэффициента диффузии (закон Архениуса); VEM – модель скорости травления; RPM – модель для R_p ; σPM – модель для σ_p ; $\beta 2M$ – модель для β_2 ; $\gamma 1M$ – модель для γ_1).

Теперь коротко остановимся на некоторых проблемах, связанных с классическими моделями.

На рис. 3 приведен только главный граф системы, но в действительности полный граф разных операций содержит 13 вершин, что показано на рис. 4.

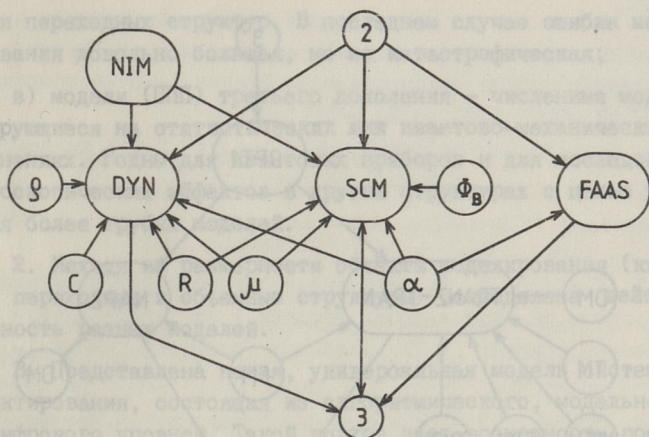


Рис. 5. Полная граф-модель для моделирования полупроводниковых структур (связи системы DEMO):

ρ - модель удельной теплоты; C - модель тепловой емкости; R - модель рекомбинации носителей; μ - модель подвижности носителей; α - модель лавинного размножения носителей; Φ_B - модель высоты барьера Шоттки; NIM - модель собственной концентрации носителей.

На рис. 5 показана полная граф-модель для моделирования структур. Используется 7 разных моделей, которые описывают разные физические явления. Как показано в работах [6-9], расхождение результатов таких моделей очень сильное и для адекватного моделирования структур требуется теоретическое и экспериментальное исследование еще сильнее.

На рис. 6 показана полная модель для моделирования схемных решений. Как видно, используется 6 разных моделей, которые в данном случае не описывают поведение физического (математического) явления, а описывают разные полупроводниковые приборы. Адекватность моделирования полупроводниковых приборов с помощью эквивалентных схем, как показано в работе [10], является очень важным моментом. Поэтому данная часть содержит и программу OPTI [11], которая выполняет оптимизацию для параметров полупроводниковых приборов, как это показано в работе [12].

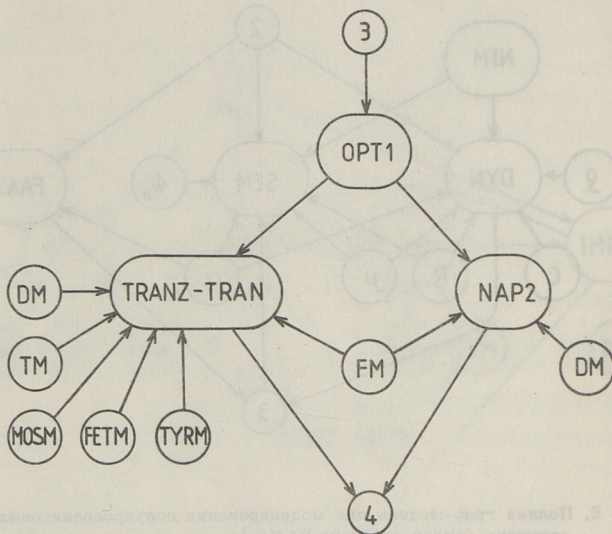


Рис. 6. Полная граф-модель для моделирования схемных решений (связи системы TARZAN):

DM – модель диода; TM – модель биполярного транзистора; MOSM – модель МОП транзистора; FETM – модель полевого транзистора с рп-переходом; TYRM – модель тиристора; FM – функциональные модели; DM – приборные модели.

