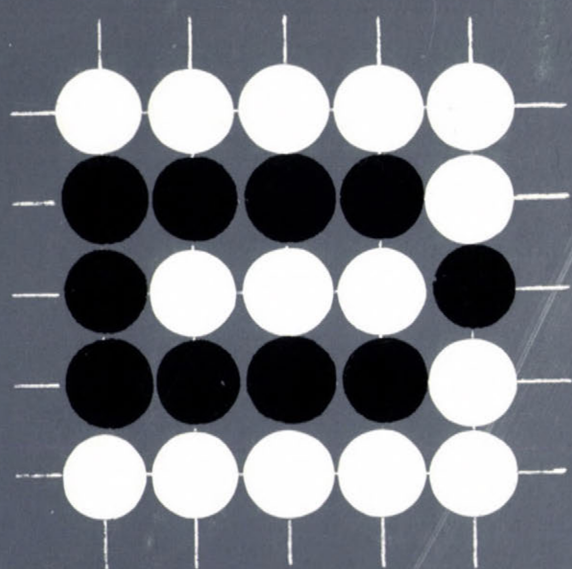
közlemények 21/1978

55807

1775

MTA Számítástechnikai és Automatizálási Kutató Intézet Budapest



MAGYAR TUDOMÁNYOS AKADÉMIA
SZÁMITÁSTECHNIKAI ÉS AUTOMATIZÁLÁSI KUTATÓ INTÉZETE

KÖZLEMÉNYEK

1978. MÁJUS

Szerkesztőbizottság:

GERTLER JÁNOS (felelős szerkesztő)
DEMETROVICS JÁNOS (titkár)
BACH IVÁN, GEHÉR ISTVÁN, GERGELY JÓZSEF, KERESZTÉLY SÁNDOR, KNUTH ELŐD, KRÁMLI ANDRÁS, PRÉKOPA ANDRÁS

Felelős kiadó:

DR VÁMOS TIBOR
igazgató

ISBN 963 311 066 1
ISSN 0133-7459

Technikai szerkesztő:
Solt Jánosné

TARTALOMJEGYZÉK

CONTENTS

Procedings of the Computer and Automation Institute
Hungarian Academy of Sciences
Vol. 21.

СОДЕРЖАНИЕ

Труды Исследовательского Института
Вычислительной Техники и Автоматизации
Венгерской Академии Наук
Выпуск 21.

MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 21/1978.

Анализатор для грамматики ван Вейнгаардена

Геревич Л.

1

1. Определение W-грамматики /грамматика ван Вейнгаардена/

Рассмотрим множества, образующие матаситемы W-грамматики:

M - конечное множество метапонятий;

τ_0 - конечное множество метатерминальных символов;

$\Pi \subset M \times (M \cup \tau_0)^*$ - конечное множество контекстно-свободных метаправил и

$\pounds_x = \{w \in \tau_0{}^* : x \underset{\Pi}{\Rightarrow} M\}$ - множество представляющее собой цепочки метатерминальных символов, порождаемых метапонятием x.

Рассотрим множества языка:

$\tau \supset \tau_0$ - конечное множество, содержащее и метатерминальные символы;

Σ - конечное множество терминальных символов языка;

$\bar{\Phi} \subset \{<\bar{w}> : \bar{w} \in (M \cup \tau)^+\}$ - конечное множество гиперпонятий;

$\bar{\Gamma} \subset \bar{\Phi} \times (\bar{\Phi} \cap \Sigma)^*$ - конечное множество гиперправил, служащих формами языка;

$\Phi \subset \{<w> : w \in \tau^+\}$ - множество нетерминалов языка, $M \cap \tau = \emptyset$;

$\Gamma \subset \Phi \times (\Phi \cup \Sigma)^*$- множество правил языка, полученное из $\bar{\Gamma}$ путем применения <u>совместимой замены</u>

$$\Psi : (M \cup \tau)^+ \rightarrow \tau^+$$

которая является однозначной функцией /полугруппой и гоморфизмом/ со следующими свойствами:

$$\Psi(x) \in \pounds_x \ (x \in M)$$
$$\Psi(w) = w \ (w \in \tau^+)$$

Из гиперправила $<\bar{w}> \rightarrow \sigma_1 <\bar{w}_1> \sigma_2 \ldots \sigma_{n-1} <\bar{w}_{n-1}> \sigma_n \in \Gamma \ (<\bar{w}, <w_i> \in \Phi, \delta_i \in \Sigma^*)$

A^+- множество всех цепочек над алфавитом A и $A^* = A^+ \cup \{\lambda\}$

получается контекстно-свободное правило:

$$\Psi(\bar{w}) \to \sigma_1\Psi(\bar{w}_1)\sigma_2\ldots\sigma_{n-1}\Psi(\bar{w}_{n-1})\ \sigma_n \in \Gamma$$

<S> является начальным символом или целью. Языком, порожденным с помощью W-грамматики, назовем следующее множество:

$$L = \{\sigma \in \Sigma^{*} : S \underset{\Gamma}{\Rightarrow} \sigma\}\ .$$

Пример 1. Рассмотрим язык $L_1 = \{x^k y^k x^k : k \geq 1\}$. W-грамматика W_1 , которая описывает язык, имеет следующий вид:

метаправила: $A \to nA,\ \ A \to \lambda$ (λ – пусто);

гиперправила: $<S>\to<aA><bA><aA>$; $<anA>\to x<aA>$; $<bnA>\to y<bA>$;
$<a>\to\lambda$; $<b>\to\lambda$,

то есть $W_1 = (M,\ \tau_0,\ \Pi,\ \Sigma,\ \bar{\Phi},\ \bar{\Gamma},\ S)$; $M = \{A\}$; $\tau_0 = \{n,\lambda\}$;

Π определено выше, $\pounds_A = \{n^{*}\}$; $\tau = \tau_0 \cup \{<S>\}$; $\Sigma = \{x,\ y\}$;

<S>- цель языка, $\bar{\Gamma}$ задано выше, $\bar{\Phi} = \{<S><aA><bA><anA><bnA><a><b>\}$.

Слова языка L порождаются с помощью грамматики W_1 с применением следующих шагов.

Пусть nn – терминальное порождение метапонятия A , тогда применяя это порождение к первому гиперправилу, получим правило языка следующего вида:

$$<S> \to <ann><bnn><ann>$$

Исходя из цели языка /из S/ и применяя полученное правило, получим строку вывода:

$$<ann><bnn><ann>$$

Для нетерминала **<ann>** можно получить правило из гиперправила <anA> → x<aA> заменой A на его терминальное порождение n , то есть можно получить правило <ann> → x<an> строгого языка. Применяя это правило к нетерминалу **<ann>** строки вывода в его обоих вхождениях, получим строку вывода:

x<an><bnn>x<an>

Исходя из того же гиперправила и заменяя A на λ , получим другое правило строгого языка:

<an> → x<a>

Применяя его к строке вывода, получаем:

xx<a><bnn>xx<a> .

В случае нетерминала <bnn> исходим из гиперправила <bnA>→y<bA>.

Аналогичным образом, заменяя сначала A на n , а затем A на λ , в результате последовательного применения этих двух правил получим строку вывода:

xx<a>yy<b>xx<a> .

Применяя последние два гиперправила - правила <a> → λ и <b> → λ –, наконец, получим:

xxyyxx .

2. Прежде чем перейти к описанию анализатора для W-грамматики, остановимся на рассмотрении таких грамматик. Попытаемся применять гиперправила таким же образом, как применялись контекстно-свободные правила, проследим, как изменяются значения метапонятий.

Определим совместимую замену по-другому и в дальнейшем всегда будем ссылаться на это определение. Будем говорить, что цепочка β <u>совместимо выведена</u> из цепочки £ , если цепочка β была порождена из £ таким образом, что все вхождения одного и того же метапонятия в $\mathcal{L}$ были заменены на одно и то же порождение этого метапонятия, не обязательно терминальное, т.е. возможно содержащее метапонятия, которые отличаются от метапонятия цепочки β. /Рассмотрим, например, α = **aAEA**, метаправило: **A** $\rightarrow$ **eE**. С помощью совместимого вывода получим цепочку β = $\mathbf{aeE_1EeE_1}$. Метапонятие $\mathbf{E_1}$ порождается как **E**, только его порождение не обязательно должно совпадать с порождением метапонятия **E** при совместимом выводе, то есть оно не подчиняется законам совместимости с метапонятием **E**./

Назовем гиперправило подходящим для некоторой цепочки δ, если из некоторого гиперпонятия α цепочки δ и из гиперпонятия, стоящего в левой части гиперправила, совместимо выводятся одинаковые гиперпонятия. Эти одинаковые гиперпонятия будем называть общими. /Пусть α = <aA> и δ = <aA><bA><aA>, метоправила для **A** : **A** $\rightarrow$ **nA**, а гиперправило <anA> $\rightarrow$ x<aA> является подходящим, поскольку <anA> совместимо выводимо из <aA> заменой **A** на **nA**./

Для порождения терминальной цепочки языка, который описывается **W**-грамматикой, будем исходить из цели грамматики. Применим к нему гиперправило, в результате получим цепочку, состоящую из гиперпонятий. Рассмотрим, как применяются гиперправила к цепочкам, состоящим из гиперпонятий. Сначала выберем некоторое гиперпонятие, а затем найдем гиперправило, подходящее к нему. После того, как найдено подходящее гиперправило, т.е. такое, что из выбранного гиперпонятия и из гиперпонятия в левой части гиперправила, подходящего для него, порождаются общие гиперпонятия, выберем одно из общих гиперпонятий, и из цепочки совместимо-порождаем такую цепочку, в которой из выбранного гиперпонятия исходной цепочки получилось общее гиперпонятие. Для этого выбранного гиперпонятия цепочки применим гиперправило, совместимо порожденное из исходного гиперправила таким образом, что левая часть есть то же самое общее гиперпонятие.

/Например, гиперправило имеет вид: <aAE> → <bA><dE> , а метаправила: A → eE и B → aeE . Рассмотрим цепочку <BE><gBC><bE>. Из гиперпонятия <BE> и из гиперпонятия <aAE> совместимо выводится гиперпонятие <aeE$_1$E> . Совместимо порожденные цепочка и гиперправило имеет вид:

<aeE$_1$E><gaeE$_1$c><aeE$_1$bE> u <aeE$_1$E> → <beE$_1$><dE>

Применим это новое гиперправило к новой цепочке, получим цепочку:

<beE$_1$><dE><gaeE$_1$C><bE> . /

Повторим эту процедуру. Будем применять к полученной цепочке гиперправила описанным выше образом до тех пор, пока не получим терминальную строку языка, описываемого с помощью W-грамматики.

При таком порождении может случиться, что получится такое гиперпонятие, для которого нет подходящего гиперправила. Такие случаи назовем тупиками.

Сущность такого порождения состоит в том, что мы пытаемся ссылаться в цепочках вывода на всевозможные выводы W-грамматики /вхождения метапонятия в цепочку вывода/. Поскольку общих гиперпонятий может быть больше, чем одно /их может быть бесконечно много/, а в таких выводах мы пользуемся только одним из них, поэтому мы можем проследить только за некоторым подмножеством множества всевозможных выводов /подмножество совпадает с множеством, если общее гиперпонятие всегда только одно/.

<u>Пример 2.</u> В качестве примера рассмотрим W-грамматику W_1 , которая задана в примере 1. Рассмотрим, как порождаются терминальные цепочки с помощью этой грамматики:

а/ Для цели языка применим гиперправило и получим следующую цепочку:

<aA><bA><aA> /1/

б/ Для того, чтобы для гиперпонятия <aA> можно было применить второе гиперправило, из цепочки /1/ нужно порождать цепочку:

<anA><bnA><anA>

После применения гиперправила получим цепочку:

x<aA><bnA><anA>

в/ Применим второе гиперправило еще раз. Подобным образом получим:

xx<aA><bnnA><annA>

г/ Четвертое гиперправило будет применимо, если из цепочки будет совместимо порождена цепочка:

xx<a><bnn><ann> .

Применяя четвертое гиперправило, получим:

xx<bnn><ann> .

д/ Для гиперпонятия <bnn> можем порождать следующее гиперправило из третьего гиперправила заменой A на n :

<bnn> → y<bn> .

С помощью этого гиперправила из этой цепочки получим следующую цепочку:

xxy<bn><ann> .

е/ Из третьего гиперправила порождаем гиперправило <bn>→y<b>. Применяя его к цепочке, получим:

xxyy<b><ann> .

ж/ Применяя пятое гиперправило, получим цепочку:

xxyy<ann> .

Аналогичным образом из гиперпонятия <ann> получим xx .

Перейдем к формальному описанию нового определения W-грамматики. Множества метасистемы W-грамматики:

$M_i = \{m_j^i\}$ – бесконечное множество, состоящее из одного метапонятия с разными индексами.

$M = \{M_i\}$ – конечное множество множеств метапонятий.

τ_0 – конечное множество метатерминальных символов.

$\Pi \subset M \times (M \cup \tau_0)^*$ – множество контекстно-свободных правил и

$\pounds_x = \{w \in M \cup \tau_0^* : x \underset{\Pi}{\overset{*}{\Rightarrow}} w\}$ – множество всех цепочек, выводимых из метапонятия x .

Множества языка и функция Ψ определяются как в предыдущем определении.

Говорим, что из цепочки α выводима цепочка β в смысле W-грамматики, если существуют α_1, α_2, ρ, δ, функции Ψ и Ψ^1 и гиперправило $\bar{w} \to \delta$, что $\alpha = \alpha_1 \rho \alpha_2$, $\Psi^1(\rho) \equiv \Psi(\bar{w})$ и $\beta = \Psi^1(\alpha_1) \Psi(\delta) \Psi^1(\alpha_2)$.

3. Основная идея анализатора W-грамматики заключается в том, что гиперправила рассматриваются как контекстно-свободные /КС/ правила, а гиперпонятия как "добавления" к ним. Для КС-правил – для метаправил и для гиперправил – используется метод рекурсивного спуска. В этом методе правила реализуются в виде булевских процедур, считывание входной строки происходит слева направо. Будем рассматривать только такие грамматики, в которых нет левой рекурсии.

Гиперправила реализуются в виде булевских процедур, а гиперпонятия рассматриваются как их параметры, являющиеся грамматическими объектами.

Поскольку для гиперпонятия существует несколько подходящих гиперправил, и заранее неизвестно, какие из гиперправил будут подходящими, так как зависит от конкретных значений метапонятий, поэтому в отличие от метода рекурсивного спуска для КС-грамматик здесь приходится вызывать все гиперправила /некоторое подмножество множества гиперправил/, чтобы найти подходящее. Для этой цели введем понятие вызова множества процедур, а затем перейдем к описанию грамматических объектов.

Под вызовом множества P булевских процедур $P = \{P_1, \ldots, P_n\}$ будем понимать последовательные вызовы элементов множества. Результат вызова будет истиной, если результат вызова некоторой процедуры P_i был истиной, и тогда последовательность вызова заканчивается. Результат вызова множества P будет ложным, если результат всех вызовов $P_1, \ldots, P_n$ был ложным. Обозначим этот вызов через P .

Введем понятия грамматической переменной и константы и операции над ними. Пусть G - контекстно-свободная грамматика ,
$G = (N, T, P, S)$,
N - конечное множество нетерминальных символов;
T - конечное множество терминальных символов;
P - конечный набор правила вида $A \to \gamma$, где $A \in N$ и $\gamma \in V^*$ $(V = N \cup T)$;
S - выделенный нетерминальный символ. Будем говорить, что β выводится из $\alpha (\alpha, \beta \in V^*)$ одним шагом, если $A \to \gamma$ есть правило в P , и существуют такие $\alpha_1, \alpha_2 \in V^*$, что $\alpha = \alpha_1 A \alpha_2$ и $\beta = \alpha_1 \gamma \alpha_2$. Тогда можно записать $\alpha \underset{G}{\to} \beta$. Выводимость является рефлексивным транзитивным замыканием выводимости с одним шагом. Обозначим выводимость через $\Rightarrow$.

Будем говорить, что A есть грамматическая переменная, если $A \in N$, а z - грамматическая константа, если $z \in T$. Грамматические переменные могут иметь значения, которые представляют собой строку, состоящую из грамматических переменных и констант. Строка $\gamma \in V^*$ является значением A в смысле G, если $A \Rightarrow \gamma$, при этом будем говорить, что γ имеет тип A. Таким образом, грамматикой определяется область значений грамматической переменной A. Задание грамматики G служит определением грамматических переменных и констант, и они являются начальными значениями и типами для себя. Под значением некоторой строки α будем понимать те строки β, которые выводим из нее, и строка α является типом β. Будем говорить, что два типа α и β подобны, если существует строка γ, типом которой является как α, так и β, и назовем ее общей.

Грамматическим переменным могут присваиваться значения только в том случае, если тип нового значения есть старое значение.

Между ними определена операция конкатентации, по значениям операндов которой строится гиперпонятие. Операцию конкатенации будем обозначать точкой /например, $A.B$/. Грамматическим выражением является такое выражение, в которое входят грамматические переменные и константы, соединенные операцией конкатенации.

Грамматические выражения могут стоять в качестве фактических параметров в обобщенном вызове P и в качестве формальных параметров в заголовках булевских процедур.

Параметры, которые являются грамматическими выражениями, передаются следующим образом. Сначала проверяется подобность фактических и формальных параметров. Если они не подобны, тогда производится возврат с результатом "ложь". В противном случае метапонятиям формального параметра присваиваются значения, на которые они были заменены при порождении выбранного общего гиперпонятия. В том случае, когда процедура заканчивается с результатом "ложь", надо выбрать другое общее гиперпонятие /только в том случае, когда их число конечно/ и присвоить другие значения метапонятиям формального параметра соответственно замене. Если нет больше гиперпонятий, тогда нужно восстановить исходные значения метапонятий формального параметра и произвести возврат с результатом "ложь". Если результат - "истина", тогда метапонятиям фактического параметра присваиваются значения, на которые они были заменены при порождении выбранного гиперпонятия при входе в данную процедуру, и управление передается вызывающей процедуре.

В определении передачи параметров, являющихся грамматическими выражениями, важное место занимает проверка подобности, или другими словами, нахождение общего гиперпонятия.

Вообще говоря, подобность двух типов - неразрешимая проблема, поскольку она является определением пустоты пересечения двух КС-языков /описанных КС-грамматикой/.

Будем рассматривать только такие **W**-грамматики, в которых пересечение языков, определенных гиперпонятиями, определено и является КС-языком.

Мы стремились выбрать для проверки подобности двух метапонятий простой, нетрудоемкий алгоритм, который тем не менее является достаточно мощным для того, чтобы он был применим для **W**-грамматики, описываемой языка АЛГОЛ-68.

Подобность двух гиперпонятий проверяется следующим образом. Сначала проверяется совпадение. Если хотя бы одна из оставшихся, несовпадающих частей начинается метапонятием, делается попытка породить из этого метапонятия часть другого. Если оба гиперпонятия начинаются метапонятием и ни одно из них не порождалось из другого, тогда делается попытка порождения такого общего гиперпонятия, которое выводимо из второго, в частности, делается подстановка вместо метапонятия его альтернативы. Если из некоторого метапонятия первого гиперпонятия выводима часть второго, но продолжить процесс дальше нельзя, тогда следует уменьшить эту часть второго гиперпонятия и продолжить проверку . Это исправление производится до тех пор, пока уменьшенная часть не станет пустой.

Рассмотрим этот метод проверки на примерах. Пусть гиперпонятия α и β имеют вид $\alpha = \langle aA \rangle$ и $\beta = \langle anA \rangle$, а метаправило – $A \to nA$. Первые символы совпадают, и из метапонятия A выводимо nA , то есть они подобны. Пусть $\alpha = \langle nnn \rangle$ и $\beta = \langle An \rangle$, а метаправило $A \to \alpha$ и $A \to nA$. Из A выводимо nnn , но тогда в β остается n . Уменьшим цепочку, выводимую из A , то есть возьмем nn в качестве порождения A . Последние символы совпадают. Подобность установлена. Пусть $\alpha = \langle Ab \rangle$ и $\beta = \langle aB \rangle$, а метаправила: $A \to aC$, $B \to Cb$. Из A не выводимо aB , а выводимо aC . Вместо $\alpha = \langle Ab \rangle$ рассмотрим в качестве α такое гиперпонятие, которое было получено из α заменой метапонятия A на его альтернативу /aC/ , то есть $\alpha = \langle aCb \rangle$. Первые символы совпадают /символ a/, и из B выводимо Cb .

Пример 3. Рассмотрим грамматику W_1 . Порождение цепочки $xxyyxx$ было рассмотрено в примере 2., рассмотрим эту же цепочку с точки зрения распознавания с помощью выше описанных идей.

Имеются 7 процедур, первые пять процедур соответствуют гиперправилам, а последние два метаправилам. Процесс распознавания начинается вызовом первой процедуры, соответствующей первым гиперправилам.

В первой процедуре формируется фактический параметр aA и обобщенно вызываются остальные процедуры, соответствующие гиперправилам. Вторая процедура является подходящей, если метапонятие A было заменено в фактическом параметре на nA.

С помощью процедур, соответствующих метаграмматике, проверяется, допустима ли замена A на nA. Во второй процедуре считывается символ x , формируется фактический параметр aA и обобщенно вызываются процедуры. Подобным образом проверяется, является ли подходящей вторая процедура. Заменим A на nA, затем считывается следующий символ x , формируется aA и вызываются процедуры. Опять устанавливается, что вторая процедура является подходящей, заменим A на nA . При попытке считывания следующего символа x в этой процедуре, оказывается, что следующий символ не x . Тогда восстанавливается старое значение A , и продолжаются вызовы процедур. Третья процедура не является подходящей, поскольку фактический параметр aA начинается метатерминалом, а первый символ формального параметра bnA есть b , то есть они не совпадают. Четвертая процедура подходит, если A заменить на "пусто". Применим эту замену и вернемся к вызывающей процедуре /второй/ с результатом "истина" и со значением λ метапонятия A . Из второй процедуры вернемся к вызывающей процедуре, из которой был сделан возврат с результатом "истина" два раза и со значением n метапонятия A . Наконец, вернемся к первой процедуре с результатом "истина" и со значением nn метапонятия A .

Сформулируем фактический параметр bnn и будем вызывать процедуры. Установим, что третья процедура является подходящей при замене A на n , то есть A присваивается значение n . Считывается символ y , формируется по грамматическому выражению bA фактический параметр bn и вызываются процедуры. Снова третья процедура является подходящей, метапонятию A присваивается значение λ . Считывается символ y , и фактический параметр будет b . Вызываются процедуры. Пятая процедура является подходящей. Вызывается эта процедура и производится возврат из нее с результатом "ложь".

Аналогичным образом считывается символ xx для нетерминала ann W-грамматики. Получим, что строка xxyyxx порождается грамматикой W_1.

4. Описание метода. Наш метод требует два стэка. Первый стэк небходим для рекурсивных вызовов с параметрами, то есть для описания грамматических выражений /для ссылок на элементы грамматического выражения/. Он предназначен также для хранения информации, связанной с рекурсивным вызовом, такой, как указатель стэка и другие указатели. Стэк используется также для хранения набора формальных параметров вызванной процедуры, то есть для гиперпонятия левой части гиперправила.

Второй стэк содержит таблицу ссылок на метапонятия гиперправила, одно и то же метапонятия гиперправила ссылаются на некоторый элемент таблицы, соответствующей ему. Стэк используется также для хранения значений метапонятий, присваиваемых им в процессе анализа. Значения метапонятий организованы в виде деревьев их порождений, сами метапонятия являются корнями деревьев.

Перед вызовом булевских процедур, соответствующих гиперправилам, необходимо занести элементы грамматического выражения в первый стэк. Если элемент является метапонятием, тогда в первый стэк помещается ссылка на соответствующий ему элемент таблицы.

Вызов булевских процедур, соответствующих гиперправилам, означает, что необходимо вызвать булевскую процедуру, соответствующую некоторому гиперправилу, и если результат вызова окажется ложью, тогда вызвать следующую процедуру. Эти вызовы продолжаются до тех пор, пока результат вызова не будет истиной или не возникает такая ситуация, что все гиперправила проверены и все результаты были "ложь", вследствие чего необходимо вернуться с результатом ложь к вызывающей процедуре.

После вызова, то есть организации рекурсии /в частности сохранения указателей/, необходимо сформировать во втором стэке таблицу ссылок /например, занести адрес процедуры, соответствующей метапонятию/ и поместить элементы формальных параметров вызванной процедуры в первый стек.

Последним шагом входа в процедуру является нахождение общего гиперпонятия δ, которое может быть выведено из грамматического выражения α и из формальных парметров β /в дальнейшем просто δ, α и β /. Для этого надо вызвать булевскую процедуру с двумя параметрами α и β, реализующую следующий алгоритм /ANAL/.

Исходные параметры алгоритма: α и β. Алгоритм состоит из следующих шагов: Шаг 1. Пропуск одинаковых частей. При $\alpha = \gamma\alpha^1$, $\beta = \gamma\beta^1$, $\gamma, \alpha^1, \beta^1 \in V^*$ в качестве α и β следует взять α^1 и β^1. Если гиперпонятие δ содержит и метапонятие, то в элемент таблицы ссылок, соответствующий метапонятию из α, нужно записать ссылку на элемент таблицы ссылок, соответствующий метапонятию из β, и поскольку они должны иметь одинаковые значения, запомнить место ссылок в первом стэке. После этого следует продолжить работу алгоритма следующим образом.

Шаг 2. Вывод первой части β /или α/ из первого символа-метапонятия α /или β/ при $\alpha = A\alpha^1$ и $\beta = \gamma\beta^1$ /или $\alpha = \gamma\alpha^1$ и $\beta = A\beta^1$ /, где $\gamma, \alpha^1, \beta^1 \in V^*$, $A \in N$, $A \underset{G}{\Rightarrow} \gamma$. Построить дерево γ, то есть перенести его элементы во второй стэк. Вместо элемента γ записать ссылку на эту информацию. Запомнить эту замену в первом стэке при $\alpha = A\alpha^1$. Взять α^1 и β^1 в качестве α и β и продолжить работу алгоритма, переходя на шаг 1.

Шаг 3. Если $\alpha = A\alpha^1$ и $\beta = B\beta^1$, где $A, B \subset N, \alpha^1, \beta^1 \in V^*$, тогда поставить вместо первого символа-метапонятия A его альтернативу. Альтернативы подставляются только вместо первого символа-метапонятия гиперпонятия α, являющегося грамматическим выражением. Записать дерево этой замены во второй стэк, ссылку на него занести в элемент таблицы, соответствующий метапонятию **A**, и запомнить это соответствие в первом стэке.

При неудовлетворительном результате, когда замена не помогает, необходимо удалить из таблицы ссылку на дерево и перейти к шагу 4.

Шаг 4. Если α исчерпано /то есть все его символы пересмотрены/, а β нет, и $\beta = A\beta^1$, $A \in N$ /или наоборот/, или никакой шаг не применим, и последним был применен шаг 2, тогда сначала надо взять в качестве α гиперпонятие α_1 / $\alpha = \alpha_1\alpha_2$, где α_1, $\alpha_2 \in V^*$ / и попытаться вывести из β, и только потом выводить α_2. Этот процесс заканчивается, когда α_2 переходит в α. Перейти к шагу 5.

/В АЛГОЛе-68 такой ситуацией является, например, вызов гиперправила 1.3.3.a (**NOTION option**), на самом деле метатерминал **option** выводим из метапонятия **NOTION**, поэтому при первой попытке из **NOTION** выводится все гиперпонятие α. Правильный вывод получается лишь тогда, когда берется $\alpha = \alpha_1 \alpha_2$, где α_2 = **option**, α_1 выводимо из **NOTION** и α_2 совпадает с **option**.

Примером иного рода является вызов гиперправила 4.8.a, где набор формальных параметров: **QUALITY NEST new PROPSETY1 QUALITY TAX PROPSETY2 definig INDICATOR with TAX**, где **QUALITY TAX** есть **PROP**, выводимое из **PROPSETY**/.

Шаг 5. Конечный шаг алгоритма. Если α и β исчерпаны одновременно, то производится останов с результатом "истина". В противном случае, когда с помощью шагов 3,4 уже проверены все возможные продолжения, необходимо аннулировать занесенные ссылки на новые деревья и произвести останов с результатом "ложь".

В шаге 2 описанного выше метода для проверки выводимости можно использовать любой метод синтаксического анализа, только выбранный метод должен быть применен для анализа строк, содержащих и метопонятия.

Скелет анализатора для **W**-грамматики является обычным анализатором для контекстно-свободных грамматик.

Попытаемся написать с использованием описанных выше идей анализатор на языке, подобном АЛГОЛ, для **W**-грамматики, заданной в примере 1. Исходным методом является метод рекурсивного спуска.

Будем использовать для обозначения элементов списков /деревья организуются в виде списков/ следующие типы: **t1** - элемент, который является метатерминалом, **t2** - ссылка на элемент типа **t1**, **t3** - ссылка на метапонятие в таблице ссылок, **t4** - ссылка на некоторый элемент таблицы ччылок, **t5** - ссылка на некоторый элемент таблицы ссылок, в который была занесена ссылка на значение, **t0** - конец списка.

Пусть z [l : u] будет входной строкой.

```
BEGIN INT ARRAY L [1:M] ; INT K; K:=1; INT H; H:=1 ;
¢  определение второго стека и его указателей ¢ ;
¢H - указатель входной строки ¢ ;
¢ - метаправила ¢ ;
¢T - такая функция, которая применяется к элементу списков и
     результатом которой является метапонятие, если оно не
     имеет только начальное значение или является метатерми-
     налом ¢ ;
¢ NEXT - такая функция, которая продвигает указатель списка ¢;

BOOL PROC A ; IF NOT T=n THEN GO TO R2; NEXT
IF NOT T=REF(A) THEN GO TO R2; NEXT; A:=TRUE; R2:A:=TRUE
END
```

```
¢ гиперправило для цели языка ¢ ;
INT ARRAY S [ 1:15 ] ;
¢ определение места для гиперпонятий - фактических парамет-
  ров, т.е. определение первого стека ¢ ;
INT I ; I:=1 ; ¢ определение указателя начала гиперпоня-
                 тия ¢ ;
INT J ; J:=K ; ¢ указатель на начало таблицы ссылок ¢ ;
¢ формирование таблиц ссылок ¢ ;
L [K] :=t3 ; L [K+1]:= REF(A); L [K+2] := Ø ; K: = K+3 ;
S [I] :=t1 ; S [I+1]:='a' ; ¢ элемент метатерминала ¢ ;
S [I+2] :=t4 ; S [v+3] :=J ; ¢ элемент метапонятия ¢ ;
S [I+4] :=t0 ; ¢ конец фактического параметра ¢ ;
IF NOT HRULE (J,S,I) THEN GO TO ERROR; ¢ вызов гиперправи-
ла ¢;
¢ формирование второго гиперпонятия правой части гиперпра-
  вила в первом стэке ¢ ;
S [I] : = t3; S [I+1] : ='b' ; S [I+1] : = t4; S [I+3] : = J ;
S [I+4] : = tØ ;
IF NOT HRULE (J,S,I) THEN GO TO ERROR;
```

Аналогичным образом формируются гиперправила третьего гиперпонятия правой части.

```
¢ гиперправила ¢ ;
BOOL PROC H1 (OJ,OS,OI); REF TO INT ARRAY OS; INT OJ ;
REF TO INT OI ; INT ARRAY S [1:14] ; ¢ первый стэк ¢ ;
INT J; J:=K; INT R; ¢ для начала таблицы изменений ¢ ;
INT I ;I:=1 ;
L [K] : = t3 ; L [K+1] : =REF(A); L [K+2] : =Ø ; K : = K+3 ;
```

```
S [I] := t1 ; S [I+1] := 'a'; S [I+2] := t1; S [I+3] := 'n' ;
S [I+4] := t4 ; S [I+5] := J; S [I+6] := tØ ; R: = I+7 ;
IF NOT ANAL (OJ,OS,OI,J,S,I,R)  THEN GO TO FALSE
IF NOT Z[H] = 'x' THEN GO TO FALSE ;  H: = H+1
S [I] := t1; S [I+1] :='a' ; S [I+2] := t4; S [I+3] := J ;
S |I+4| : tØ ;
IF NOT HRULE (J,S,I)  THEN GO TO FALSE;  H1: = TRUE ;
FALSE: WHILE  S[R] = t5 DO BEGIN L[S[R+1]] : = Ø ; R:=R+2 END
HI: = FALSE END
```

Аналогично формируются и остальные гиперправила.

ЛИТЕРАТУРА

1. Van Wijngaarden, A., Maillonx, B.J., Peck, J.E.L., Koster, C. H. A., Sintzoff, M., Lindsey, C. H., Meertens, L. G. L. T. and Fisker. R. G. Revised Report on the Algorithmic Language ALGOL-68. Supplement to ALGOL BULLETIN № 36, March 1974, Edmonton, Alberta.

2. Маслов, А.Н.: Лекции по алгоритмическому языку АЛГОЛ-68.

3. Baker, J.L. Grammars with structured vocabulary: A model for the ALGOL-68 definition. Infomation and Control, Volume 20, /1972/, 351-395.

4. Greibach, S.A. Some restrictions on W-grammars. ACM proceedings of the 6th symposium on the theory of computing, Seattle, May, 1974.

5. Deussen, P.A. Decidability criterion for van Wijngaarden grammars . Acta Informatica 5., /1975/, 353-374.

6. Цейтин, Г.С. и другие. АЛГОЛ-68. Методы реализации, изд-во ЛГУ, 1976 г.

7. Братчиков, И.Л. Синтаксис языков программирования. Изд-во Наука, 1975.

8. Маслов, А.Н. Индексные грамматики и грамматика Вейнгаардена. Проблемы передачи информации. Том., Xl., Вып. 3, 1975, 81-89.

S u m m a r y

Syntax-parser of W-grammars

L. Gerevich

A new definition is given for W-grammar. The method of recursive descent is made applicable for W-grammars through the introduction of a new, grammar-type variable.

Ö s s z e f o g l a l ó

Szintaktikai analizátor kétszintü nyelvtanokra

Gerevich László

A dolgozat uj meghatározását adja a Wijngaarden-féle kétszintü nyelvtannak. Egy uj tipusu, nyelvtan-tipusu változó bevezetésével teszi alkalmassá a rekurziv leereszkedés módszerét kétszintü nyelvtanokra.

MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 21/1978.

RECOGNITION OF THE CHANGE OF A STOCHASTIC PROCESS ON THE BASIS OF THE PREDICTION ERROR

T. Várady

ABSTRACT

Two statistical tests are introduced by means of which one can decide if an unknown stochastic process has changed or not. The first one is based on the examination of the linear least square prediction error, the second one on the evaluation of the sum of consecutive prediction errors. In connection with this it is examined how the prediction error "reacts" to the change of the covariance function and what is the limit distribution of the above mentioned sum. For the investigation of the discussed and similar problems a computer program system was developed which generates and analyses stochastic processes.

INTRODUCTION

The aim of this paper is to study how to recognize the change of a stocahstic process having unknown statistical parameters, more precisely, how to set up mathematical criteria on the basis of which one can consider a process changed or unchanged.

The above mentioned alarming problem is generally examined in cases where the set of possible stochastic processes and probability characteristics describing them are known. In this case recognition of change means that we suppose the process to have one of the possible distributions and we examine when a stepping over to another class takes place. In our case the situation is somewhat different: we know nothing about the possible distribution of the process either before or after the change and still have to alarm when the change occurs.

Recognition of change of stochastic processes is used in several areas; quality control, process control, controlling and checking of various industrial systems, description of communication and computer technique, several biological and medical applications etc. A typical example of the latter is the intensive guarding room, where on the basis of controlling the ECG of serious cardiac patients forecasting of fatal arrhytmia can be given.

Our method for recognition will be as follows. Let the process be $\{ \xi_1, \xi_2, \dots \xi_n, \dots \}$. In the interval $[1, n]$ (in the so called first phase) we suppose the process to be stationary. We estimate here some statistical parameters (connected with linear prediction error) which we shall need in the interval $[n+1,\infty)$ (in the so called second phase). Our null hypothesis is that the coefficients $\{c_i\}$ minimizing the square prediction error in the first phase give the process correctly in the second phase. Actually we base our test not on the whole interval $[n+1,\infty)$ i.e. on the sequence $\xi_{n+1}, \xi_{n+2} \dots$ but only on subsequences $\xi_{n+k+1}, \xi_{n+k+2}, \dots \xi_{n+k+m}$ where m fixed and k runs through the values $k = 1, 2, \dots$. We construct two statistics $t = t(\xi^{(n)}, \xi'^{(m)})$ where

$$\xi^{(n)} = \{ \xi_1, \xi_2, \dots \xi_n \}$$

$$\xi'^{(m)} = \{ \xi_1', \xi_2', \dots \xi_m' \} = \{ \xi_{n+k+1}, \xi_{n+k+2}, \dots \xi_{n+k+m} \}$$

and base our test on them.

LINEAR PREDICTION AND EXPECTED VALUE OF ITS LEAST SQUARE ERROR

Linear prediction approximates the random variable ξ_{k+1} by a linear combination of the previous s random variables with error δ_{k+1} in the following form: (s is a positive integer, $k \geqq s$).

$$\xi_{k+1} = \sum_{i=1}^{s} c_i \xi_{k-s+i} + \delta_{k+1}$$

The minimization of $E[|\delta_{k+1}|]$ is rather difficult, on the contrary the expected value of the square error, $E[\delta^2_{k+1}]$, can be minimized in a very simple way. Assuming that the process is stationary we can drop the index κ. The extremum problem

$$E[\delta^2] = E[(\sum_{i=1}^{s} \xi_i c_i - \xi_{s+1})^2] = \min.$$

leads to a system of s linear equations. The coefficients $\{c_i\}$ which minimize the square prediction error are obtained by solving this system. The solutions is

$$c_{(s)} = Z_s^{-1} \cdot m_{(s)}$$

assuming that Z_s^{-1} exists. Here we used the following notations:

$$R(i) = E\,[\xi_k\ \xi_{k+i}] \qquad i = 0, 1, 2, \dots$$

$$Z_s = \{Z_{ij}\}\ \begin{matrix} i=1,s \\ j=1,s \end{matrix} = \{R\,(|i-j|)\}\ \begin{matrix} i = 1, s \\ j = 1, s \end{matrix}$$

$$c_{(s)} = \{c_i\} \quad i=1,s$$

$$m_{(s)} = \{m_i\} \quad i=1,s \quad = \{R\ (s+1-i)\}\ i = 1,s$$

(We notice that we determine the coefficients $\{c_i\}$ not by matrix inversion, but by Robbins-Monroe stochastic approximation.)
Finally we formulate a known theorem.