Заключение. Итак, дан обзор о главных моментах моделирования элементов полупроводниковой и интегральной электроники (ЭПИЭ). Разработан новый подход к моделированию структур и созданию систем проектирования ЭПИЭ. На основании вышеизложенного можно сделать следующие окончательные выводы:

1. Определен уровень развития моделирования отдельных ППП, исходя из сущности и построения. На этой основе ППП разделяются на следующие три класса:

а) модели (ППП) первого поколения главным образом аналитические и аналитико-численные, годны только для объемных структур;

б) модели (ППП) второго поколения – численные, базирующиеся на дифференциальных уравнениях, годны для объем-

ных и переходных структур. В последнем случае ошибка моделирования довольно большая, но не катастрофическая;

в) модели (ППП) третьего поколения – численные модели, базирующиеся на статистических или квантово-механических уравнениях. Годны для квантовых приборов и для исследования микроскопических эффектов в других структурах с целью уточнения более грубых моделей.

2. Исходя из размерности объекта моделирования (квантовые, переходные и объемные структуры) определена действительность разных моделей.

3. Представлена новая, универсальная модель системы проектирования, состоящая из алгоритмического, модельного и цифрового уровней. Такой подход дает возможность построить систему проектирования, учитывая элементы создания искусственного интеллекта.

4. Для системного анализа более подходящими являются модели, базирующиеся на интегральных уравнениях (теория развивающихся систем). Таким образом, на единой математической базе можно реализовать очень разные задачи, как, например, системный анализ CAD или CIM и анализ материальной базы CAD и CIM – сети ЭВМ.

Л и т е р а т у р а

1. Р а н г Т. Инженерное обеспечение САПР элементов РЭА : Сб. тез. респ. совещ. Численные методы и средства проектирования и испытания элементов РЭА. – Таллин, 7-II дек. 1987. – С. 12-23.

2. Y o s h i k a w a Н. Computer integrated manufacturing // Comput. in Industry. – 1987. 8, – N 283. – P. 71-85.

3. Г л у ш к о в В.М., П о п о в А.А., П е т р у х и н В.А. Системный подход к моделированию в медицине // Кибернетика и вычисл. техн. – Киев: Наукова думка, 1977. – Вып. 36. – С. 3-6.

4. Г л у ш к о в В.М., И в а н о в В.В., Я н е н к о В.М. Моделирование развивающихся систем. – М.: Наука, 1983. – 350 с.

5. Я н е н к о Н.Н. Вопросы модульного анализа и параллельных вычислений в задачах математической физики: Комплексы программ математической физики. - Новосибирск: ИПМ СО АН СССР, 1980. - С. 3-II.

6. R a n g Т. Computer aided examination of the avalanche multiplication in Si, GaAs and InP in complementary pn-junctions // Modeling, Simulation and Control, A, AMSE Press. - 1985. - Vol. 3, N 1.-P. 23-40.

7. Р а н г Т. Моделирование подвижностей носителей I. Рассеивание на акустических фононах и на нейтральной и ионизированной примеси // Тр. Таллинск. политехн. ин-та. - 1986. - № 629. - С. I03-II6.

8. Р а н г Т. Моделирование подвижностей носителей II. Рассеивание на носителях и поверхности, эффект насыщения дрейфовой скорости носителей // Тр. Таллинск. политехн. ин-та. - 1986. - № 629. - С. II7-I30.

9. Р а н г Т. Моделирование концентрации носителей заряда в полупроводниках // Тр. Таллинск. политехн. ин-та. - 1987. - № 639. - С. 26-36.

10. Р а н г Т. Влияние параметров эквивалентных моделей на характеристики полупроводниковых приборов // Тр. Таллинск. политехн. ин-та. - 1987.-№ 643. - С. 72-76.

11. Р а н г Т. Программа определения параметров "OPTI" // Тр. Таллинск. политехн. ин-та. - 1986. - № 617. - С. 21-26.

12. Р а н г Т. Сравнение разных методов оптимизации // Тр. Таллинск. политехн. ин-та. - 1987. - № 643. - С. 77-84.

13. К о р я ч к о В.П., К у р е й ч и к В.М., Н о р е н к о в И.П. Теоретические основы САПР. - М.: Энергоатомиздат, 1987. - 398 с.

Computer Aided Design System for Electron Devices
and Integrated Circuits

Abstract

The developing of the CAD systems consists of many parts. Five key-words (production - design - system - model - phenomena) have been chosen to give the explanation. A hierarchical three-level model for the design system has been proposed, which consists of algorithmic, model and numerical levels.

T. Rang

Pooljuhtseadiste ja integraalskeemide
projekteerimissüsteem

Kokkuvõte

Projekteerimissüsteemide (CAD) loomine koosneb mitmest etapist. Probleemi selgitamiseks on valitud järgmised võtmesõnad: tootmine - kujundamine - süsteem - mudel - nähtus. Esitatakse hierarhiline algoritm projekteerimissüsteemi kirjeldamiseks, mis koosneb algoritmilisest, mudelilisest ja numbrilisest tasemest.

ФИЗИЧЕСКИЕ АСПЕКТЫ УЧЕТА ВЛИЯНИЯ ЭФФЕКТА ЭЛЕКТРОННО-ДЫРОЧНОГО РАССЕЯНИЯ ПРИ МАТЕМАТИЧЕСКОМ МОДЕЛИРОВАНИИ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР

Феноменологически адекватный учет влияния электронно-дырочного рассеяния (ЭДР) на транспорт носителей заряда (НЗ) в полупроводниках требует не только перерасчета подвижности НЗ [1], но и модификации уравнений токов электронов и дырок [2]. Начиная с конца 60-х годов, этим подходом стали широко пользоваться отечественные авторы, в частности, при аналитическом исследовании характеристик силовых полупроводниковых приборов. Зарубежные авторы придерживались традиционного подхода Флетчера [1]. В практику численного моделирования модифицированные уравнения токов (подход Авакьянца и др. [2]) были впервые введены авторами работ [3, 4].

В данной работе рассмотрим некоторые физические аспекты представления уравнения токов НЗ и определения соответствующих кинетических коэффициентов (подвижностей и коэффициентов диффузии).

Плотность тока НЗ в полупроводнике можно записать в следующем общем виде [5]:

$$j_{\alpha} = q_{\alpha} N_{\alpha} \mu_{\alpha} c_{\alpha}, \quad (I)$$

где α, β — индексы типа НЗ ($\alpha, \beta \rightarrow n$ или p ; $\alpha \beta \rightarrow np$ или pn , но не nn или pp);

q_{α} — знак заряда НЗ ($q_n = -1$, $q_p = 1$);

c_{α} — концентрация НЗ ($c_n \equiv n$, $c_p \equiv p$).

Сила, действующая на НЗ при условии изотермичности ($T_{\alpha} = T$ и $\text{grad} T = 0$) и в отсутствие магнитного поля, в одномерном приближении:

$$H_{\alpha} = q_{\alpha} q E_{\alpha} - k_B T F_{\alpha} \frac{1}{c_{\alpha}} \frac{dc_{\alpha}}{dx}, \quad (2)$$

где E_{α} — напряженность электрического поля, действующая на НЗ типа α ;

q — единичный заряд ($q = |e|$);

k_B — постоянная Больцмана.

Фактор F_{α} учитывает возможное вырождение газа НЗ в изотропной параболической зоне

$$F_{\alpha} = \frac{2\mathcal{F}_{1/2}(\eta_{\alpha})}{\mathcal{F}_{-1/2}(\eta_{\alpha})},$$

где $\mathcal{F}_r(\eta_{\alpha})$ — интеграл Ферми;

η_{α} — приведенный квазиуровень Ферми НЗ типа α .

Обобщенный кинетический коэффициент μ_{α} , входящий в уравнение (I), имеет следующий вид:

$$\mu_{\alpha} = \frac{q_{\alpha} q H_n}{c_{\alpha} H_{\alpha}} \cdot \frac{\begin{vmatrix} 0 & b'_{\alpha} \\ b & G \end{vmatrix}}{|G|}, \quad (3)$$

где векторы (штрих обозначает транспонирование)

$$\begin{aligned} b'_n &= (\beta_{n0}, 0, \beta_{n1}, 0, \dots, \beta_{ns_1}, 0), \\ b'_p &= (0, \beta_{p0}, 0, \beta_{p1}, \dots, 0, \beta_{ps_1}), \\ b' &= (\beta_{n0}, -R\beta_{p0}, \beta_{n1}, -R\beta_{p1}, \dots, \beta_{ns_1}, -R\beta_{ps_1}) \end{aligned} \quad (4)$$

и

$$R = H_p / H_n.$$

Элементы блочно-квадратичной матрицы G содержат интегралы столкновений для различных типов рассеивающих центров:

$$G = \begin{pmatrix} G_{00} & G_{01} & \dots & G_{0s_1} \\ G_{10} & G_{11} & \dots & G_{1s_1} \\ \vdots & \vdots & & \vdots \\ G_{s_1 0} & G_{s_1 1} & \dots & G_{s_1 s_1} \end{pmatrix}, \quad G_{rs} = \begin{pmatrix} \Gamma_{rs}^{nn} & \Gamma_{rs}^{np} \\ \Gamma_{rs}^{pn} & \Gamma_{rs}^{pp} \end{pmatrix}.$$

Детальную расшифровку использованных символов можно найти в [5].