Theorem 1. /See Appendix I./
The expected value of the square prediction error is

$$E\,[\delta^2] = E\,[\xi_{s+1}^2] - (m_{(s)}, c_{(s)}) = \frac{\det Z_{s+1}}{\det Z_s}$$

EXPECTED VALUE OF THE PREDICTION ERROR AT THE CHANGED PROCESS

The condition of applicability of the prediction tests is to show that the expected value of the prediction error changes if the covariance function of the process has changed. In the first phase the original process was characterized by random variables $\xi_1, \xi_2, \dots \xi_s, \dots$, in the second phase by $\xi_1', \xi_2', \dots \xi_s', \dots$. The prediction error for the changed process is the following (using the notation $c_{s+1} = -1$):

$$E\,[\delta'^2] = E[(\sum_{i=1}^{s+1} c_i\xi_i')^2] = R'(o)\sum_{i=1}^{s+1} c_i^2 + 2\ R'(1) \sum_{i=1}^{s} c_i c_{i+1} \dots$$

$$\dots 2\ R'(j) \sum_{i=1}^{s+1-j} c_i c_{i+j} + \dots + 2\ R'(s-1)\ [c_1 c_s + c_2 c_{s+1}] +$$

$$2\ R'(s)\ c_1 c_{s+1} = \sum_{k=0}^{s} R'(\kappa)\ C_k\ .$$

Denoting the covariance function of the first phase by $R(k)$, let the covariances of the second phase be $R'(k) = R(k) + \Delta_k$ $(k = 0, 1, 2, ..)$
In this case

$$E[\delta'^2] = E[\delta^2] + \sum_{k=o}^{s} \Delta_k C_k = E[\delta^2] + \Delta$$

Obviously, if the observed (ξ) and (ξ') processes differ only by one $R(k)$ value, the change of the prediction error depends on the value of C_k. If $R(o)$ changes, then the prediction error increases since $C_o > 0$. In general, however, we can not tell how the changes of Δ_o, Δ_1, ... Δ_s compensate each other and how the prediction error behaves.

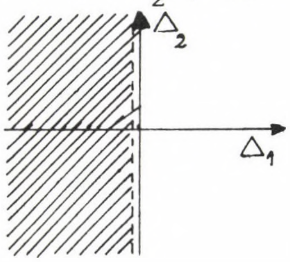
Let us consider thecase $s = 2$.
Suppose that $E[\xi_i^2] = E[\xi_i'^2] = 1$. In this case

$$\Delta = 2[\Delta_1 (c_1 c_2 - c_2) + \Delta_2 (-c_1)] \qquad (c_3 = -1)$$

$$c_1 = \frac{R(2) - R^2(1)}{1 - R^2(1)} \qquad c_2 = \frac{R(1)(1 - R(2))}{1 - R^2(1)}$$

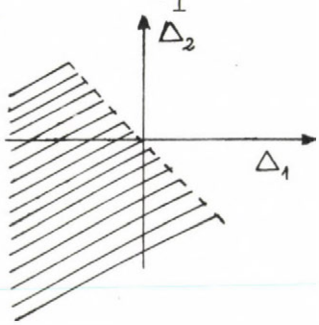
Examine the expression Δ :

I. If $c_1 = c_2 = 0$ then $\Delta = 0$.
II. $c_1 = 0$ and $c_2 \neq 0$.



If $\Delta_1 = 0$ then $\Delta = 0$.
If sign Δ_1 = $-$sign c_2 then $\Delta > 0$,
otherwise $\Delta < 0$.
$\Delta > 0$
$\Delta < 0$
$\Delta = 0$

III. $c_1 \neq 0$ and $c_2 \neq 0$.
Let $\beta = \dfrac{c_2 c_1 - c_2}{c_1}$



If $\Delta_1 \beta = \Delta_2$ then $\Delta = 0$.
If $c_1 > 0$ and $\Delta_1 \beta > \Delta_2$ or
$c_1 < 0$ and $\Delta_1 \beta < \Delta_2$ then $\Delta > 0$,
otherwise $\Delta < 0$.
$\Delta > 0$
$\Delta < 0$
$\Delta = 0$

Consequently, the plane (Δ_1, Δ_2) can be divided into three regions; in one halfplane Δ is positive, in the other one it is negative and on the boundary line Δ equals to zero, i.e. if the changes of the covariances are Δ_1, Δ_2 and $\Delta_1 \beta = \Delta_2$ then the expected value of the prediction error is unchanged. In our example the change of the process implies the change of the prediction error with probability one.

RECOGNITION OF CHANGE I.

Let us return to our problem concerning the change of processes. In the first phase, supposing that n is large enough, a good estimate can be obtained for the coefficients $\{c_i\}$, as well as for $E[\xi]$ and $E[\xi^2]$. Let our statistic be

$$t \quad (\xi^{(n)}, \xi^{\prime(m)}) = \delta = \sum_{i=1}^{m-1} c_i \; \xi_i' - \xi_m' \qquad (s = m-1)$$

The distribution of the random variable δ is unknown but the expected value and the variance of δ can be determined provided that the null hypotheses is true.

$$E[\delta] = E[\xi]\left(\sum_{i=1}^{m-1} c_i - 1\right)$$

$$D^2[\delta] = E\,[\delta^2] - (E\,[\,\delta])^2$$

$E\,[\delta^2]$ can be calculated with the aid of Theorem I. Supposing that the expected value and the variance are known, by means of the Chebisev inequality we can obtain a critical region for given level of significance ε .

$$P(|\delta - \varepsilon[\delta]| > \lambda_\varepsilon D[\delta]) < \frac{1}{\lambda_\varepsilon^{\;2}} = \varepsilon$$

$$X_{cr} \simeq (\delta < E[\delta] - \lambda_\varepsilon \, D[\delta] \text{ or } \quad \delta > E[\delta] + \lambda_\varepsilon \, D[\delta])$$

If the value δ, calculated on the basis of the actual subsequence of the second phase falls into the critical region we consider the process changed. (If the exact distribution of δ were known we would get a greater critical region, i.e. our test would be more powerful.) In the next section, using sums of prediction errors instead of individual errors, we give a better test depending on a limit distribution theorem.

RECOGNITION OF CHANGE II.

We first formulate a central limit theorem for stationary mixing sequences. The stationary process X_t is said to be uniformly strongly mixing if

$$\phi(\tau) = \sup_{A\varepsilon\, \sigma_{-\infty}^{t^*},\ B\varepsilon\, \sigma_{t^*+\tau}^{\infty}} \frac{P(AB) - P(A)\,P(B)}{P(A)}$$

and

$$\lim_{\tau\to\infty} \phi(\tau) = 0$$

where $\sigma_{-\infty}^{t^*}$ and $\sigma_{t^*+\tau}^{\infty}$ are the σ-fields generated by the process X_t in the intervals $(-\infty, t^*]$ and $[t^*+\tau, \infty)$, respectively. $\phi(\tau)$ is called the mixing coefficient.

<u>Theorem 2.</u> [2] (18.5.2.)
Let X_t be a uniformly strongly mixing stationary sequence with $E[X_t] = 0$ such that

$$\sum_n (\phi(n))^{1/2} < \infty$$

Let
$$\sigma_n^2 = E\,[\,(\sum_{j=1}^{n} X_j\,)^2\,]$$

moreover
$$\sigma^2 = \lim_{n\to\infty} \frac{1}{n}\,\sigma_n^2 < \infty \qquad \text{and} \quad \sigma \neq 0.$$

Then
$$\lim_{n\to\infty} (p\,(\frac{1}{\sigma\sqrt{n}} \sum_{j=1}^{n} X_j < z)) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{u^2}{2}}\,du$$

In the following we shall examine processes for which $\sum_{k=0}^{\infty} |R(k)| < \infty$ In order to apply Theorem 2. let the size of the subsequence of the second phase (m) be much greater then the number of prediction coefficients (s). We shall use $E[\xi] = 0$. (If this is not satisfied we replace the random variable ξ by $\zeta = \xi - E[\xi]$.) Let us suppose that the sequence ξ_k satisfies the mixing condition. (As it is known, from this it follows that $\sum_{k=0}^{\infty} |R(k)| < \infty$.). The statistic $t(\xi^{(n)}, \xi'^{(m)})$ is defined by

$$t\ \xi^{(n)}, \xi'^{(m)} = \delta_m^* = \sum_{k=0}^{m-s-1} \delta_k .$$

where

$$\delta_k = \sum_{i=1}^{s} \xi'_{i+k} \; c_i - \xi'_{s+k+1}$$

Since $E[\xi_k] = 0$, $E[\delta_k]$ equals zero. The sequence δ_k generates a σ-field not greater than those generated by ξ_k, thus the mixing condition is valid for the sequence δ_k.

Finally for applying Theorem 2. we have to show that $\lim_{m\to\infty} \frac{1}{m} D^2 [\delta_m^*]$ exists. This follows from Theorem 3.

Theorem 3. /See Appendix II./

If

$$\sum_{k=0}^{\infty} |R(k)| < \infty \qquad \text{and} \qquad E[\delta_k] = 0$$

then

$$\lim_{m\to\infty} \frac{1}{m} D^2 [\delta_m] = \lim_{m\to\infty} \frac{1}{m} D^2 [\sum_{k=0}^{m-s-1} \delta_k] =$$

$$= (\sum_{j=1}^{s+1} c_j)^2 (R(0) + 2 \sum_{k=1}^{\infty} R(k))$$

$$(c_{s+1} = -1)$$

Consequently, we construct our test for recognition of change in the following way. Given ε, we detemine the values c_{o1}, c_{o2} - using the fact that δ_m^* is approximately normally distributed - such that

$$P\,(\delta_m^* < c_{o1} \quad \text{or} \quad \delta_m^* > c_{o2} \mid H_o) = \varepsilon$$

If the value δ_m^* calculated on the actually observed subsequence does not fall into the region of acceptance $[c_{o1}, c_{o2}]$ we decide that the process has changed.

COMPUTER IMPLEMENTATION

For the investigation of the above mentioned and similar statistical tests a computer program system was developed on the configuration TPA 70 - GD'71 in the Computer and Automation Institute of the Hungarian Academy of Sciences by means of which, after generating various deterministic and stochastic processes, one can follow the changes of the defined processes [3]. All of this is carried out in a simple and flexible way using the possibilities given by the interactive technique and graphical display. This is illustrated by Fig. 1-3. (Figure 1 shows how the Robbins-Monroe adaptive algorithm converges in case of a deterministic process. In Figure 2 and 3 we can see how the process generated by prediction (dashed * line) approximates the original process before and after change in case of Gaussian processes.)

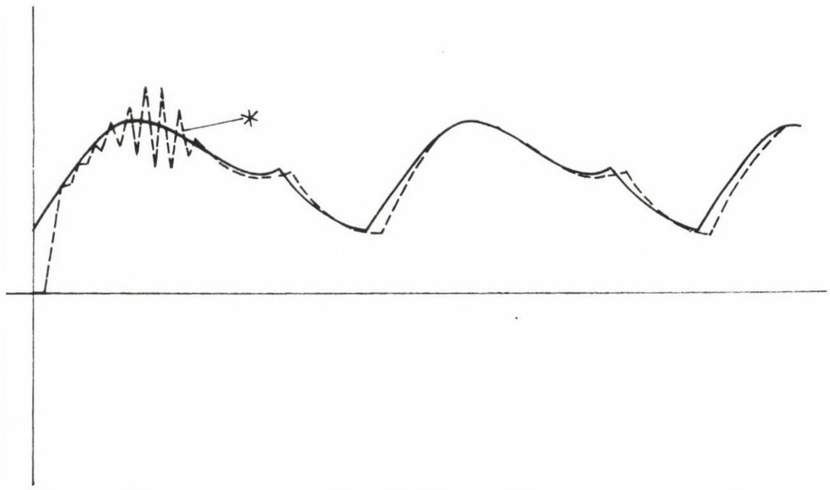


Fig.1 DETERMINISTIC PROCESS AND PREDICITON

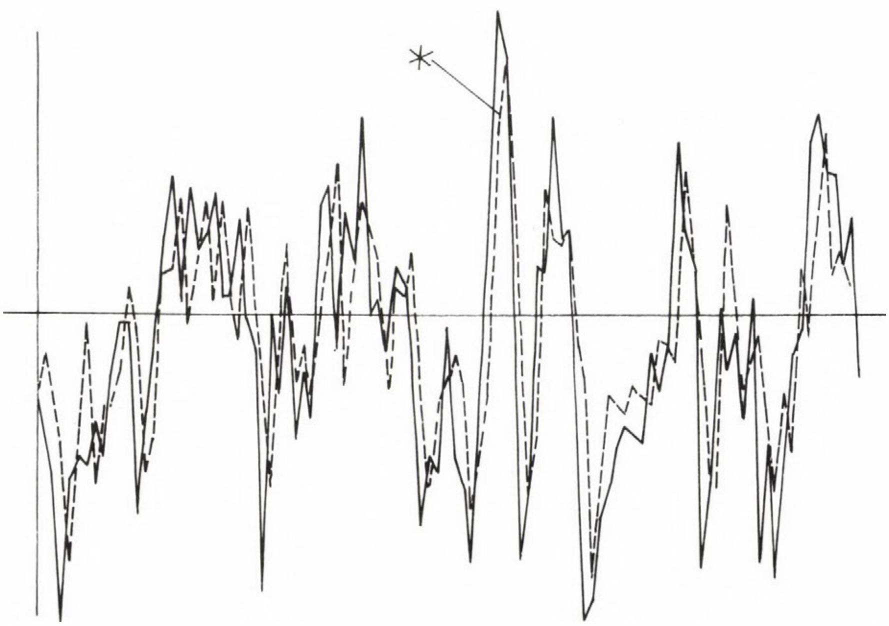


Fig. 2 GAUSSIAN PROCESS AND PREDICTION

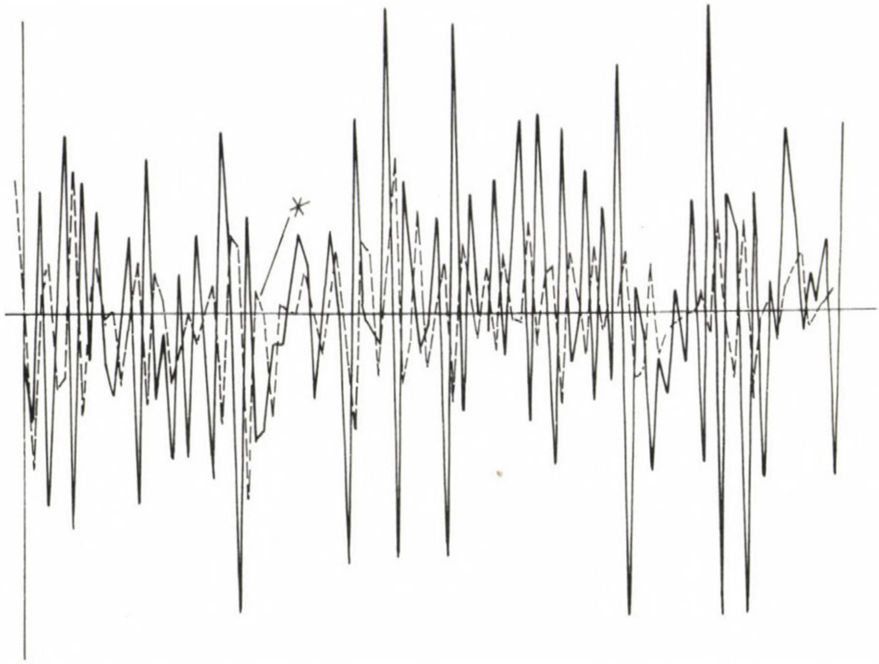


Fig. 3 CHANGED GAUSSIAN PROCESS AND PREDICTION

APPENDIX I.

First we give an algebraic proof of Theorem 1. We shall use the same notations as before, adding the following ones, the matrix $Z_{(s)}(i)$ is obtained by replacing the i-th column of $Z_{(s)}$ by the vector $m_{(s)}$, the matrix $D_{ij}^{(s)}$ by omiting the i-th row and the j-th column of $Z_{(s)}$.

PROOF

The expected value of prediction error in case of arbitrary vector $c'_{(s)}$ is

$$E[\delta^2] = E[((\xi_{(s)}, c'_{(s)}) - \xi_{s+1})^2] =$$

$$= (c'_{(s)}, Z_{(s)}c'_{(s)}) - 2\,(c'_{(s)}, m_{(s)}) + E[\xi^2_{s+1}].$$

Replacing $c'_{(s)}$ by the vector $c_{(s)} = Z^{-1}_{(s)}\, m_{(s)}$, which minimizes the prediction error, we get the left side of the equation in Theorem 1.

$$E[\delta^2] = E[\xi^2_{s+1}] - (m_{(s)}, Z^{-1}_{(s)}\, m_{(s)}) =$$

$$= R(0) - (m_{(s)}, c_{(s)})$$

By solving the previously discussed system of linear equations using Cramer's rule, where

$$c_i = \frac{\det Z_{(s)}(i)}{\det Z_{(s)}} ,$$

we obtain

$$E[\delta^2] = R(0) - \sum_{i=1}^{s} m_i \frac{\det Z_{(s)}(i)}{\det Z_{(s)}} =$$

$$= \frac{1}{\det Z_{(s)}} [R(0) \det Z_{(s)} - \sum_{i=1}^{s} m_i \sum_{j=1}^{s} (-1)^{i+j} R(s+1-j) \det D^{(s)}_{ij}] =$$

$$= \frac{1}{\det Z_{(s)}} [R(0) \det Z_{(s)} + \sum_{i=1}^{s} \sum_{j=1}^{s} (-1)^{i+j+1} R(s+1-i) R(s+1-j) \det D^{(s)}_{ij}]$$

The determinant of $Z_{(s)}(i)$ was developed by its i-th column.

We have to prove also that the expression in the square brackets equals $\det Z_{(s+1)}$. Developing the determinant of $Z_{(s+1)}$ by its first row and the determinant of matrices $D^{(s+1)}_{1,k+1}$ by their first column we obtain

$$\det Z_{(s+1)} = R(0) \det Z_{(s)} + \sum_{k=1}^{s} (-1)^k R(k) \det D^{(s+1)}_{1,k+1} =$$

$$= R(0) \det Z_{(s)} + \sum_{k=1}^{s} \sum_{l=1}^{s} (-1)^{k+l+1} R(k) R(l) D^{(s)}_{k,l}$$

The matrix $Z_{(s)}$ is symmetrical to both diagonal, i.e.

$$D^{(s)}_{ij} = D^{(s)}_{ji}$$

and

$$D^{(s)}_{ij} = D^{(s)}_{s+1-j,s+1-i} .$$

Consequently, after replacing the indices

$$i = s+1-k$$
$$j = s+1-l$$

the expression in the square brackets can be obtained thus the right side of the equation in Theorem 1 is also proved.