Обобщенный кинетический коэффициент μ_{α} по формуле (3) соответствует s_1 -приближению к решению кинетического уравнения Больцмана, полученному с помощью вариационного метода

Колера [6]. Бесконечное приближение ($s_1 \rightarrow \infty$) приводит к точному решению уравнения Больцмана. Отметим еще, что величину μ_α нельзя рассматривать как обыкновенную подвижность, поскольку она содержит силовой фактор R по формуле (4). Для раскрытия этой зависимости используем разложение определителя в (3) по первому столбцу:

$$\begin{vmatrix} 0 & b'_\alpha \\ b & g \end{vmatrix} = -\frac{n}{m_n} \sum_{s=0}^{s_1} \beta'_{ns} D_{2s+2,1} - \frac{H_p}{H_n} \cdot \frac{p}{m_p} \sum_{s=0}^{s_1} \beta'_{ps} D_{2s+3,1}, \quad (5)$$

где D_α - миноры определителя, если вычеркнуть $(2s+2)$ -ю или $(2s+3)$ -ю строку и первый столбец;

m_α - эффективная масса НЗ типа α .

и
$$\beta'_{\alpha s} = \frac{m_\alpha \beta_{\alpha s}}{c_\alpha}.$$

Подставляя (5), (4) и (3) в (I), получим

$$\begin{aligned} j_n &= -n H_n \mu_{n2} - p H_p \mu_{n3}, \\ j_p &= p H_p \mu_{p2} + n H_n \mu_{p3}, \end{aligned} \quad (6)$$

где

$$\begin{aligned} \mu_{n2} &= \frac{q}{m_n} \cdot \frac{\sum_{s=0}^{s_1} \beta'_{ns} D_{2s+2,1}}{|G|}, & \mu_{n3} &= \frac{q}{m_p} \cdot \frac{\sum_{s=0}^{s_1} \beta'_{ps} D_{2s+3,1}}{|G|}, \\ \mu_{p2} &= \frac{-q}{m_p} \cdot \frac{\sum_{s=0}^{s_1} \beta'_{ps} D_{2s+3,1}}{|G|}, & \mu_{p3} &= \frac{-q}{m_n} \cdot \frac{\sum_{s=0}^{s_1} \beta'_{ns} D_{2s+2,1}}{|G|}. \end{aligned} \quad (7)$$

Кинетические коэффициенты $\mu_{\alpha 2}$ и $\mu_{\alpha 3}$ уже не зависят от силового фактора R и имеют физический смысл подвижности. Сомножители с общим видом $\pm \sum / |G|$ можно рассматривать как некоторые усредненные времена релаксации.

В случае отсутствия ЭДР из (I) получим

$$j_{no} = -H_n \mu_{no} n, \quad j_{po} = H_p \mu_{po} p. \quad (8)$$

Определяя из (8) силы H_α и подставляя их в (6), получим

$$\begin{aligned} j_n &= \beta_{nn} j_{no} - \beta_{np} j_{po}, \\ j_p &= -\beta_{pn} j_{no} + \beta_{pp} j_{po}, \end{aligned} \quad (9)$$

где

$$\begin{aligned}\beta_{nn} &= \frac{\mu_{n2}}{\mu_{n0}}, & \beta_{np} &= \frac{\mu_{n3}}{\mu_{p0}}, \\ \beta_{pn} &= \frac{\mu_{p3}}{\mu_{n0}}, & \beta_{pp} &= \frac{\mu_{p2}}{\mu_{p0}}.\end{aligned}\quad (I0)$$

Уравнения токов, представленные в виде (9), были приняты за основу численной модели, описанной в [4].