The geometric interpretation of Theorem 1 is the following. Let $\{\varepsilon_i\}_{i=1,s}$ be a basis of the s-dimensional linear space, then the random variables $\{\xi_i\}_{i=1,s}$ can be written as follows:

$$\xi_i = \sum_j a_{ij}\,\varepsilon_j$$

The volume of the parallelepiped spanned by the vectors $\{\xi i\}_{i=1,s}$ equals det $A_{(s)}$, where

$$A_{(s)} = \{a_{ij}\}_{\substack{i=1,s\\ j=1,s}} \quad .$$

It is well-known, that random variable ξ_{s+1} can be obtained as the sum of two orthogonal vectors; the first one is the projection of ξ_{s+1} to the s-dimensional space - this is the best linear prediction of ξ_{s+1}, while the second one is the prediction error - its absolute value is δ. The formula "volume equals area multiplied by height" leads to the following:

$$\det A_{(s+1)} = \delta \det A_{(s)} \ .$$

Since

$$\xi_i \xi_j = \sum_\kappa a_{i\kappa}\,\varepsilon_\kappa \sum_l a_{jl}\,\varepsilon_l = \sum_\kappa a_{i\kappa} a_{j\kappa},$$

we obtain

$$E[\det A_{(s)} \det A^*_{(s)}] = E\,[\det A_{(s)} A^*_{(s)}] =$$
$$= \det Z_{(s)} \ ,$$

thus

$$E[\delta^2] = \frac{\det Z_{(s+1)}}{\det Z_{(s)}}$$

which was to be proved.

APPENDIX II.

First we formulate a lemma, by means of which we shall prove Theorem 3.

Lemma

If $$\sum_{k=1}^{\infty} |a_k| < \infty$$

then

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} a_k(m-k) = \sum_{k=1}^{\infty} a_k$$

Proof

For all m, it is true, that

$$\left|\frac{1}{m} \sum_{k=1}^{m} a_k(m-k) - \frac{1}{m} \sum_{k=1}^{m} m a_k\right| = \left|\frac{1}{m} \sum_{k=1}^{m} a_k\, k\right| \leq \frac{1}{m} \sum_{k=1}^{m} |a_k| k$$

We have to show that

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} |a_k| k = 0$$

Let

$$f(x) = \begin{cases} 0; & \text{if } x \leq 0 \\ |a_k|; & \text{if } k-1 < x \leq k \quad k = 1,2,\dots \end{cases}$$

$$f_m(x) = \begin{cases} 0; & \text{if } x \leq 0 \text{ or } x > m \\ \frac{1}{m} |a_k| k & \text{if } k-1 < x \leq k \quad k = 1,2,\dots m \end{cases}$$

then $\int_{-\infty}^{\infty} f(x)dx < \infty$ and for all x $|f_m(x)| \leq |f(x)|$.

Using Lebesgue's theorem

$$\lim_{m\to\infty} \int_{-\infty}^{\infty} f_m(x)dx = \int_{-\infty}^{\infty} [\lim_{m\to\infty} f_m(x)\, dx] = 0$$

so

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=1}^{m} |a_k|\, k = 0$$

- . - . -

<u>Proof of Theorem 3.</u>

As it was discussed previously

$$\delta_m^* = \left(\sum_{k=0}^{m-s-1} \cdot \sum_{i=1}^{s+1} c_i \, \xi'_{i+k} \right)$$

(using the notation $c_{s+1} = -1$).

The following figure shows the structure of coefficients $\{c_i\}$ connecting to the random variables $\{\xi_i\}$.

Let k equal m-2s.

$$\begin{array}{cccccccccc}
 & & & & \multicolumn{3}{c}{\overbrace{\hphantom{\xi'_{s+1}\quad\xi'_{s+2}\qquad\quad\xi'_{m-s}}}^{k}} & & & \\
\xi'_1 & \xi'_2 & & \xi'_s & \xi'_{s+1} & \xi'_{s+2} & \xi'_{m-s} & \xi'_{m-s+1} & \xi'_{m-1} & \xi'_m \\
 & & & & & & & & & \\
c_1 & c_2 & c_3 & c_s & c_{s+1} & & & & & \\
 & c_1 & c_2 & & c_s & c_{s+1} & & & & \\
 & & c_1 & & & & & & & \\
 & & & & & & & & & \\
 & & & c_1 & & & c_{s+1} & & & \\
 & & & & c_1 & & c_s & c_{s+1} & & \\
 & & & & & & & & & \\
 & & & & & & c_1 & c_2 & c_{s+1} & \\
 & & & & & & & c_1 & c_s & c_{s+1}
\end{array}$$

We use the following notations:

$$\gamma_i = \sum_{j=1}^{i} c_j \qquad 1 \leq i \leq s$$

$$\gamma = \sum_{j=1}^{s+1} c_j$$

$$\gamma_i^* = \sum_{j=i}^{s} c_{j+1} \qquad 1 \leq i \leq s$$

Then

$$E[\delta_m^{*2}] = E[(\sum_{i=1}^{s}\gamma_i\xi_i + \sum_{i=1}^{k}\gamma\xi_{s+i}\sum_{i=1}^{s}\gamma_i^{*}\xi_{s+k+i})^2] =$$

$$= \overset{1}{E[(\sum_{i=1}^{s}\gamma_i\xi_i)^2]} + \overset{2}{E[(\sum_{i=1}^{k}\gamma\xi_{s+i})^2]} + \overset{3}{E[(\sum_{i=1}^{s}\gamma_i^{*}\xi_{s+k+i})^2]} +$$

$$+ \overset{4}{2E[(\sum_{i=1}^{s}\gamma_i\xi_i)(\sum_{i=1}^{m-2s}\gamma\xi_{s+i})]} + \overset{5}{2\,E[(\sum_{i=1}^{s}\gamma_i\xi_i)(\sum_{i=1}^{s}\gamma_i^{*}\xi_{s+k+i})} +$$

$$+ \overset{6}{2\,E[(\sum_{i=1}^{k}\gamma\xi_{s+i})(\sum_{i=1}^{s}\gamma_i^{*}\xi_{s+k+i})]}$$

We want to determine $\lim_{m\to\infty}\frac{1}{m}E[\delta_m^{*2}]$.

Since the terms *1*, *3*, *5* don't depend on m, dividing them by m, they converge to zero, if $m\to\infty$. The term *2*, dividing by m is the following:

$$2 \quad \frac{\gamma^2}{m}E[(\sum_{i=1}^{m-2s}\xi_{s+i})^2] = \frac{\gamma^2}{m}\sum_{j=1}^{m-2s}\sum_{i=1}^{m-2s} R(s+i-(s+j))=$$

$$= \frac{\gamma^2}{m}\sum_{j=1}^{m-2s}\sum_{i=1}^{m-2s} R(i-j) = \frac{\gamma^2}{m}[R(o)(m-2s) + 2\,R(1)(m-2s-1) +$$

$$+ 2\,R(m-2s-1)].$$

By means of our lemma $(a_1 \sim R(o),\quad a_2 \sim 2R(1),\quad a_k \sim R(k-1)\ldots)$.

$$\gamma^2 \lim_{m\to\infty}\frac{1}{m}[R(o)\ (m-2s) + \ldots 2R(m-2s-1)] =$$

$$\gamma^2\,[R(o) + 2\sum_{k=1}^{\infty} R(k)]$$

The term *4*, dividing by m converges to zero, if $m\to\infty$, since we can give an upper bound in the following way:

$$4 \le \gamma \max\gamma_i \quad E[\sum_{i=1}^{s}\xi_i \sum_{j=1}^{m-2s}\xi_{s+j} \le$$

$$\le \gamma \max\gamma_i\; s \sum_{j=1}^{m-s-1} R(j) \le \gamma \max\gamma_i\; s \sum_{j=1}^{\infty} R(j)$$

We obtain similar result to *6*.

This completes the proof.

REFERENCES

1 Gihmann, I.I. - Skorohod, A.V.: Introduction to the theory of stochastic processes /Russian/

2 Ibragimov, I.A. - Linnik, J.V.: Independent and stationary dependent random variables /Russian/ Edition "Nauka", Moscow, 1965

3 Várady, T.: Statistical methods and computer graphics for the recognition of the change of stochastic processes Dissertation, Technical University /BME HEI/, Budapest, 1977

Р Е З Ю М Е

Опознание изменения случайных процессов на основе ошибки линейного прогнозирования

Тамаш Варади

В статье излагаются две статистических пробы, с помощью которых можно решить изменился ли неизвестный случайный процесс или нет. Первая проба основывается на исследовании ошибки прогнозирования, во второй пробе оценивается сумма последовательных ошибок. В статье исследовано, как ошибка прогнозирования изменяется при изменении корреляционной функции, дальше, что будет предельным распределением вышеупомянутой суммы. Для исследования рассмотренных и подобных этому проблем интерактивная система была разработана на малую ЭВМ, при помощи которой можно генерировать и анализировать различные случайные процессы.

S u m m a r y

Statisztikus folyamatok állapotváltozásának felismerése a predikciós hiba alapján

Várady Tamás

Két statisztikai próbát ismertetünk, amelyek segitségével eldönthető, hogy egy ismeretlen sztochasztikus folyamat megváltozott avagy nem. Az első próba a lineáris predikció hibájának vizsgálatán alapszik, a második próbával az egymást követő predikciós hibák összegét értékeljük ki. Ennek kapcsán megvizsgáljuk, hogyan változik a predikciós a kovariancia függvény megváltozásakor, továbbá, hogy mi a fent emlitett összeg határeloszlása. A tárgyalt és hasonló jellegü problémák vizsgálatára egy kisszámitógépes interaktiv programrendszert fejlesztettünk ki, melynek segitségével különböző sztochasztikus folyamatok generálhatók és analizálhatók.

MTA Számítástechnikai és Automatizálási Kutató Intézete, Közlemények 21/1978.

О ПРИМЕНЕНИИ НЕЯВНОЙ РАЗНОСТНОЙ СХЕМЫ ДЛЯ РЕШЕНИЯ ОДНОМЕРНОГО ДИФФЕРЕНЦИАЛЬНОГО УРАВНЕНИЯ ПАРАБОЛИЧЕСКОГО ТИПА СО СЛАБОЙ НЕЛИНЕЙНОСТЬЮ

Д. Молнарка, Р.Х. Фарзан, Л. Фаи

Использование неявной разностной схемы для приближенного численного решения линейного дифференциального уравнения параболического типа, по сравнению с явной схемой, ослабляет условия, накладываемые на параметры схемы, при которых доказана устойчивость схемы.

В настоящей работе показаны достаточные условия для сходимости в равномерной метрике неявной разностной схемы для первой краевой задачи, для одномерного параболического уравнения со слабой нелинейностью и устойчивости схемы по начальным данным и нелинейной части. При этом условия, накладываемые на функцию f достаточно слабые. В частности, не требуется ее монотонности или непрерывности ее производных. Показано, что и для нелинейного случая при использовании неявной схемы ограничения, накладываемые на параметры схемы, значительно слабее, и в частности, нет соотношений, связывающих между собой шаги по времени τ и по пространству h, как это имело место для схемы [1].

Рассматривается следующая краевая задача:

$$a(x,t)\frac{\partial^2 u}{\partial x^2} + b(x,t)\frac{\partial u}{\partial x} + c(x,t)u - \frac{\partial u}{\partial t} = -f(x,t,u),$$

$$(x,t)\in\Omega \cdot \Omega\ \{x,t : x\in(0,1), t\in(0,T)\}\ ; \tag{1}$$

$$u(x,0) = g(x), \quad x\in[0,1]; \qquad \begin{array}{l} u(0,t) = g_1(t), \\ u(1,t) = g_2(t),\ t\in[0,T]. \end{array} \tag{2}$$

Здесь a $(x,t) > 0$, b (x,t), c $(x,t) \in C(\bar{\Omega})$, $g(x) \in C([0,1])$, $g_1(t)$, $g_2(t) \in C([0,T])$, причем в угловых точках границы граничные условия согласованы:

$g(0) = g_1(0)$, $g(1) = g_2(0)$.

В отношении $f(x,t,u)$, кроме непрерывности по всем аргументам, предполагаем, что она ограничена при всех u и удовлетворяет условию Липшица по t и u, т.е. существуют постоянные L_t, L_u, независимые от x, t, u такие, что

$$|f(x,t_1,u_1) - f(x,t_2,u_2)| \leqslant L_t |t_1 - t_2| + L_u |u_1 - u_2|, \quad (x,t)\in\overline{\Omega}, \quad -\infty < u < \infty. \tag{3}$$

Такого типа краевые задачи встречаются при моделировании химических процессов.

Будем предполагать, что решение задачи (1), (2) существует и единственно.

На области $\overline{\Omega}$ введем сетку $\overline{\Omega}_h\{x_i, t^n : x_i = ih,\ i = 0,1,\dots M,\ h = 1/M;\ t^n = n\tau,\ n = 0,1,\dots N,\ \tau = T/N\}$. Используем обычные в теории разностных схем обозначения [1], [2]. Разностную задачу, аппроксимирующую дифференциальную задачу (1), (2) запишем для $y_h = \{y_i^n,\ i = 0,1,\dots M,\ n = 0,1,\dots N\}$. Обозначим:

$$\underline{y}_n = (y_i^n,\ i = 0,1,\dots M)^T,\quad \underline{f}_n(\underline{y}_n) = (f(x_i,t^n,y_i^n),\quad i = 0,1,\dots M)^T$$

Введем линейный оператор D_n:

$$\{D_n\underline{y}_n\}_i = \left(\frac{\tau}{h^2}a_i^n - \frac{\tau}{2h}b_i^n\right)y_{i-1}^n + \left(-\frac{2\tau}{h^2}a_i^n + \tau c_i^n - 1\right)y_i^n + \left(\frac{\tau}{h^2}a_i^n + \frac{\tau}{2h}b_i^n\right)y_{i+1}^n,\quad i = 1,\dots M-1,\quad n = 1,\dots N. \tag{4}$$

С его помощью дифференциальный оператор из (1) аппроксимируется с точностью $O(h^2 + \tau)$ [1]. Будем рассматривать параллельно две разностные схемы, аппросимирующие уравнение (1), у которых много общего в теоретическом плане, хотя с точки зрения алгоритма численного расчета они различны. Первая:

$$D_n\underline{y}_n = -\underline{y}_{n-1} - \tau\underline{f}_{n-1}(\underline{y}_{n-1}),\quad n = 1,\dots N, \tag{5}$$

вторая:

$$D_n \underline{y}_n = -\underline{y}_{n-1} - \tau \underline{f}_n(\underline{y}_n), \qquad n = 1,\dots N. \tag{6}$$

Начальные и граничные условия для y_h одинаково аппроксимируют условия (2) :

$$y_i^0 = g(x_i), \quad i = 0,1,\dots M, \quad \begin{matrix} y_0^n = g_1(t^n), \\ y_m^n = g_2(t^n), \end{matrix} \quad n = 1,\dots N. \tag{7}$$

На основании (7) можно продолжить D_n и на граничные точки:

$$\{D_n \underline{y}_n\}_{i=0} = y_0^n, \ \{D_n \underline{y}_n\}_{i=M} = y_M^n.$$

Тогда вместо (7) можно записать:

$$\{D_n \underline{y}_n\}_0 = g_1(t^n), \quad \{D_n \underline{y}_n\}_M = g_2(t^n). \tag{8}$$

Таким образом, оператор D_n определен при всех i.

Различие между двумя схемами существенное: первая схема (5) , (7) является линейной системой уравнений относительно y_i^n , в то время ка вторая (6) , (7) - нелинейной. Систему уравнений (5) , (7) можно решать методом прогонки на каждом слое. Для устойчивости метода, из которой следует и сходимость, достаточно выполнения следующих условий [2] :

$$\begin{gathered} \frac{\tau}{h^2} a_i^n - \frac{\tau}{2h} b_i^n > 0, \\ \frac{\tau}{h^2} a_i^n + \frac{\tau}{2h} b_i^n > 0, \\ -\left(-\frac{2\tau}{h^2} a_i^n + \tau c_i^n - 1\right) \geqslant \left(\frac{\tau}{h^2} a_i^n - \frac{\tau}{2h} b_i^n\right) + \left(\frac{\tau}{h^2} a_i^n + \frac{\tau}{2h} b_i^n\right). \end{gathered} \tag{9}$$

Первые два из условий (9) ставят ограничение сверху /вообще говоря легко выполнимое/ на величину шага h:

$$h < \min_{i,n} \frac{2a_i^n}{|b_i^n|} . \tag{10}$$

Из третьего условия получаем ограничение для τ:

$$\tau \leqslant \frac{1}{\max\limits_{i,n} |c_i^n|}, \qquad (11)$$

причем, это условие также легко выполняется. Условия (10) , (11) являются, таким образом, и условиями существования и единственности решения системы уравнений (5) , (7). Отсюда же следует существование обратного оператора D_n^{-1}

В случае нелинейной системы (6), (7), ее можно решать методом простой итерации /с использованием прогонки на каждом шаге итерации/:

$$D_n \overset{(s+1)}{\underline{y}_n} = -\underline{y}_{n-1} - \tau \underline{f}_n(\overset{(s)}{\underline{y}_n}), \qquad s = 0,1,\ldots \qquad (12)$$

где s - номер итерации. В качестве начального приближения $\overset{(0)}{\underline{y}_n}$ можно использовать значение на предыдущем слое $\underline{y}_{n-1}$. Очевидно, предел $\underline{y}_n$ последовательности $\overset{(S)}{\underline{y}_n}$, $S = 0,1,\ldots$, если он существует, является решением уравнения (12).

Для сходимости и устойчивости метода прогонки в (12) достаточно выполнения условий (9). Рассмотрим условия, при которых сходится итерационный процесс. Известно [3], что для сходимости итерационного процесса

$$\overset{(S+1)}{\underline{x}} = g(\overset{(S)}{\underline{x}})$$

достаточно существования $q < 1$ такого, что

$$\rho(\underline{g}(x_1), \underline{g}(x_2)) \leqslant q \cdot \rho(\underline{x}_1, \underline{x}_2) \qquad (13)$$

Пусть

$$\rho(\underline{y}_n^*, \underline{y}_n^{**}) = \|\underline{y}_n^* - \underline{y}_n^{**}\|_n,$$

где "норма на слое":

$$\|\underline{y}_n\|_n = \max_i |y_i^n|.$$

Перепишем уравнение (12) .

$$\overset{(s+1)}{\underline{y}_n} = D_n^{-1}\,(-\underline{y}_{n-1} - \tau \underline{f}_n(\overset{(s)}{\underline{y}_n})).$$

Тогда, используя (3) ,

$$\|D_n^{-1}\,(-\underline{y}_{n-1} - \tau \underline{f}_n(\underline{y}_n^{*})) - D_n^{-1}(-\underline{y}_{n-1} - \tau \underline{f}_n(\underline{y}_n^{**}))\|_n \leqslant$$
$$\leqslant \tau \cdot L_u \cdot \|D_n^{-1}\|_n \cdot \|\underline{y}_n^{*} - \underline{y}_n^{**}\|_n, \qquad (14)$$

где норма оператора согласована с нормой вектора на слое.

Поскольку в граничных точках /при $i=0$, $i=M$/, согласно (7), значения $\overset{(S)}{\underline{y}_n}$ определяются точно, т.е. $\overset{(S)}{y}{}_o^n = g_1(t^n)$, $\overset{(S)}{y}{}_M^n = g_2(t^n)$, для определения нормы оператора D_n^{-1} достаточно рассмотреть уравнение:

$$D_n\,\underline{V} = \underline{W} \qquad (15)$$

где $\underline{V}$, $\underline{W} \in R^{M+1}$, $W_o = W_M = 0$, откуда

$$V_o = V_M = 0. \qquad (16)$$

Будем рассматривать уравнение (15), вместе с (16), как систему уравнений относительно V_i при заданных W_i.

Пусть в точке x_{io} V_i достигает максимума /в силу (16), он будет неотрицательным/. Если таких точек несколько, то пусть i_o наименьшее. Тогда возможны два случая:

|i| $i_o > 0$,

|ii| $i_o = 0$.

Из |ii| следует, что $V_i \leq 0$, $i = 0,1,\ldots M$. Для случая |i| $V_{io} > 0$. Перепишем уравнение (15) в точке x_{io} в следующем виде:

$$\left(\frac{\tau}{h^2} a_{i_0}^n - \frac{\tau}{2h}\, b_{i_0}^n\right)(V_{i_0} - V_{i_0}) + \left(\frac{\tau}{h^2} a_{i_0}^n + \frac{\tau}{2h}\, b_{i_0}^n\right)(V_{i_0+1} - V_{i_0}) -$$
$$- (1 - \tau c_{i_0}^n)V_{i_0} = W_{i_0} \qquad (17)$$

Первые два члена, вследствии (9), неположительны. Поэтому,

$$-(1-\tau c^n_{i_0})\, V_{i_0} \geqslant W_{i_0},$$

откуда, вследствие (11) ,

$$0 \leqslant (1-\tau \max_{i,n} |C^n_i|)\cdot V_{i_0} \leqslant -W_{i_0} \leqslant \max_i |W_i|. \tag{18}$$

Пусть теперь в точке x_{i1} V_i достигает /неположительного/ минимума. Тогда возможны два случая:

(а) $i_1 > 0$,

(б) $i_1 = 0$.

Из (б) следует, что $V_i \geq 0$, $i = 0,1,\ldots M$. Для случая (а) $V_{i1} < 0$. Тем же приемом , как и выше, получим, что в этом случае

$$0 \leqslant (1-\tau \max_{i,n} |c^n_i|)\cdot |V_{i_1}| \leqslant W_{i_1} \leqslant \max_i |W_i|. \tag{19}$$

Пусть мы имеем случай |i|, (а). Из неравенств (18), (19) следует, что

$$\max_i |V_i| \leqslant \frac{\max_i |W_i|}{1-\tau \max_{i,n} |c^n_i|}. \tag{20}$$

В случае |i|, (б) из (18), а в случае |ii|, (а) из (19) получаем такую же оценку (20) . В случае |ii|, (б) $V_i \equiv 0$.

Обозначим

$$c = \max_{(x,t)\in\overline{\Omega}} |c(x,t)|. \tag{21}$$

Тогда, используя определение нормы на слое, из (20) получаем:

$$\|\underline{V}\|_n \leqslant \frac{\|\underline{W}\|_n}{1-\tau\cdot C},$$

причем, здесь вместо (11) потребуем, чтобы $\tau < 1/c$. Если переписать (15) в виде:

$$\underline{V} = D_n^{-1}\,\underline{W},$$

то из предыдущей оценки следует, что

$$\|D_n^{-1}\|_n \leqslant \frac{1}{1-\tau c}\,. \tag{22}$$

Вернемся к соотношению (14). Для удовлетворения условия (13) необходимо выполнение неравенства

$$\tau \leqslant \frac{q}{L_u + cq} \qquad \text{где} \quad q < 1 \tag{23}$$

При этих условиях итерационный процесс (12) сходится, т.е. при выполнении (10), (23) решение сеточной задачи (6), (7) существует и единственно.

Перейдем к вопросу о сходимости решения y_h сеточной задачи к решению $u(x,\ t)$ дифференциальной задачи. Введем погрешность разностной схемы

$$z_h = \{\, z_i^n,\ i = 0,1,\ldots M,\ n = 0,1,\ldots N \,\},\ z_i^n = y_i^n - u_i^n,$$

где $u_i^n = u\,(x_i, t^n)$ - значения функции u в узлах сетки. Введем аналогично вектору y_n вектора $\underline{z}_n$, $\underline{u}_n$.

Теорема 1. При выполнении условий (3), (9), а также неравенства

$$\tau \leqslant \frac{1}{2(c + L_u)} \tag{24}$$

имеет место соотношение в равномерной метрике на $\bar{\Omega}_h$:

$$\|z_h\| \leqslant 0(h^2 + \tau), \tag{25}$$

т.е. решение сеточной задачи сходится к решению дифференциальной задачи.

Доказательство. Запишем разностные уравнения для погрешности схемы $\underline{z}_n$ для схем (5), (6) :

$$D_n \underline{z}_n = - \underline{z}_{n-1} - \tau(\underline{f}_{n-1}(\underline{y}_{n-1} \qquad - \underline{f}_n(\underline{u}_n)) + \tau \cdot 0(h^2 + \tau), \tag{26}$$

$$D_n \underline{z}_n = - \underline{z}_{n-1} - \tau(\underline{f}_n(\underline{y}_n) - \underline{f}_n(\underline{u}_n)) + \tau 0(h^2 + \tau). \tag{27}$$

Разрешим уравнение (26) относительно z_n:

$$\underline{z}_n = D_n^{-1} [- \underline{z}_{n-1} - \tau(\underline{f}_{n-1}(\underline{y}_{n-1}) - \underline{f}_{n-1}(\underline{u}_{n-1})) - \tau(\underline{f}_{n-1}(\underline{u}_{n-1}) - \underline{f}_n(\underline{u}_n)) + \tau 0(h^2 + \tau)].$$

откуда

$$\|\underline{z}_n\|_n \leqslant \|D_n^{-1}\|_n [\|\underline{z}_{n-1}\|_{n-1} + \tau \max_i |f(x_i, t^{n-1}, y_i^{n-1},) - f(x, t^{n-1}, u_i^{n-1})| + $$
$$+ \tau \max_i |f(x_i t^{n-1}, u_i^{n-1}) - f(x_i, t^n, u_i^n)|] + \tau 0(h^2 + \tau) \cdot \|D_n^{-1}\|. \tag{28}$$

В силу (24) и (22), $\| D_n^{-1} \|_n \leq 0(1)$. Далее, оценивая члены справа в (28) с помощью (3) и пользуясь свойством непрерывности u'_t, получим:

$$\|\underline{z}_n\|_n \leqslant \frac{1 + \tau L_u}{1 - \tau c} \|\underline{z}_{n-1}\|_{n-1} + \tau 0(h^2 + \tau). \tag{29}$$

Поскольку это соотношение верно для всех $n \leq N$, продолжая (29) в цепочку неравенств до $n = 1$, получим:

$$\|\underline{z}_n\|_n \leqslant \left(\frac{1 + \tau L_u}{1 - \tau c}\right)^n \cdot \|\underline{z}_0\|_0 + \tau \cdot 0(h^2 + \tau) \sum_{j=0}^{n-1} \left(\frac{1 + \tau L_u}{1 - \tau c}\right)^j. \tag{30}$$

Аналогично можно получить из уравнения (27) :

$$\|z_n\|_n \leqslant \frac{1}{1 - \tau(c + L_u)} \|z_{n-1}\|_{n-1} + \tau \cdot 0(h^2 + \tau)$$

Продолжая эту цепочку неравенств, получим:

$$\|\underline{z}_n\|_n \leqslant \left(\frac{1}{1 - \tau c - \tau L_u}\right)^n \cdot \|\underline{z}_0\|_0 + \tau \cdot 0(h^2 + \tau). \sum_{j=0}^{n-1} \left(\frac{1}{1 - \tau(c + L_u)}\right)^j \tag{31}$$

Для оценки сумм в (30), (31) используем соотношения

$$1 - \tau L_u \leqslant e^{\tau L_u},\ 1 - \tau c \geqslant e^{-2\tau c},\ 1 - \tau(c + L_u) \geqslant e^{-2\tau(c + L_u)} \tag{32}$$

которые выполняются при выполнении (24). Тогда

$$\left(\frac{1 + \tau L_u}{1 - \tau c}\right)^n \leqslant e^{\tau n L_u} e^{2\tau c n} \leqslant e^{\tau N(L_u + 2c)} = e^{T(L_u + 2c)},$$

$$\left(\frac{1}{1 - \tau(c + L_u)}\right)^n \leqslant e^{2n\tau(c+L_u)} \leqslant e^{2\tau N(c+L_u)} = e^{2T(c+L_u)} .$$

Подставляем эти выражения в (30), (31) и учитывая, что $z_i^0 = 0$, получим:

$$\|\underline{z}_n\|_n \leqslant \tau \cdot 0(h^2 + \tau) \cdot N \cdot e^{T(L_u + 2c)} = 0(h^2 + \tau)'$$

$$\|\underline{z}_n\|_n \leqslant \tau \cdot 0(h^2 + \tau) \cdot N \cdot e^{2T(c + L_u)} = 0(h^2 + \tau).$$

Таким образом, для обеих схем (5), (6) получили

$$\max_i \ |z_i^n| \leq 0(h^2 + \tau).$$

Поскольку правая часть не зависит от n , эта оценка выполняется для всех n, и, следовательно,

$$\max_{i,n} |z_i^n| \leq 0(h^2 + \tau).$$

Теорема доказана.

Замечание 1. Можно показать, что требование (24) можно ослабить. В частности, вместо (32) можно записать:

$$1 - \tau c \geq e^{-\tau k c},$$

что выполняется для тем большего интервала τ, чем больше k, и для больших k граница сверху приближается к $\tau \leq 1/C$. Однако, в этом нет необходимости, т.к. требование (24) легко выполнимо.

Замечание 2. Легко видеть, что из оценки (24) следует (23) для значения

$$q = \frac{L_u}{2L_{\ddot{u}} + c} < 1.$$

Перейдем к вопросу об устойчивости разностных схем (5), (6) относительно начальных условий и функции f.

Пусть y_h - точное решение разностной задачи (5), (7) /или (6), (7) /. Далее, пусть функции f и g получили некоторое возмущение. Обозначим эти возмущенные функции через $\tilde{f}$ и $\tilde{g}$, а через

$\tilde{y}_h$ - решение системы уравнений

$$D_n \underline{y}_n = -\underline{y}_{n-1} - \tau\underline{\tilde{f}}_{n-1}(\underline{y}_{n-1}),$$

или (33)

$$D_n y_n = -y_{n-1} - \tau\tilde{\tilde{f}}_n(y_n)$$

с начальными и граничными условиями:

$$y_i^0 = \tilde{g}(x_i), \quad y_0^n = g_1(t^n), \quad y_M^n = g_2(t^n) \tag{34}$$

Относительно возмущенной функции $\tilde{f}$ пологаем, что она удовлетворяет условию Липшица

$$|\tilde{f}(x_i,t^n,y_1) - \tilde{f}(x_i,t^n,y_2)| \leqslant \tilde{L} \cdot |y_1 - y_2|, \tag{35}$$

а функция $\tilde{g}$ согласована с граничными условиями.

Теорема 2. При выполнении условий (9), (35), а также

$$\tau \leqslant \frac{1}{2(c+\tilde{L})} \tag{36}$$

схемы (5), (7) и (6), (7) устойчивы по начальным данным и функции f.

Доказательство. Обозначим:

$$\tilde{z}_h = \tilde{y}_h - y_n$$

Тогда для $\underline{\tilde{z}}_h$ получим уравнения

$$\begin{aligned} D_n\underline{\tilde{z}}_n &= -\underline{\tilde{z}}_{n-1} - \tau(\underline{f}_{n-1}(\underline{y}_{n-1}) - \underline{\tilde{f}}_{n-1}(\underline{\tilde{y}}_{n-1})), \\ D_n\underline{\tilde{z}}_n &= -\underline{\tilde{z}}_{n-1} - \tau(\underline{f}_n(\underline{y}_n) - \underline{\tilde{f}}_n(\underline{\tilde{y}}_n)), \end{aligned} \tag{37}$$

причем $\tilde{z}_i^o = g(x_i) - \tilde{g}(x_i)$, $\tilde{z}_o^n = \tilde{z}_M^n = 0$. Обозначая

$$r = \max_n \|f_n(\underline{y}_n) - \underline{\tilde{f}}_n(\underline{y}_n)\|_n$$

и применяя рассуждения, использованные при доказательстве Теоремы 1, получим из уравнений (37) для двух схем соответственно:

$$\|\tilde{\underline{z}}_n\|_n \leq e^{T(\tilde{L}+2c)} \cdot \|\underline{z}_0\|_0 + T \cdot e^{T(\tilde{L}+2c)} \cdot r,$$
$$\|\tilde{\underline{z}}_n\|_n \leq e^{2T(c+\tilde{L})} \cdot \|\tilde{\underline{z}}_0\|_0 + T \cdot e^{2T(c+\tilde{L})} \cdot r \qquad (38)$$

Правые части не зависят от n, поэтому неравенства (38) выполняются для всех n. Следовательно, можно записать:

$$\max_{i,n} |\bar{y}_i^n - \tilde{y}_i^n| \leq K_1 \cdot \max_i |g(x_i) - \tilde{g}(x_i)| + K \cdot \max_{i,n} |f(\bar{y}_i^n) - \tilde{f}(\bar{y}_i^n)| \qquad (39)$$

где значения K_1 и K очевидны. Итак, Теорема 2 доказана.

Относительно условия (36) можно сказать то же, что было сказано в Замечании 1 после Теоремы 1.

Нами были проведены расчеты для задачи (1), (2) с использованием разностной схемы, несколько упрощенным вариантом которой является схема (6), (7), а также для разностного уравнения (6), но с несколько более сложными граничными условиями. Результаты численного расчета показали достаточно быструю сходимость итерационного процесса и устойчивость выбранной схемы.

Литература

1. Д. Молнарка, Р.Х. Фарзан. О приближенном решении методом конечных разностей краевой задачи для одномерного дифференциального уравнения параболического типа со слабой нелинейностью. 1. Явная схема.

Annales Univ. Sci. Budapest
Sectio Commutatorica (в печати).

2. А.А. Самарский. Введение в теорию разностных схем.

Москва, "Наука", 1971.

3. Н.С. Бахвалов. Численные методы.
Москва, "Наука", 1973.

Összefoglaló

Implicit differencia séma a parabolikus, egydimenziós gyöngén nemlineáris parciális egyenlet numerikus megoldására

D. Molnárka – R.H. Farzan – L. Fái

A munka parabolikus tipusú, gyöngén nemlineáris, egydimenziós parciális differenciálegyenlet numerikus megoldásával foglalkozik. Elégséges feltételt kapunk, mely biztositja a differenciálegyenletre felirt differencia séma stabilitását és konvergenciáját egyenletes normában. A differenciáloperátort approximáló véges differencia operátor inverzének normájára kapunk becslést. Ennek felhasználásával lehet vizsgálni a véges differencia egyenlet, mint nemlineáris egyenletrendszer, iterációs módszerrel való megoldhatóságának feltételét és a megoldásnak az eredeti feladat megoldásához való konvergenciáját.

Summary

About the implicit difference schemes for the solution of weakly nonlinear parabolic differential equation with one space dimension

D. Molnárka – R.H. Farzan – L. Fái

In this paper a suefficient condition is examined for uniform convergence and stability of difference schemes for numerical solution of weakly nonlinear parabolic partial differential equation. The norm of the inverse oparator of the finite difference operator was estimated. Using up this estimation it can be showed that the difference scheme, what is a system of nonlinear algebraic equation, can be solved by iterational processes and its solution converes to the solution of differential equation.

MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 21/1978.

ПРИМЕНЕНИЕ ТЕОРИИ РАЗМЫТЫХ МНОЖЕСТВ В ФОРМАЛЬНОЙ СИСТЕМЕ КЛАССИФИКАЦИИ

Вольфганг Бельке, Бернд Кёппен, Хартмут Еттрих

I. Основы теории размытых множеств.

В науке и её конкретной реализации существует множество проблем, при которых точное описание исследуемых объектов невозможно. Если эта неопределённость вызвана случайными величинами, то решение задач может быть найдено методами теории вероятности или математической статистики. Но если эти условия не выполняются, тогда необходимо применение других методов /см. KUMMER 77 /. Например, в этом случае применима теория размытых множеств, которая была развита Заде /ZADEH 65/ в середине 60-х годов. Обзор литературы к этой теории и главные области её применения можно найти у /ГУСЕВА, 73/.

В обычной теории множеств вопрос о том, является ли некоторый объект x элементом множества $\underline{M}$ решается всегда альтернативно, т.е. $x \in \underline{M}$ или $x \notin \underline{M}$. В теории размытых множеств такого ограничения нет и допускается задание степени принадлежности элемента к множеству при помощи вещественного числа.

Размытое множество определено следующим образом:

Размытое множество $\underset{\sim}{M}$ над множеством $\underline{X}$ представляет собой отображение множества $\underline{X}$ на интервал $[0,1]$.

$$(1.1.) \qquad \underset{\sim}{M}: \ \underline{X} \longrightarrow [0,1]$$

Каждому $x \in \underline{X}$ ставится в соответствие некоторое число интервала $[0,1]$, которое называется функцией или степенью принадлежности элемента x к множеству $\underset{\sim}{M}$.

Когда множество $\underline{X}$ является дискретным, тогда размытое множество можно написать как множество пар $(x, f_{\underset{\sim}{M}}(x))$

$$(1.2.) \qquad \underset{\sim}{M} = \{ (x_1, f_{\underset{\sim}{M}}(x_1)), (x_2, f_{\underset{\sim}{M}}(x_2)), \dots \}$$

В общем случае отображение может осуществляться на произвольный интервал L:

$$(1.3.) \qquad \underset{\sim}{M}: \underline{X} \longrightarrow L$$

Для большинства случаев применения нормированный интервал $[0,1]$ имеет особое значение. Если $L = [0,1]$, т.е. $\sup f_{\underset{\sim}{M}}(x) = 1$, то говорят, что размытое множество $\underset{\sim}{M}$ является нормализованным. В этой статье рассматриваются только нормализованные размытые мно-

множества.

Каждое размытое множество $\underset{\sim}{M}$ имеет своё множество-носитель, которое определяется следующим образом:

Множество-носитель $\underline{S}$ некоторого размытого множества $\underset{\sim}{M}$ содержит все элементы размытого множества, для которых степень принадлежности больше нуля.

$$(1.4.) \qquad \underline{S}(\underset{\sim}{M}) = \{ x \mid f_{\underset{\sim}{M}}(x) > 0 \}$$

Множество-носитель размытого множества представляет таким образом неразмытое /чёткое/ множество.