Уравнения токов (9) можно представить в виде матричного уравнения:

$$\mathbf{j} = \mathbf{B} \mathbf{j}_0, \quad (II)$$

где

$$\mathbf{j} = \begin{pmatrix} j_n \\ j_p \end{pmatrix}, \quad \mathbf{j}_0 = \begin{pmatrix} j_{n0} \\ j_{p0} \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} \beta_{nn} & -\beta_{np} \\ -\beta_{pn} & \beta_{pp} \end{pmatrix}.$$

Согласно классической феноменологии явлений переноса при пренебрежении ЭДР ток НЗ можно представить в виде суммы дрейфовой и диффузионной составляющей. В случае принятой здесь матричной записи

$$\begin{aligned}\mathbf{j}_0 &= \mathbf{j}_{0dp} + \mathbf{j}_{0dif}, \\ \mathbf{j}_{0dp} &= q M_0 \mathbf{E} \mathbf{c}, \\ \mathbf{j}_{0dif} &= k_B T M_0 F \frac{d\mathbf{c}}{dx},\end{aligned}\quad (I2)$$

где

$$M_0 = \begin{pmatrix} \mu_{n0} & 0 \\ 0 & \mu_{p0} \end{pmatrix}, \quad \mathbf{E} = \begin{pmatrix} E_n & 0 \\ 0 & E_p \end{pmatrix}, \quad \mathbf{F} = \begin{pmatrix} F_n & 0 \\ 0 & -F_p \end{pmatrix}, \quad \mathbf{c} = \begin{pmatrix} n \\ p \end{pmatrix}.$$

Подставляя (I2) в (II), получим

$$\mathbf{j} = q M \mathbf{E} \mathbf{c} + k_B T M F \frac{d\mathbf{c}}{dx}, \quad (I3)$$

где

$$\mathbf{M} = \mathbf{B} \mathbf{M}_0.$$

Из последнего соотношения видна важность матрицы $\mathbf{B}$, преобразующей матрицу подвижностей M_0 (без учета ЭДР) в матрицу подвижностей $\mathbf{M}$ (с учетом ЭДР):

$$\mathbf{M} = \begin{pmatrix} \mu_{n2} & -\mu_{n3} \\ -\mu_{p3} & \mu_{p2} \end{pmatrix}.$$

В случае отсутствия ЭДР недиагональные элементы матриц $\mathbf{B}$ и $\mathbf{M}$ становятся равными нулю.

Обозначив в (I3) диффузионный коэффициент

$$\frac{1}{q} k_B T M F \equiv D,$$

можно записать соотношение Эйнштейна в матричной форме

$$M^{-1} D = \frac{1}{q} k_B T F.$$

Аналогичный результат был получен в работе [7], но при менее общих предположениях.

Из матричного уравнения (I3) легко получить токи электронов и дырок в общем виде:

$$j_n = q n \mu_{n2} E_n - q p \mu_{n3} E_p + k_B T \mu_{n2} F_n \frac{dn}{dx} + k_B T \mu_{n3} F_p \frac{dp}{dx};$$

$$j_p = q p \mu_{p2} E_p - q n \mu_{p3} E_n - k_B T \mu_{p2} F_p \frac{dp}{dx} - k_B T \mu_{p3} F_n \frac{dn}{dx}.$$

Эти уравнения учитывают вырождение газов НЗ, влияние квазиэлектрических полей (сужение запрещенной зоны) и эффект увлечения неосновных НЗ основными.

В отсутствие квазиэлектрических полей $E_n = E_p = E$ и

$$j_n = q n \mu_{n1} E + k_B T \mu_{n2} F_n \frac{dn}{dx} + k_B T \mu_{n3} F_p \frac{dp}{dx}, \quad (I4)$$

$$j_p = q p \mu_{p1} E - k_B T \mu_{p2} F_p \frac{dp}{dx} - k_B T \mu_{p3} F_n \frac{dn}{dx},$$

где $\mu_{n1} = \mu_{n2} - \frac{p}{n} \mu_{n3}, \quad \mu_{p1} = \mu_{p2} - \frac{n}{p} \mu_{p3}.$

К аналогичным уравнениям токов приводит феноменологический подход, развитый Авакьянцем и др., в работах [2, 8].