Принципиальное различие теории размытых множеств от теории вероятности находит своё явное выражение в определениях операций над размытыми множествами. Таким образом, например, операции объединение и пересечение определены как применение операторов **max** и **min** над функциями принадлежности. Основные определения операции теории размытых множеств можно найти у (ZADEH 65) и (KUMMER 77).

2. Основы классификации

Исходным пунктом классификации является множество объектов. Под понятием объект подразумеваются как реально существующие так и абстрактные объекты / абстракции, модели, проблемы, алгоритмы/.

$$(2.1.) \qquad \underline{O} = \{ o_r \,,\ r = 1, 2, \dots R \}$$

Эти объекты формально описываются в соответствии с выбранным аспектом рассмотрения, т.е. свойства, которые при аспекте рассмотрения являются существенными, формулируются в виде признаков. Каждый признак задан множеством его значений $\underline{M}_i$, все признаки образуют множество признаков $\underline{M}$.

$$(2.2.) \qquad \underline{M} = \{ M_i \,,\ i = 1, 2, \dots I \}$$

$$(2.3.) \qquad \underline{M}_i = \{ m_{ij} \,,\ j = 1, 2, \dots J_i \}$$

Эта запись действительна для дискретных и конечных множеств значений признаков, которые рассматриваются в рамках этой статьи. Совокупность всех признаков и их значений в виде таблицы называется классификатором:

$$(2.4.)\quad \underline{K} = \begin{array}{c} \begin{array}{cccc} M_1 & M_2 & \dots & M_I \end{array} \\ \left[\begin{array}{cccc} m_{11} & m_{21} & \dots & m_{I1} \\ m_{12} & m_{22} & & m_{I2} \\ \vdots & \vdots & & \vdots \\ m_{1k} & m_{2l} & & m_{Im} \end{array}\right] \end{array}$$

Строго говоря, эта таблица представляет собой лишь схему классификации, а к классификатору относятся и правила применения схемы, которые обычно можно считать тривиальными / см. BELKE u.a. 77 /. Описание объектов таким образом можно рассматривать как отображение этого множества на классификатор $\underline{K}$.

$$(2.5.)\quad \underline{OB} : \underline{O} \longrightarrow \underline{K},$$

причём $\underline{OB}$ - множество описаний объектов.

$$(2.6.)\quad \underline{OB} = \{ob_1, ob_2, \dots ob_R\}$$

$$card(\underline{OB}) = card(\underline{O})$$

Для одного объекта o_r при одном аспекте рассмотрения существует только одно описание объекта ob_r. Такое описание объекта представляет собой n-набор, элементами которого являются конкретные значения признаков, отражающие свойства объекта.

$$(2.7.)\quad ob_r = (m_{1i}, m_{2j}, \dots m_{Il})$$

Когда для каждого признака можно указать точно одно значение, т.е. когда каждое место n-набора занято, тогда имеет место следующее:

$$(2.8.)\quad ob_r \in \prod_{i=1}^{I} \underline{M}_i$$

Описания объектов такого типа называются полными; когда хотя бы для одного признака заданы больше одного или никакого значения, тогда можно написать:

$$(2.9.)\quad ob_r \in \prod_{i=1}^{I} P(\underline{M}_i)$$

Если, по крайней мере, для одного признака значение не задано, то описание объекта называется неполным.

Под классификацией понимается процесс разложения заданного множества объектов $\underline{O}$ в соответствии с определёнными признаками в классы k_l. Под классом подразумевается подмножество множества объектов, выбранное при определённом аспекте рассмотрения / определённая задача, цель /.

$$(2.10.) \qquad \underline{k}_l = \{ o_i, o_j, \dots \}$$

$$(2.11.) \qquad \underline{k}_l \subseteq \underline{O}$$

Предписание, которое определяет, какие объекты следует отнести к классу $\underline{k}_l$, называется описанием класса kb_l и представляет собой логическое выражение / т.е. значения признаков, связанные друг с другом логическими операторами дисъюнкция, конъюнкция и отрицание /. Класс $\underline{k}_l$ содержит все объекты o_r, для которых высказывание $A(ob_r, kb_l)$, состоящее в том, что описание объекта соответствует описанию класса, является истинным, т.е. равно логическому значению 1.

$$(2.12.) \qquad \underline{k}_l = \{ o_r \mid A(ob_r, kb_l) = 1 \}$$

Процесс образования классов можно рассматривать как отображение множества объектов на множество логических значений 0 и I.

$$(2.13.) \qquad \underline{k}_l : \ \underline{O} \longrightarrow \{0, 1\}$$

3. Размытая классификация

Применение теории размытых множеств в формальной системе классификации расширяет её и даёт возможность решения новых задач. В рамках этой статьи рассматривается только применение этой теории на неметрическую классификацию / множества значений признаков являются неметрическими, дискретными и конечными /. При этом можно различить 3 случая использования размытых множеств, которые рассматриваются ниже.

3.1. Размытое описание объектов

Предполагается, что существуют объекты, которые нельзя точно описать из-за объективных причин, т.е. для которых задание значений признаков в смысле альтернативной постановки вопроса / значение признака m_{ij} является для объекта o_r релевантным или нет / невозможно. Тогда описание объектов осуществляется путём задания значения признака и его степени принадлежности к объекту. Эта степень принадлежности выражается числом из интервала $[0, 1]$ и полностью соответствует понятию функции принадлежности из теории размытых множеств. Описания объектов тогда имеет следующий вид:

$$(3.1.) \quad ob_r = ((m_{1i}, f_{or}(m_{1i})), (m_{2j}, f_{or}(m_{2j})), \ldots)$$

Описания объектов представляют таким образом n- наборы, элементами которых являются пары значений признака - функция принадлежности. Размытое описание объектов есть процесс размытого соотнесения значений признаков к объектам. Этот процесс можно рассматривать как отображение классификатора на интервал $[0, 1]$.

$$(3.2.) \quad \underset{\sim}{\underline{OB}}: \quad \underline{O} \longrightarrow (\underline{K} \longrightarrow [0, 1])$$

Определение степени /функции/ принадлежности является в настоящее время субъективным процессом, т.е. оно проводится самим использователем. Необходимо создать более объективные способы /алгоритмы/ для определения $f_{or}(m_{ij})$. Для объективизации можно использовать методы теории экспертных оценок.

3.2. Размытый классификатор

Особый интерес представляют признаки, имеющие структуру, т.е. между значениями которых существуют взаимосвязи, отношения. Формально такой признак можно написать в виде структурированного множества:

$$(3.3.) \quad \underline{M}_i^S = [m_{i1}, m_{i2}, \ldots m_{ij_i}; \underline{R}]$$

$$(3.4.) \quad \underline{R} = \{ \underline{M}_i \times \underline{M}_i, R_1, R_2, \ldots \}$$

В качестве примера приводятся следующие отношения:

R_1 - m_{ij} имеет такое же значение /является синонимом/ как m_{ik}

R_2 - m_{ij} подчиняется m_{ik}
/при определённом аспекте/

R_3 - m_{ij} является противоположным к m_{ik}

Эти отношения используются в рамках информационно-поисковых систем /ИПС/ для структурирования множества значений признака "дескрипторы". Если такое множество имеет достаточно высокую мощность, то оно называется тезаурусом. Используя структуру значений признака /дескрипторов тезауруса/ при информационном поиске, можно достичь значительного улучшения работы ИПС, особенно улучшается при этом удобство пользования.

При разработке тезаурусов оказывается, что не всегда возможно однозначное отнесение дескрипторов к отношениям. Существует много слов, смысловые содержания которых похожи, но не полностью равны, т.е. вопрос об отнесении этих дескрипторов к отношению "быть синонимом" нельзя решить альтернативно. В этом случае оказывается целесообразным применение понятия "размытое отношение".

$$(3.5.) \qquad \underset{\sim}{R}_p = \{((m_{ik}, m_{il}), f_{\underset{\sim}{R}_p}(m_{ik}, m_{il})), \ldots\}$$

$f_{\underset{\sim}{R}_p}(m_{ik}, m_{il})$ - функция принадлежности пары (m_{ik}, m_{il}) в отношении $\underset{\sim}{R}_p$. Применение понятия размытого отношения даёт возможность формально описать малые, но существующие и в определённом аспекте важные семантические разницы. Такие понятия как ЭВМ и вычислительная машина часто используются как синонимы, но всё-таки существует разница в семантическом содержании / к вычислительным машинам относятся и неэлектронные ВМ /. Это обстоятельство можно формально описать таким образом, что отнесение к отношению "быть синонимом" осуществляется с функцией принадлежности меньше 1. Наиболее часто встречающиеся отношения в ИПС можно делить в зависимости от их свойств на следующие типы:

- отношение эквивалентности / свойства: транзитивность,
  рефлексивность,
  симметричность /
- отношение полуупорядоченности / свойства: транзитивность,
  антисимметричность /

Разделение отношений на эти типы имеет значение для применяемого алгоритма классификации.

3.3. Размытое описание классов

В параграфе 2 этой статьи класс был определён как множество объектов, для которых высказывание, состоящее в соответствии описания объектов и описания класса, является истинным, т.е. имеет логическое значение 1. В некоторых практических случаях оказывается целесообразным отнести к классу объекты, описания которых не полностью соответствуют заданному описанию класса, т.е. для которых уравнение следующее из формулы /2.12./

$$(3.6.) \qquad A(ob_r, kb_l) = 1$$

не выполняется. Это условие приводит к условию:

$$(3.7.) \qquad A(ob_r, kb_l) \geq \gamma \qquad 0 < \gamma \leq 1$$

Таким образом, теперь становится ясным, что применение теории размытых множеств требует использования многозначной логики. В общем случае достаточно допустить дискретно-многозначное множество значений истинности, т.к. функции принадлежности определяются измерениями, вычислениями или оценкой и в этих процессах всегда существует предел точности. Принципиально возможны непрерывно-значные множества значений истинности / см. JAHN 75, стр. 38/. Неполное соответствие описания объекта и описания класса учитывается таким образом, что объекты будут отнесены к классу с функцией принадлежности меньше 1. Процесс образования классов представляет собой отображение множества объектов $\underline{O}$ на интервал $[0,1]$:

$$(3.8.) \qquad \underset{\sim}{k_l} : \underline{O} \longrightarrow [0,1]$$

Каждый класс k_l есть размытое множество:

$$(3.9.) \qquad \underset{\sim}{k_l} = \{(o_r, f_{\underset{\sim}{k_l}}(o_r))\}$$

Из формулы / 3.7. / следует, что всегда

$$(3.10.) \qquad f_{\underset{\sim}{k_l}}(o_r) \geq \gamma$$

При заданных объектах количество объектов, находящееся в определённом классе, не только зависит от описания этого класса, но и от γ. Для двух значений $\gamma_1 > \gamma_2$ имеет место отношение:

$$(3.11.) \qquad \underset{\sim}{k_{l\gamma_1}} \subset \underset{\sim}{k_{l\gamma_2}}$$

Выбирая параметр γ можно управлять размытием класса, т.е. чем ближе γ к 1, тем меньше размытым /"более чётким"/ становится класс. Для $\gamma = 1$ получается случай неразмытого /чёткого/ класса обычной классификации.

4. Алгоритмы классификации

Под классификацией понимается процесс отнесения формально описанных объектов к классам. Если исходят от заданных,полных и неразмытых /чётких/ описаний объектов, то алгоритм классификации реализует сравнение описаний объектов с описаниями классов и относит объекты при положительном результате сравнения к классу. Если используется один из описанных в §3 видов размытия, то главная задача алгоритма классификации заключается в вычислении функции принадлежности объектов к классам.
Алгоритм классификации, позволяющий обрабатывать все возможные виды размытия, является очень сложным. На практике редко встречаются все виды размытия одновременно; таким образом оказывается целесообразным создать алгоритмы для выбранных специальных случаев.

4.1. Алгоритм классификации для объектов с размытым описанием

Этот алгоритм используется при заданных размытых описаниях объектов /формула 3.1./ и неразмытом классификаторе /форм. 2.4./. Описания классов представляют собой логические выражения, состоящие из значений признаков, связанные логическими операторами конъюнкции, дисъюнкции и отрицания. Они определены без размытия, но из-за размытого описания объектов классы являются также размытыми множествами, причём необходимо определить степень принадлежности объектов к классу /см. ф. 3.9.:

$$\underset{\sim}{k}_l = \{ (o_r, f_{\underset{\sim}{k}_l}(o_r)) \}$$

В случае, когда описание класса состоит только из одного значения признака

$$kb_l = m_{ij}$$

степень принадлежности любого объекта o_r в описанном таким образом классе вычисляется как степень принадлежности этого значения признака m_{ij} в описании объекта ob_r :

$$(4.1.) \qquad f_{\underset{\sim}{kl}}(o_r) = f_{ob_r}(m_{ij})$$

Описание класса любой сложности всегда можно разложить в более простые описания. Над функциями принадлежности выполняется операция из теории размытых множеств, которая соответствует логическому оператору разложения.

- Разложение описания класса на две конъюнктивно связанные части

$$(4.2.) \qquad kb_l = kb_k \wedge kb_m$$

$$f_{\underset{\sim}{kl}}(o_r) = \min\left(f_{\underset{\sim}{kk}}(o_r),\ f_{\underset{\sim}{km}}(o_r)\right)$$

- Разложение описания класса на две дисъюнктивно связанные части

$$(4.3.) \qquad kb_l = kb_k \vee kb_m$$

$$f_{\underset{\sim}{kl}}(o_r) = \max\left(f_{\underset{\sim}{kk}}(o_r),\ f_{\underset{\sim}{km}}(o_r)\right)$$

- Представление описания класса как отрицание другого описания

$$(4.4.) \qquad kb_l = \neg\, kb_k$$

$$f_{\underset{\sim}{kl}}(o_r) = 1 - f_{\underset{\sim}{kk}}(o_r)$$

При помощи рекурсивного применения формул 4.2. - 4.4. можно вычислить степень принадлежности объектов к классам, имеющим описания любой сложности.

4.2. Алгоритм классификации при существовании структурированных множеств значений признаков

Исходными являются предпосылки, сделанные в параграфе 4.1. Кроме того, считается, что множества значений признаков являются структурированными, т.е. над ними определены неразмытые отношения. Эта структура должна учитываться при классификации, т.е.

при отнесении объектов к классам и при вычислении степени принадлежности. К классу следует отнести также объекты, которые содержат в своём описании не те значения признаков, которые требует описание класса, а те, которые находятся в отношении к требуемым.

$$(4.5.)\qquad \begin{array}{l} kb_l = m_{ij} \\ m_{ij} \notin \underline{S}(ob_r) \\ \exists_k (m_{ij} R m_{ik} \wedge m_{ik} \in \underline{S}(ob_r)) \end{array}$$

Алгоритм вычисления для $f_{\underset{\sim}{kl}}(o_r)$ зависит от типа отношения.

4.2.1. Отношение эквивалентности

Из всех значений признаков m_{ik}, находящихся в отношении R к m_{ij} учитывается то, для которого степень принадлежности в описании объекта максимальна.

$$(4.6.)\qquad f_{\underset{\sim}{kl}}(o_r) = \max_{\forall_k (m_{ij} R m_{ik})} f_{ob_r}(m_{ik})$$

4.2.2. Отношение полуупорядочения

В связи с тем, что отношение полуупорядочения имеет свойство антисимметричности, для определения $f_{\underset{\sim}{kl}}(o_r)$ существуют две возможности /с учётом предпосылок 4.5./:

$$(4.7.)\qquad f_{\underset{\sim}{kl}}(o_r) = \max_{\forall_k (m_{ij} R m_{ik})} f_{ob_r}(m_{ik})$$

$$(4.8.)\qquad f_{\underset{\sim}{kl}}(o_r) = \max_{\forall_k (m_{ik} R m_{ij})} f_{ob_r}(m_{ik})$$

Выбор формул /4.7./,/4.8./ зависит от конкретного семантического содержания отношения и от цели классификации. Если R, например, является отношением "быть более общим", то при применении формулы /4.7./ к классу относятся все те объекты, которые описаны значениями признаков, являющимися более специфическими, чем заданные в описании класса. При применении формулы /4.8./ учиты -

ваются все более общие значения признаков.
Если предпосылка /4.5./ не выполняется, т.е.

$$m_{ij} \in \underset{\sim}{S}(ob_r)$$

то можно применить более простой алгоритм 4.1.
Разложение более сложных описаний классов опять проводится при помощи рекурсивного применения формул 4.2 - 4.4.

4.3. Алгоритм классификации для размыто описанных объектов при наличии размытых отношений

Допустим, задано множество объектов, которые размыто описаны по формуле /3.1./ и над множествами значений признаков определены размытые отношения по формуле /3.5./.
Наглядное решение этой проблемы можно получить с использованием теории графов. При этом структурированное множество значений признаков изображается как ориентированный нагруженный граф. Значения признаков представляют нагрузки вершин. Если два значения признаков являются элементами отношения, т.е.

$$(m_{ij}, m_{ik}) \in \underset{\sim}{S}(\underset{\sim}{R})$$

то эквивалентные им вершины связываются дугой, нагруженной степенью принадлежности пары (m_{ij}, m_{ik}) в размытом отношении.
Если речь идёт об отношении эквивалентности, то его можно изобразить как неориентированный граф. Алгоритм, который будет описан ниже, использует отношение эквивалентности.

Пример:

M_1 - средство труда

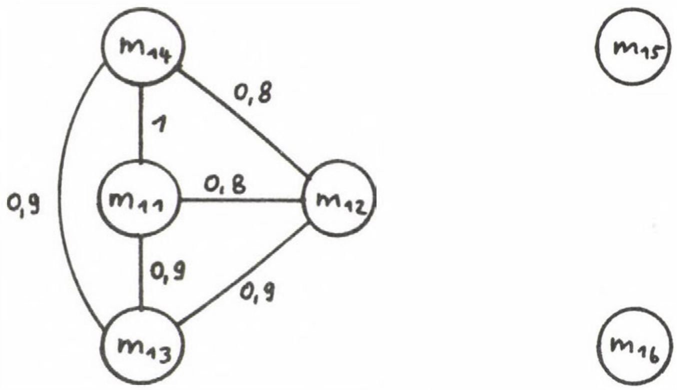
m_{11} - ЭВМ
m_{12} - вычислительная машина
m_{13} - вычислительное оборудование
m_{14} - компьютер
m_{15} - токарный станок
m_{16} - сверлильный станок

В качестве отношения определяется отношение " быть синонимом".
Из свойств этого отношения /симметричность, рефлексивность,

транзитивность/ следует, что оно является отношением эквивалентности. Для задания степеней принадлежности пар в отношении целесообразно использовать таблицу:

	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_{16}
m_{11}	1	0,8	0,9	1	0	0
m_{12}		1	0,9	0,8	0	0
m_{13}			1	0,9	0	0
m_{14}				1	0	0
m_{15}					1	0
m_{16}						1

граф:



Как мера связи взаимоотношения значений признаков вводятся коэффициенты эквивалентности. Коэффициент эквивалентности a_{jk} двух значений признаков m_{ij} , m_{ik} определяется как максимальная нагрузка всех путей, которые существуют между ними. Нагрузка пути $m_{ij}, m_{ip}, \ldots , m_{ik}$ вычисляется умножением нагрузок дуг, образующих путь. Таким образом достигается цель, состоящая в том, что нагрузка пути уменьшается с каждым значением признака, т.е. с удлинением пути, с увеличением расстояния. Всегда имеет место:

$$(4.10.) \qquad 0 \leq f_{\underset{\sim}{R}}(m_{ij}, m_{ik}) \leq 1$$

a_{jk} можно вычислить по следующей формуле:

$$(4.11.) \qquad a_{jk} = \begin{cases} 1 & j=k \\ \max\limits_{p,q} \left(f_{\underline{R}}(m_{ij}, m_{ip}) \cdot \ldots \cdot f_{\underline{R}}(m_{iq}, m_{ik}) \right) & \\ 0 & \text{нет пути } m_{ij} - m_{ik} \end{cases}$$

Коэффициенты эквивалентности можно написать в форме треугольной матрицы:

$$A = \| a_{jk} \| \qquad \begin{array}{c|l} & m_{ik} \\ \hline m_{ij} & 1 \\ & -1 \quad a_{jk} \\ & —\ \cdot \\ & ——\ \cdot \\ & ———\ \cdot \\ & ————\ 1 \end{array}$$

Так как эта матрица не зависит от классов / т.е. от описаний классов /, то её можно считать заданной. Её всегда можно вычислить при помощи ф./4.11/ из заданного классификатора и определённых над ним отношений. В дальнейшем рассматривается только случай, когда описание класса kb_e состоит из одного значения признака

$$kb_e = m_{ip}$$

Более сложные описания классов можно разложить по формулам /4.2.-4.4./ в простые описания. Кроме того, имеет место:

$$m_{ip} \notin \underline{S}(ob_r)$$

/ для случая $m_{ip} \in \underline{S}(ob_r)$ можно применить алгоритм, описанный в пп. 4.1./

Пусть существуют значения признаков, которые содержатся в описании объекта ob_r и находящиеся в отношении к m_{ip}

$$\exists_k (a_{pk} > 0)$$

Объекты, описанные такими значениями признаков, также следует отнести к классу k_e, но их степень принадлежности к классу k_e зависит не только от функции принадлежности соответствующего значения признака в описании объекта, но и от связи /расстояния/ значений признаков в описании объекта и описании класса, т.е. от коэффициентов эквивалентности. Степень принадлежности объекта o_r в классе k_l $(kb_l = m_{ip})$ можно тогда вычислить по следующей формуле:

$$(4.12.) \quad f_{k_l}(o_r) = \max_k \left(f_{ob_r}(m_{ik}) \cdot a_{pk} \right)$$

Для случая, когда отношение, определённое над множеством значений признаков, не является отношением эквивалентности, можно создать аналогичный алгоритм; отношения в этом случае изображаются в виде ориентированного графа. Меняя определение коэффициентов эквивалентности, можно учесть специфику данного отношения. Алгоритм, описанный в пп. 4.2. является специальным случаем описанного выше алгоритма.
Если запрещаются размытые отношения, т.е.

$$(4.13.) \quad f_R(m_{ij}, m_{ip}) \in \{0, 1\}$$

то коэффициенты эквивалентности могут принимать только значения 0 и 1

$$a_{jp} \in \{0, 1\}$$

и из ф./ 4.11./ получается ф./ 4.6./.
В заключение, ещё раз следует обратить внимание на то, что исходным для предложенных алгоритмов классификации является общее описание объектов по формуле /3.1./. Оно было выведено из чёткого описания по ф. /2.7.-2.9./. Применение алгоритмов классификации таким образом многократно /т.е. и по отношению к различным M_i рекурсивно.

5. Применение при автоматизированном синтезе программ

Принцип работы системы автоматизированного синтеза программ, который был разработан на основе применения многоуровневой архитектуры системы базы программ и с использованием библиотеки алгоритмов, был описан в работе / BELKE u.a. 77 /. Основные компоненты этой системы: банк проблем, банк алгоритмов и банк программ. Банк проблем содержит классы проблем, для которых существует решение /в смысле ссылок на алгоритмы и программы/. Пользователь формулирует на языке описания проблем /дескрипторный язык/ свою проблему.

В связи со сложностью проблемы не всегда возможно однозначное отнесение значений признаков, т.е. имеется размытое описание объектов. Например, к признаку "тип производства" при исследованиях конкретных предприятий иногда нельзя чётко отнести одно из значений: единичное производство, ... , крупносерийное, массовое. Сложные проблемы обычно из-за большого количества признаков описываются неполно. Применяя структурированный, размытый классификатор, можно провести автоматизированное дополнение описания проблемы. Несмотря на это, процесс описания проблемы является итеративным процессом, который можно провести только в диалоге человека с машиной. На первых этапах этого процесса однозначное отнесение /нечётко и неполно/ описанной проблемы к классам невозможно - отнесение будет размытым. Цель диалога человека с машиной состоит в уменьшении размытия до такой степени, что достигается однозначное отображение на соответствующие алгоритмы и программы.

Для проведения экспериментов был спроектирован пакет программ "Размытая классификация" и был создан первый вариант. Пакет программ имеет модульную структуру и осуществляет размытое отнесение размыто описанных объектов к классам. Предусмотрено применение размытого структурированного классификатора. Все массивы /массив описаний объектов, классификатора и описаний классов/ являются индексно-последовательными массивами на магнитных дисках. Таким образом, количество обрабатываемой информации ограничивается лишь числом сменных магнитных дисков. Имя признаков и значения признаков при вводе автоматически кодируются, внутри системы используются только коды. Пакет программ написан на языке PL/I и испытан на ЭВМ ЕС 1022.

Литература:

BELKE u.a. 77/1: Belke, W.; Graichen, D.; Starruß, M.: Nichtmetrische Klassifizierung von Informationen; Buchmanuskript, Dresden 1977

BELKE u.a. 77/2: Belke, W.; Oettrich, H.: Algorithmenbank und Synthese von problemorientierten systemtechnologischen Unterlagen (POS), TU-Informationen 08-10-77

GUSEV 73: Gusev, I.A.; Smirmova, I.A.: Razmytye mnozestva, in: Avtomatika i telemechanika 1973, 5, S. 66

JAHN 75: Jahn, K.-U.: Anwndungen von fuzzy sets; in: Vorträge zur Algorithmen und Automatentheorie, Heft 16/76, Weiterbildungszentrum für mathematische Kybernetik und Rechentechnik, Sektion Mathematik, Technische Universität Dresden

KUMMER 77: Kummer, B.; Straube, B.: Eine Einführung in die Theorie unscharfer Mengen; in: Wissenschaftliche Zeitschrift d. TU Dresden 26 (1977),2, S. 363-369

ZADEH 65: Zadeh, L.A.: Fuzzy sets; in: Information and control 8 (1965) s.338-353

S u m m a r y

The application of fuzzy set theory in formal systems of classification

W. Belke – B. Kappen – H. Ettrich

The author describes the ground concepts of classification and fuzzy set-theory. He introduces the fuzzy-description of objects, the concepts of fuzzy-relation and fuzy-classification. Using these he gives a method and an algorithm to solve not exactly describable classification problems.

Ö s s z e f o g l a l ó

Fuzzy-halmazok elméletének alkalmazása klasszifikáció formális rendszerében

W. Belke – B. Kappen – H. Ettrich

Szerző ismerteti a fuzzy-halmazok és a klasszifikáció elméletének alapjait. Bevezeti objektumok fuzzy-leirását, a fuzzy-reláció és fuzzy-klasszifikáció fogalmát. Ezek felhasználásával módszert és algoritmust ad nem pontosan leirható klasszifikációs problémák megoldására.

MTA Számitástechnikai és Automatizálási Kutató Intézete, Közlemények 21/1978.

К В А З И М О Д У Л И , I.

До Лонг Ван, Нгуен Куок Тоан

§ 0. Введение

В этой заметке авторы предлагают одну алгебраическую систему так называемую квазимодулем.

По определению квазимодуля, введенному в §1, класс всех квазимодулей включает в себя, в частности, все группы и все модули. Некоторые конструкции, позволяющие построить новые квазимодули из заданных квазимодулей указаны. Благодаря этому, в частности, можно построить простые примерыквазимодулей, не являющихся ни группами ни модулями. Одно условие, необходимое и достаточное для того, чтобы одна группа с кольцом унарных операций была квазимодулем, потом, сформулировано.

В §2 сформулируются характеристики некоторых классов знакомых групп /Бернсайдовых групп, ΠD - групп, D - групп/ в терминах квазимодулей. Это, с одной стороны, дает более глубокие примеры квазимодулей, и с другой стороны, наводит на мысль одного подхода к группам со стороны колец.

§ 1. Определение и примеры.

1. Под кольцом мы будем понимать ассоциативное кольцо /не обязательно с единицей/. Квазимодулем над кольцом Ω , или короче, Ω-квазимодулем будем называть любую группу G с кольцом Ω унарных операций, удовлетворяющих, для любых $a \in G$; $\alpha, \beta \in \Omega$, следующим условиям

/КВМ.0/ $\quad e^{\alpha} = e$;

/КВМ.1/ $\quad a^{\alpha+\beta} = a^{\alpha} . a^{\beta}$;

/КВМ.2/ $\quad a^{\alpha\beta} = (a^{\alpha})^{\beta}$, где через a^{α} обозначаем результат применения операции $\alpha \in \Omega$ к элемнту $a \in G$, e - единица группы G .

Если, кроме этого, кольцо Ω обладает единицей ε и для любого $a \in G$

/КВМ.3/ $\quad a^{\varepsilon} = a$,

то Ω - квазимодуль G называется унитарным.

Непосредственно из /КВМ.1/ и /КВМ.2/ получаем

/КВМ.4/ $\quad a^{0} = e$,

и также

/КВМ.5/ $\quad a^{-\alpha} = (a^{\alpha})^{-1}$

где 0 - нулевой элемент кольца Ω .

А если G унитарен, то, ввиду /КВМ.2/ и /КВМ.3/, мы имеем также

/КВМ.6/ $\quad a^{-\alpha} = (a^{-1})^{\alpha}$

Замечание 1. Класс всех квазимодулей над заданным кольцом составляет многообразие Ω - групп /об Ω - группах см., например, [1] или [2] /. Многие понятия и результаты из теории Ω - групп, поэтому автоматически перенесены на случай квазимодулей. Мы сможем говорить, например, о подквазимодулях и об идеалах данного квазимодуля, о фактор-квазимодуле по его идеалу, о гомоморфизме квазимодулей и т.д. Во избежание громоздкости, в дальнейшем такие понятия и факты мы будем использовать, считая известными.

В качестве первых примеров квазимодулей мы можем взять следующие две известные алгебраические системы.

Пример 1. Каждый Ω - модуль /унитарный/ является Ω - квазимодулем /унитарным/.

Пример 2. Каждая группа G превратится в унитарный квазимодуль над кольцом Z целых чисел если под "действием" операции $n \in Z$ на элемент $a \in G$ мы понимаем воздействие в степени n элемента a в группе G. Наоборот, если G любой унитарный Z-квазимодуль, то, ввиду /КВМ. 1/ и /КВМ. 3/ результат применения операции $n \in Z$ к $a \in G$ не что иное как возведение в сепени n элемента a в группе G. Итак, в каком-то смысле мы можем отождествлять группу G с унитарным Z-квазимодулем G.

2. Естественно возникает такой вопрос: существуют ли квазимодули, не являющиеся ни унитарными Z-квазимодулями, ни модулями. В следующем мы покажем, что таких квазимодулей много.

Прямой проверкой легко доказать следующую теорему

<u>Теорема 1.</u> Пусть I – любое множество индексов и для каждого $i \in I$, G_i есть квазимодуль /унитарный/ над некоторым кольцом Ω_i. Тогда декартово произведение G групп G_i, $G = \prod_I G_i$, будет квазимодулем /унитарным/ над декартовым произведением Ω колец Ω_i, $\Omega = \prod_I \Omega_i$.

<u>Пример 3</u>. В теореме 1 для каждого выбора квазимодулей G_i таких, что группа G не абелева и кольцо Ω неизоморфно кольцу Z целых чисел мы получаем квазимодуль /именно Ω - квазимодуль G /, не являющийся ни унитарным Z - квазимодулем, ни модулем. В частности, благодаря Теореме 1 можно простроить квазимодули, не являющиеся ни унитарными Z - квазимодулями, ни модулями, исходя только из унитарных Z - квазимодулей и модулей.

Система подгрупп $\{A_i \mid i \in I\}$ группы G такая, что $\bigcup_I A_i = G$ называется <u>покрытием</u> группы G. Если, кроме этого, подгруппы A_i абелевы, то это покрытие называется <u>абелевым покрытием</u>.

<u>Теорема 2</u>. Пусть G - произвольная группы и $\{ A_i \mid i \in I \}$ некоторое ее покрытие такое, что каждая подгруппа A_i явля-

ется квазимодулем /унитарным/ над некоторым кольцом Ω_i . Через p_i обозначением i -ую проекцию декартова произведения $\prod_I \Omega_i$. Тогда

$$\Omega = \{ \alpha \in \prod_I \Omega_i \mid \forall_{i,j} (a \in A_i \cap A_j \to a^{p_i \alpha} = a^{p_j \alpha})\}$$

будет кольцом с одними и теми же операциями сложения и умножения, определенными в $\prod_I \Omega_i$ /подкольцом кольца $\prod_I \Omega_i$ /, а G будет квазимодулем над Ω , если действие $\alpha \in A_i$ на $a \in G$ определяется следующим образом: для любого $a \in A_i$, $a^\alpha = a^{p_i \alpha}$

<u>Доказательство</u>. Видно, что для любых $a \in A_i \cap A_j$, $\alpha, \beta \in \Omega$, мы имеем

$$a^{p_i(\alpha+\beta)} = a^{p_i\alpha + p_i\beta} = a^{p_i\alpha} . a^{p_i\beta} = a^{p_j\alpha} . a^{p_j\beta} = a^{p_j\alpha + p_j\beta} = a^{p_j(\alpha+\beta)};$$

$$a^{p_i(-\alpha)} = a^{-p_i\alpha} = (a^{p_i\alpha})^{-1} = (a^{p_j\alpha})^{-1} = a^{-p_j\alpha} = a^{p_j(-\alpha)};$$

$$a^{p_i(\alpha\beta)} = a^{p_i\alpha . p_i\beta} = (a^{p_i\alpha})^{p_i\beta} = (a^{p_j\alpha})^{p_j\beta} = a^{p_j\alpha . p_j\beta} = a^{p_j(\alpha\beta)}.$$

/заметим, что $a^{p_i\alpha} = a^{p_j\alpha} \in A_i \cap A_j$ /. Это означает, что $\alpha + \beta$, $-\alpha$, $\alpha\beta$ также принадлежат Ω . Таким образом Ω является кольцом с одними и теми же операциями сложения и умножения, определенными в $\prod_I \Omega_i$.

Так как $e \in A_i$ для всех $i \in I$, то для любого $\alpha \in \Omega$

$$e^\alpha = e^{p_i\alpha} = e.$$

Потом при любых $a \in A_i$, $\alpha, \beta \in \Omega$

$$a^{\alpha+\beta} = a^{p_i(\alpha+\beta)} = a^{p_i\alpha + p_i\beta} = a^{p_i\alpha} . a^{p_i\beta} = a^\alpha . a^\beta ;$$

$$a^{\alpha\beta} = a^{p_i(\alpha\beta)} = a^{p_i\alpha . p_i\beta} = (a^{p_i\alpha})^{p_i\beta} = (a^\alpha)^\beta .$$

/заметим, что $a^{p_i\alpha} \in A_i$ /. Итак G является Ω - квазимодулем.

Теперь допустим, что квазимодули A_i унитарны. Тогда ясно, что кольцо $\prod_I \Omega_i$ имеет единичный элемент ε, i-й компонент которого есть единичный элемент ε_i кольца Ω_i.
Но тогда, для любого $a \in A_i \cap A_j$, мы имеем

$$a^{p_i \varepsilon} = a^{\varepsilon_i} = a = a^{\varepsilon_j} = a^{p_j \varepsilon} ,$$

значит $\varepsilon \in \Omega$, и тогда Ω является подкольцом кольца $\prod\limits_I \Omega_i$. Затем, при любом $a \in A_i$

$$a^{\varepsilon} = a^{p_i \varepsilon} = a^{\varepsilon_i} = a ,$$

значит, в этом случае, Ω - квазимодуль G унитарен.

Покрытие $\{A_i \mid i \in I\}$ группы G называется <u>непересекающимся</u> если $\forall i,j(i \neq j \to A_i \cap A_j = E)$, где E - единичная группа.

Непосредственно из Теоремы 2 следует.

<u>Следствие 1</u>. Если $\{A_i \mid i \in I\}$ непересекающееся покрытие группы G и для каждого $i \in I$, A_i есть квазимодуль /унитарный/ над некоторым кольцом Ω_i , то G будет квазимодулем /унитарным/ над $\prod\limits_I \Omega_i$.

Как известно, что каждую абелеву группу A можно считать модулем над Z или над $Hom\ (A,\ A)$. Следовательно, из Теоремы 2 и Следствия 1 следует

<u>Следствие 2</u>. Если $\{A_i \mid i \in I\}$ - абелево покрытие /непересекающееся/ группы G , то G будет унитарным квазимодулем над некоторым подкольцом кольца $\prod\limits_I \Omega_i$ /над $\prod\limits_I \Omega_i$ /, где Ω_i , $i \in I$, есть либо Z либо $Hom\ (A_i,\ A_i)$.

<u>Пример 4</u>. Рассмотрим абелеву группу из четырех элементов $G = \{e,a,b,c\}$, заданную следующей таблицей.

	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Видно, что циклические подгруппы второго порядка $A_1 = \{e,a\}$, $A_2 = \{e,b\}$, $A_3=\{e,c\}$ составляют непересекающееся абелево покрытие группы G . По Следствию 2 можно считать G унитарным квазимодулем, например, над Z_2^3 , где Z_2 - кольцо вычетов по модулю 2. Ясно, что этот квазимодуль не является унитарным Z - квазимодулем; он также не является модулем так как, например

$$a^{(1,1,0)}\ b^{(1,1,0)} = ab = c ,$$

в то время

$$(ab)^{(1,1,0)} = c^{(1,1,0)} = e .$$

Пример 5. Рассмотрим симметрическую группу третьей степени

$$S_3 = \{e, (12), (13), (23), (123), (132)\} .$$

Легко видеть, что циклические подгруппы второго и третьего порядка

$$A_1=\{e,(12)\},\ A_2=\{e,(13)\},\ A_3=\{e,(23)\},\ A_4=\{e,(123),(132)\}$$

составляют непересекающееся абелево покрытие группы S_3. Снова по Следствию 2 можно считать S_3 унитарным квазимодулем, например, над $Z_2^3 \times Z_3$. Видно, что этот не является унитарным Z - квазимодулем; он также не является модулем так как S_3 не коммутативна.