Использование физически адекватных уравнений тока (9) для количественных расчетов требует знания элементов матрицы В или, другими словами, подвижностей $\mu_{\alpha 2}$ и $\mu_{\alpha 3}$. Конечно, их можно определить теоретически, решив соответствующие интегралы столкновений и приняв приближение достаточно высокого порядка к вариационному решению уравнения Больцмана. Однако это ведет к очень громоздким выражениям, точность которых не гарантирована ввиду трудностей задания некоторых параметров процессов на микроскопическом уровне описания.

В случае приближения нулевого порядка можно, исходя из (7) и (I0), элементы матрицы В записать в следующем достаточно простом виде:

$$\begin{aligned}\beta_{nn}^{(0)} &= |B^{(0)}| (1 + n \mu_{p0}^{(0)} J^{(0)}), & \beta_{np}^{(0)} &= |B^{(0)}| n \mu_{n0}^{(0)} J^{(0)}, \\ \beta_{pn}^{(0)} &= |B^{(0)}| p \mu_{p0}^{(0)} J^{(0)}, & \beta_{pp}^{(0)} &= |B^{(0)}| (1 + p \mu_{n0}^{(0)} J^{(0)}),\end{aligned}\quad (I5)$$

где

$$|B^{(0)}| = [1 + (p \mu_{n0}^{(0)} + n \mu_{p0}^{(0)}) J^{(0)}]^{-1},$$

$$\mu_{\alpha 0}^{(0)} = \frac{q c_{\alpha}}{m_{\alpha}^2 d_{00}^{\alpha}}, \quad J^{(0)} = - \frac{m_n m_p}{q n p} \Gamma_{00}^{np},$$

$\Gamma_{00}^{np} = \Gamma_{00}^{pn}$ — интеграл ЭД-столкновений;

$d_{00}^{\alpha\alpha}$ — интеграл столкновений электронов (дырок) с центрами рассеяния вне электронно-дырочной системы (входит в элементы $\Gamma_{00}^{\alpha\alpha}$ матрицы G).

Из (I5) легко получить выражения для подвижностей $\mu_{\alpha 2}$, $\mu_{\alpha 3}$ и $\mu_{\alpha 1}$. Если приравнять приближения нулевого порядка к реальным величинам подвижностей ($\mu_{\alpha 0}^{(0)} = \mu_{\alpha 0}$) и обозначить $(n J^{(0)})^{-1} = \mu_{pn}$ и $(p J^{(0)})^{-1} = \mu_{np}$, придем к известным выражениям подвижностей, используемых при анализе полупроводниковых структур, следуя подходу Авакьянца и др.

Поскольку теоретический расчет подвижностей в случае ЭДР затруднителен, оказывается целесообразным использовать в практических расчетах полуэмпирические зависимости подвижностей $\mu_{\alpha 2, \alpha 3} = f(\mu_{\alpha 0}, \mu_{\beta 0}, n, p, J^{np})$. Для этого приравняем $\mu_{\alpha 0}^{(0)} = \mu_{\alpha 0}$ (погрешность не более 20 % [5]) и $J^{(0)} = J^{np}$, а функцию J^{np} , характеризующую "внутреннее трение" электронного и дырочного потоков, определим из подходящих экспериментальных результатов по измерению кинетических эффектов. Поскольку эти вопросы требуют более подробного рассмотрения, ограничимся лишь только кратким изложением сущности методики.

Наиболее достоверные данные в литературе можно найти по измерению суммы дрейфовых подвижностей $\mu_{n1} + \mu_{p1}$ в кремний в зависимости от плотности ЭД-плазмы $n \approx p \approx \bar{n}$ [9, 10]. Используя формулы (I4), можно записать следующее уравнение:

$$J^{np} = \frac{1}{\bar{n}(\mu_{n0} + \mu_{p0})} \left[\frac{\mu_{n0} + \mu_{p0}}{(\mu_{n1} + \mu_{p1}) \exp} - 1 \right]. \quad (I6)$$

Для определения J^{np} можно также использовать данные по измерению амбиполярной холловской подвижности в кремний [11].

В этом случае

$$J^{np} = \frac{1}{\bar{n}(\mu_{n0} + \mu_{p0})} \left[\left(\frac{\mu_{n\delta 0}}{\mu_{n\delta}} \right) \exp -1 \right]. \quad (I7)$$

Результаты расчетов показаны на рисунке. Данные работы [II], снятые при $T = 290$ К, приведены к $T = 300$ К с помощью теоретической зависимости $J^{(0)} \sim T^{-3/2}$. Пунктирная кривая на рисунке соответствует аппроксимирующей зависимости

$$J^{np} = 1,4 \cdot 10^{-20} \frac{1 + 3,8 \cdot 10^{-18} \bar{n}}{1 + 2,8 \cdot 10^{-17} \bar{n}}.$$

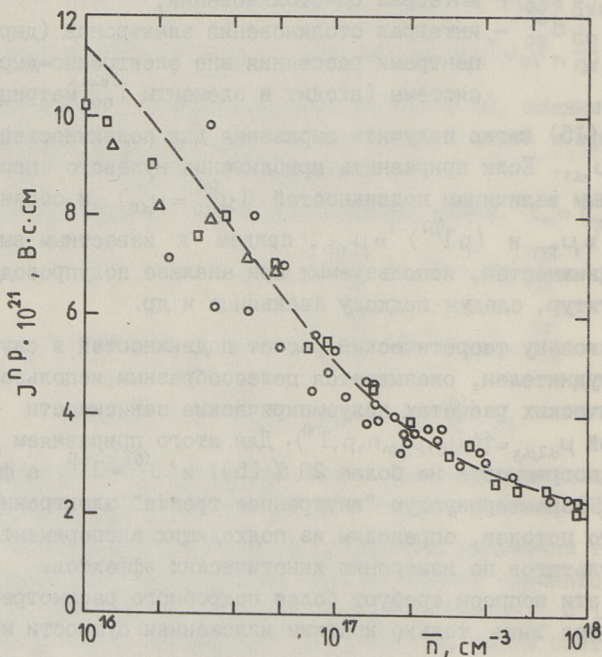


Рис. Зависимость параметра J^{np} от концентрации носителей заряда ($n \approx p = \bar{n}$) в кремнии при $T = 300$ К. Точки получены в результате расчета с помощью формул (16), (17), используя экспериментальные данные работ [9] (Δ), [10] (○) и [11] (□).

Как видно из рисунка, величины J^{np} , найденные из различных кинетических эффектов, хорошо согласуются между собой. При концентрациях $\bar{n} \lesssim 10^{16}$ см⁻³ влияние ЭДР ослабляется, величины $(\mu_{n0} + \mu_{p0}) / (\mu_{n1} + \mu_{p1}) \exp$ и $(\mu_{n\delta 0} / \mu_{n\delta}) \exp$ стремятся к единице и погрешность расчета по формулам (I6), (I7) сильно растет.

1. F l e t c h e r N.H. The high current limit for semiconductor junction devices // Proc. IRE. - 1957. - Vol. 45. N 6. - P. 862-872.
2. А в а к њ я н ц Г.М., М у р ы г и н В.И., С а н д л е р Л.С., Т е ш а б а е в А., Ю р о в с к и й А.В. Прямая ветвь вольт-амперной характеристики тонких диодов // Радиотехника и электрон.- 1963. - Т. 8.- № II. - С. 1919-1926.
3. В е л м р е Э.Э., У д а л А.Э. Численный расчет диода с точным учетом электронно-дырочного рассеяния // Силовые быстродействующие полупроводниковые приборы: Сб. статей. - Таллин: Валгус, 1984. - С. 39-43.
4. В е л м р е Э.Э., П и р о ж е н к о А.А., У д а л А.Э. Численное моделирование электрофизических процессов в полупроводниках с учетом электронно-дырочного рассеяния // Электрон. моделирование. - 1985. - Т. 7. - № 4. - С. 66-71.
5. M e y e r J.R. Effect of electron-hole scattering on ambipolar diffusion in semiconductors // Phys. Rev. - 1980. - Vol. 21, N 4. - P. 1554-1558.
6. K o h l e r H. Transporterscheinungen in Elektrogenas // Z. Phys. - 1949. - Bd. 125, N 11/12. - S. 679-693.
7. М н а ц а к а н о в Т.Т., Р о с т о в ц е в И.Л., Ф и л а т о в Н.И. О соотношении Эйнштейна в полупроводниках в условиях сильного электронно-дырочного рассеяния // Физ. и техн. полупр. - 1984. - Т. 18. - Вып. 7. - С. 1293-1296.
8. А в а к њ я н ц Г.М., Л а з а р е в Е.В. О влиянии электронно-дырочного рассеяния на вольт-амперную характеристику полупроводниковых приборов // Изв. АН Арм. ССР. - Физика. - 1969.- Т. 4. - С. 83-90.
9. D a n n h ä u s e r F. Die Abhängigkeit der Trägerbeweglichkeit in Silizium von der Konzentration der freien Ladungsträger I // Solid-State Electron. - 1972. - Vol. 15, N 12. - P. 1371-1375.
10. K r a u s s e J. Die Abhängigkeit der Trägerbeweglichkeit in Silizium von der Konzentration der freien Ladungsträger II // Solid-State Electron. - 1972. - Vol. 15, N 12. - P. 1377-1381.

E. Velmre

Some Physical Aspects of Taking into Account the Effect
of Electron-Hole Scattering in Mathematical Modelling
of Semiconductor Structures

Abstract

Starting from the generalized current and transport coefficient expressions obtained from the Boltzmann transport equation using Kohler's variational principle, the electron and hole current formalism, describing the electron-hole scattering (EHS) and the drag effect as well, have been discussed. A method for approximate calculation of the elements of matrix B in the current equation $j = B j_0$, where j_0 is the matrix-vector of electron and hole currents in case of negligible small EHS effect, has been also proposed.

E. Velmre

Füüsikalisi aspekte elektron-aukhajumise arvestamisel
pooljuhtstruktuuride matemaatilisel modelleerimisel

Kokkuvõte

Lähtudes üldistatud vooluvõrrandist ja vastavast ki-
neetiliseist tegurist, mis on saadud Boltzmanni kineetili-
se võrrandi lahendamisel Kohleri variatsioonprintsipi abil,
on esitatud ja põhjendatud elektron- ja aukvoolude võrran-
deid, mis võtavad arvesse laengukandjate nn. kaasahaaramise
nähtuse. Viimase põhjustab elektron-aukhajumise (EAH) meh-
hanism. Samuti on esitatud võrrandi $j = B j_0$ matriksi B
elementide ligikaudse määramise meetod. Matriksi-vektor j_0
esitab elektron- ja aukvoolusid pooljuhis, kui EAH mõju
võib mitte arvestada.

С о д е р ж а н и е

I.	А. Кеэваллик, М. Круус, К. Салум. О декомпозиции конечных автоматов на базе сдвиговых регистров.....	3
2.	А. Кеэваллик, Б. Беркман. Алгебра пар неполностью определенных разбиений и критерий информресурса при декомпозиции автоматов.....	12
3.	Б. Беркман. Декомпозиция микропрограммных автоматов в классе гомоморфизмов с использованием информресурса.....	21
4.	К. Таммемяэ. Выбор разбиений для итеративной декомпозиции конечных автоматов.....	35
5.	В. Аланго. Функциональный подход к созданию математических моделей дискретных объектов на обобщенных альтернативных графах.....	43
6.	Т. Коньт. Подход к автоматизации построения тестов для микропроцессорных БИС.....	55
7.	Р. Убар, Т. Лохуару, Б. Штраубе, Г. Элст. Построение тестов для проверки операционных частей дискретных систем.....	65
8.	В. Аланго, Т. Коньт, Р. Убар. Формирование тест-программ для микропроцессорных БИС при автоматическом генерировании тестов.	78
9.	Б. Бомбик, В. Вийес. Пакет графических программ для курса моделирования цифровых схем.....	88
10.	Ш. Галлаи, Т. Ранг. Компараторы с обратной связью.....	92
II.	А. Лаансоо, В. Мяннама, О. Пикков. Измерение параметров аналоговых микросборок применением компенсационного метода... ..	105
12.	А. Лаансоо, В. Мяннама, О. Пикков. Вопросы разработки коммутационно-согласующего устройства автоматизированного комплекса измерения параметров микросборок.....	113

13.	Б. Гурьянов. Проектирование электронных устройств для медицины.....	I20
14.	А. Опоцкий. Времена жизни неосновных носителей заряда в полупроводниках с учетом Оже-процессов на локализованных состояниях.....	I34
15.	Т. Ранг. Комплексный подход к системам проектирования элементов полупроводниковой и интегральной электроники.....	I49
16.	Э. Велмре. Физические аспекты учета влияния эффекта электронно-дырочного рассеяния при математическом моделировании полупроводниковых структур.....	I66

руб. 1.80

EESTI AKADEEMILINE RAAMATUKOGU



1 0200 00082437 9