Замечание 2. Пусть $\varphi = \Delta \to \Omega$ гомоморфизм кольца Δ в кольцо Ω. Тогда, если G есть Ω - квазимодуль /унитарный/, то G будет Δ - квазимодулем /унитарным/, где действие элемента $\alpha\in\Delta$ на элемент $a\in G$ определяется следующим образом: $a^{\alpha} = a^{\varphi(\alpha)}$. В частности, если G есть квазимодуль /унитарный/ над кольцом Ω, то G будет также квазимодулем /унитарным/ над любым подкольцом кольца Ω.

Пусть G - группа с кольцом Ω унарных операций и пусть $\{A_i \mid i\in I\}$ - некоторое покрытие группы G. Если для каждого $i\in I$, подгруппа A_i вместе с операциями из Ω составляет Ω - квазимодуль, то мы будем называть $\{A_i \mid i\in I\}$ покрытием Ω - квазимодулей группы G. А если, A_i вместе с операциями из Ω составляет Ω - модуль, то мы имеем покрытие Ω -модулей группы G.

Из Теоремы 2 и из Замечания 2 следует

Следствие 3. Если G - группа с кольцом Ω унарных операций и $\{A_i \mid i \in I\}$ - некоторое ее покрытие Ω - квазимодулей /унитарных/, то G будет Ω - квазимодулем /унитарным/.

3. Ω - квазимодуль, порожденный одним элементом a называется циклическим и обозначается (a) .

Лемма 1. Любой унитарный циклический Ω - квазимодуль является унитарным циклическим Ω - модулем.

Доказательство. Пусть (a) есть унитарный циклический Ω - квазимодуль, порожденный элементом a . Ввиду унитарности легко видеть, что $(a) = \{a^{\alpha} \mid \alpha \in \Omega\}$. Затем при любых $\alpha, \beta, \gamma \in \Omega$ имеем

$$a^{\alpha} . a^{\beta} = a^{\alpha+\beta} = a^{\beta+\alpha} = a^{\beta} . a^{\alpha} ;$$

$$(a^{\alpha} . a^{\beta})^{\gamma} = (a^{\alpha+\beta})^{\gamma} = a^{(\alpha+\beta)\gamma} = a^{\alpha\gamma+\beta\gamma} = a^{\alpha\gamma} . a^{\beta\gamma} = (a^{\alpha})^{\gamma} . (a^{\beta})^{\gamma}$$

Итак (a) является унитарным циклическим Ω - модулем.

Теорема 3. Пусть Ω - кольцо с единицей и G - группа с кольцом Ω унарных операций. Для каждого $a \in G$ поставим $A_a = \{a^{\alpha} \mid \alpha \in \Omega\}$. Тогда G будет унитарным Ω - квазимодулем тогда и только тогда, когда $\{A_a \mid a \in G\}$ составляет покрытие унитарных циклических Ω - модулей группы G .

Доказательство. Достаточность непосредственно следует из Следствия 3. Пусть G есть унитарный Ω - квазимодуль. Тогда при каждом $a \in G$, A_a должен быть унитарным циклическим подквазимодулем, порожденным элементом a , и следовательно, по Лемме 1, A_a будет унитарным циклическим Ω - модулем. Таким образом система $\{A_a \mid a \in G\}$ составляет покрытие унитарных циклических Ω - модулей группы G .

§ 2. Характеристики некоторых классов групп в терминах квазимодулей.

1. Как известно в теории групп, бернсайдовское многообразие экспоненты n определяется тождеством

$$x^n = e .$$

следующая теорема дает одну характеристику групп из бернсайдовского многообразия экспоненты n в терминах квазимодулей.

Теорема 4. Группа G принадлежит бернсайдовскому многообразию экспоненты n тогда и только тогда, когда G является унитарным квазимодулем над кольцом Z_n вычетов по модулю n .

Доказательство. Пусть группа G принадлежит бернсайдовскому многообразию экспоненты n . Через $[r]$ обозначаем класс числа r по модулю n и для любого $a \in G$ поставим

$$a^{[r]} =_{df} a^r$$

где a^r есть r -ая степень элемента a . Очевидно для любого r

$$e^{[r]} = e .$$

Затем для любого $a \in G$ и лбых r, λ имеем

$$a^{[r]+[\lambda]} = a^{[r+\lambda]} = a^{r+\lambda} = a^r . a^\lambda = a^{[r]} . a^{[\lambda]} ;$$

$$a^{[r] \cdot [\lambda]} = a^{[r\lambda]} = a^{r\lambda} = (a^r)^\lambda = (a^{[r]})^{[\lambda]} ;$$

$$a^{[1]} = a^1 = a .$$

Таким образом G является унитарным Z_n - квазимодулем.

Обратно, пусть G унитарный Z_n - квазимодуль. Тогда для любого $a \in G$, ввиду /КВМ.3/, /КВМ.1/ и /КВМ.4/, имеем

$$a^n = \underbrace{a.a \ldots a}_{n \text{ раз}} = \underbrace{a^{[1]}.a^{[1]} \ldots a^{[1]}}_{n \text{ раз}} = a^{\overbrace{[1]+[1]+\ldots+[1]}^{}}_{\;} = a^{[n]} = e$$

n раз

Итак группа G должна принадлежать бернсайдовскому многообразию экспоненты n .

2. Пусть П некоторое множество простых чисел. Любое целое число r , все простые делители которого принадлежат П, называется _П-числом._ Группа G называется _ПD-группой,_ если для любого $a \in G$ и любого П-числа $r > 0$, уравнение

$$x^r = a \qquad (1)$$

имеет ровно одно решение. В частности, если П совпадает с множеством всех простых чисел, то ПD-группу будем называть просто _D-группой._ ПD-группы и D-группы были рассмотрены впервые Баумслагом [3] и оказалось, что они обладают многими интересными свойствами. В следующем мы сформулируем для этих групп характеристики в терминах квазимодулей.

Через $Z[\Pi^{-1}]$ обозначаем кольцо всех рациональных чисел вида $\frac{n}{r}$, где n - целое, r - положительное П-число или $r = 1$.

Теорема 5. Группа G будет ПD-группой тогда и только тогда, когда она является унитарным квазимодулем над кольцом $Z[\Pi^{-1}]$.

Доказательство. Пусть G произвольная данная ПD-группа. Для каждого $a \in G$ и каждого $\frac{n}{r} \in Z[\Pi^{-1}]$ мы обозначаем через $a^{\frac{n}{r}}$ единственный элемент группы G такой, что

$$(a^{\frac{n}{r}})^r = a^n .$$

Тогда каждый элемент из колца $Z[\Pi^{-1}]$ однозначно определяет одну унарную операцию на группе G , потому что если $\frac{n}{r} = \frac{m}{\lambda}$, т.е. $n\lambda = mr$, то $(a^{\frac{n}{r}})^{r\lambda} = a^{n\lambda} = a^{mr} = (a^{\frac{m}{\lambda}})^{r\lambda}$ и следовательно $a^{\frac{n}{r}} = a^{\frac{m}{\lambda}}$.

Теперь при любом $\frac{n}{r} \in Z[\Pi^{-1}]$ имеем $(e^{\frac{n}{r}})^r = e^n = e = e^r$, отсюда $e^{\frac{n}{r}} = e$.

Затем, при любых $a, b \in G$ и любых $\frac{n}{r}, \frac{m}{\lambda} \in Z[\Pi^{-1}]$, если $ab = ba$, то $a^{\frac{n}{r}} b^{\frac{m}{\lambda}} = b^{\frac{m}{\lambda}} a^{\frac{n}{r}}$. В самом деле, достаточно докажем для случая $\frac{m}{\lambda} = 1$. Из коммунитативности элементов a и b следует, что $(b^{-1} a^{\frac{n}{r}} b)^r = b^{-1} (a^{\frac{n}{r}})^r b = b^{-1} a^n b = a^n$. отсюда $b^{-1} a^{\frac{n}{r}} b = a^{\frac{n}{r}}$, или $a^{\frac{n}{r}} b = b\, a^{\frac{n}{r}}$.

В качестве частного случая доказанного утверждения имеем

$$a^{\frac{n}{r}} a^{\frac{m}{\lambda}} = a^{\frac{m}{\lambda}} a^{\frac{n}{r}}$$

при любых $a \in G$; $\frac{n}{r}, \frac{m}{\lambda} \in Z[\Pi^{-1}]$.

Используя последний факт получаем

$$\begin{aligned}(a^{\frac{n}{r}+\frac{m}{\lambda}})^{r\lambda} &= (a^{\frac{n\lambda+mr}{r\lambda}})^{r\lambda} = a^{n\lambda+mr} = a^{n\lambda} . a^{mr} \\ &= (a^{\frac{n\lambda}{r\lambda}})^{r\lambda} (a^{\frac{mr}{r\lambda}})^{r\lambda} = (a^{\frac{n}{r}})^{r\lambda} (a^{\frac{m}{\lambda}})^{r\lambda} \\ &= (a^{\frac{n}{r}} . a^{\frac{m}{\lambda}})^{r\lambda},\end{aligned}$$

отсюда

$$a^{\frac{n}{r}+\frac{m}{\lambda}} = a^{\frac{n}{r}} . a^{\frac{m}{\lambda}}.$$

Затем мы имеем также

$$\begin{aligned}((a^{\frac{n}{r}})^{\frac{m}{\lambda}})^{r\lambda} &= (((a^{\frac{n}{r}})^{\frac{m}{\lambda}})^{\lambda})^r = ((a^{\frac{n}{r}})^m)^r \\ &= ((a^{\frac{n}{r}})^r)^m = a^{nm},\end{aligned}$$

отсюда

$$(a^{\frac{n}{r}})^{\frac{m}{\lambda}} = a^{\frac{nm}{r\lambda}} = a^{\frac{n}{r}\cdot\frac{m}{\lambda}}$$

Итак мы доказали, что G является $Z[\Pi^{-1}]$ - квазимодулем. Унитарность этого квазимодуля очевидна.

Обратно, пусть G есть унитарный $Z[\Pi^{-1}]$ - квазимодуль. Тогда для каждого $a \in G$ и каждого П-числа $r > 0$, уравнение /1/ имеет $a^{\frac{1}{r}}$ в качестве решения потому, что ввиду /КВМ.2/ и /КВМ.3/

$$\left(a^{\frac{1}{r}}\right)^{r} = a^{\frac{1}{r}\cdot r} = a^{1} = a \quad .$$

Затем, если b - любое решение уравнения /1/, то

$$b = b^{r\cdot\frac{1}{r}} = \left(b^{r}\right)^{\frac{1}{r}} = a^{\frac{1}{r}}$$

Таким образом уравнение /1/ имеет единственное решение в группе G ; иначе говоря, G является ПD-группой.

Заметим, что когда П совпадает с множеством всех простых чисел, то $Z\,[П^{-1}]$ становится полем Q всех рациональных чисел. Таким образом, в качестве прямого следствия Теоремы 5 мы получаем следующую теорему

Теорема 6. Группа G будет D -группой тогда и толко тогда, когда она является унитарным квазимодулем над полем Q всех рациональных чисел.

Цитированная литература

1 Курош А.Г., Лекции по общей алгебре, Москва, 1962.

2 Курош А.Г., Теория групп, Москва, 1967

3 Baumsiag G., Some aspects of groups with unigue roots, Acta Math. 104 /1960/, 217-303.

S u m m a r y

Quasimodules

Do Long Van – Nguen Kuok Toan

The notion of guasimodule is introduced, by definition of which the class of all guasimodules contains, in particular, all group and modules. It is shown that there are many guasimodules which are neither groups nor modules. A necessary and sufficient condition for a group with a ring of unary operations being a guasimodule is shown. The characteristics of some well-known classes of groups in terms of guasimodules are formulated.

Ö s s z e f o g l a l ó

Kvázimodulok

Do Long Van – Nguen Kuok Toan

A cikk bevezeti a kvázimodulok fogalmát, melynek definiciója alapján a kvázimodulok osztálya tartalmazza speciálisan az összes csoportokat és modulokat. Megmutatja, hogy sok kvázimodul van amelyik sem csoport sem modul. Szükséges és elégséges feltételt ad arra, hogy egy olyan csoport amelyik rendelkezik egyváltozós operációk egy gyűrűjével kvázimodul legyen. A csoportok néhány jól ismert osztályának jellemzőit leirja kvázimodul terminólogiában.

MTA Számitástechnikai és Automatizálási Kutató Intézete, **Közlemények** 21/1978.

К В А З И М О Д У Л И, II.

До Лонг Ван, Нгуен Куок Тоан

0. Понятие квазимодуля было введено в [1]. В насточщей заметке мы быдем определять и рассматривать некоторые специальные типы подквазимодулей, которые, как увидим, играют здесь такие же роли, какие играют в теории групп соответствующие типы подгрупп.

1. Пусть дан Ω - квазимодуль G. Для любых $a, b \in G$; $\beta \in u\{1\}$ поставим

$$[a,b]_\beta =_{df} a^{-\beta} b^{-\beta} (ab)^\beta ;$$

$$[a]^b_\beta =_{df} b^{-\beta} (ab)^\beta$$

В следующем, вместо $[a, b]_1$ и $[a]^b_1$ мы тоже пишем просто $[a, b]$ и $[a]^b$ соответственно.

<u>Лемма 1.</u> Пусть G произвольный Ω - квазимодуль. Тогда для любых a, b из G и любых элементов $\beta_1, \beta_2, \ldots, \beta_n$ $(n \geq 1)$ из Ω мы имеем

$$[a]^b_{\beta_1+\beta_2+\ldots+\beta_n} = [[a]^b_{\beta_2}]^{b^{\beta_2+\ldots+\beta_n}} \cdot [[a]^b_{\beta_2}]^{b^{\beta_3+\ldots+\beta_n}} \ldots [[a]^b_{\beta_{n-1}}]^{b^{\beta_n}} \cdot [a]^b_{\beta_n}.$$

<u>Доказательство.</u> Мы докажем индукцией по n. При $n = 1$ лемма верна тривиально. При $n = 2$ имеем

$$[a]^b_{\beta_1+\beta_2} = b^{-(\beta_1+\beta_2)} (ab)^{\beta_1+\beta_2} = b^{-\beta_2} b^{-\beta_1} (ab)^{\beta_1} (ab)^{\beta_2} =$$

$$= b^{-\beta_2} b^{-\beta_1} (ab)^{\beta_1} b^{\beta_2} b^{-\beta_2} (ab)^{\beta_2} =$$

$$= b^{-\beta_2} [a]^b_{\beta_2} b^{\beta_2} [a]^b_{\beta_2} = [[a]^b_{\beta_1}]^{b^{\beta_2}} \cdot [a]^b_{\beta_2} ,$$

т.е. лемма верна и в этом случае. Теперь допустим, что $n > 2$ и лемма была доказана для $n-1$. Тогда, из верности леммы для $n = 2$ и из индуктивного предложения следует

$$[a]^b_{\beta_1+\beta_2+\ldots+\beta_n} = [a]^b_{\beta_1+(\beta_2+\ldots+\beta_n)} = [[a]^b_{\beta_1}]^{b^{\beta_2+\ldots\beta_n}} \cdot [a]^b_{\beta_2+\ldots+\beta_n} =$$

$$= [[a]^b_{\beta_1}]^{b^{\beta_2+\ldots+\beta_n}} [[a]^b_{\beta_2}]^{b^{\beta_3+\ldots+\beta_n}} \ldots [[a]^b_{\beta_{n-1}}]^{b^{\beta_n}} \cdot [a]^b_{\beta_n}$$

значит лемма справедлива для n. Лемма доказана.

<u>Лемма 2.</u> Пусть G произвольный заданный Ω - квазимодуль и пусть $\beta_1, \beta_2, \ldots, \beta_n$ $(n \geq 1)$ есть элемент из Ω такие, что каждый β_i, $i = 1,2,\ldots,n$ является эндоморфизмом или антиэндоморфизмом группы G. Тогда для любых a, b из G мы имеем

$$[a]^b_{\beta_1+\beta_2+\ldots+\beta_n} = [a^{\beta_1}]^{b^{\beta_1'+\beta_2+\ldots+\beta_n}} \cdot [a^{\beta_2}]^{b^{\beta_2'+\beta_3+\ldots+\beta_n}} \ldots [a^{\beta_{n-1}}]^{b^{\beta_{n-1}'+\beta_n}} \cdot [a^{\beta_n}]^{b^{\beta_n'}},$$

где β_2' равно β_i или 0 в зависимости от того, является ли β_i эндоморфизмом или антиэндоморфизмом группы G, $i = 1,2,\ldots,n$.

<u>Доказательство.</u> Легко видеть, что если β_i является эндоморфизмом группы G, то $[a]^b_{\beta_i} = [a^{\beta_i}]^{b^{\beta_i}}$; а если β_i - антиэндоморфизмом группы G, то $[a]^b_{\beta_i} = [a^{\beta_i}]^{b^0}$. От этого, для каждого i, $1 \leq i \leq n$, имеем

$$[[a]^b_{\beta_i}]^{b^{\beta_{i+1}+\ldots+\beta_n}} = [[a^{\beta_i}]^{b^{\beta_i'}}]^{b^{\beta_{i+1}+\ldots+\beta_n}}$$

$$= b^{-(\beta_{i+1}+\ldots+\beta_n)} b^{-\beta_i'} a^{\beta_i} b^{\beta_i'} b^{\beta_{i+1}+\ldots+\beta_n}$$

$$= b^{-(\beta_i'+\beta_{i+1}+\ldots+\beta_n)} a^{\beta_i} b^{\beta_i'+\beta_{i+1}+\ldots+\beta_n}$$

$$= [a^{\beta_i}]^{b^{\beta_i'+\beta_{i+1}+\ldots+\beta_n}}.$$

Но тогда справедливость леммы 2 сразу следует из леммы 1.

2. Пусть дан Ω - квазимодуль G. Тогда множество

$$Z(G) =_{df} \{a\in G \mid \forall b\in G\ \forall \alpha,\beta\in \Omega\cup\{1\} = [a^{\alpha},b]_{\beta} = e\}$$

мы будем называть <u>центром</u> Ω - квазимодуля G.

По определению видно, что $Z(G) \le Z^{gr}(G)$, где $Z^{gr}(G)$ обозначает центр группы G. Итак, каждый элемент из $Z(G)$ перестановочен с любым элементом из G; в частности, элементы из $Z(G)$ перестановочны между собой.

<u>Теорема 1.</u> Центр $Z(G)$ Ω - квазимодуля /унитарного/ является Ω - модулем /унитарным/ и является идеалом квазимодуля G.

<u>Доказательство.</u> Из Определения центра видно, что если $a\in Z(G)$ то $a^{\gamma}\in Z(G)$ при любом $\gamma\in\Omega$.

Пусть $a, c\in Z(G)$. Тогда при любых $b\in G$; $\alpha,\beta\in\Omega\cup\{1\}$ мы имеем

$$((ac)^{\alpha}b)^{\beta} = (c^{\alpha}a^{\alpha}b)^{\beta} = (c^{\alpha}(a^{\alpha}b))^{\beta} = (a^{\alpha}b)^{\beta}c^{\alpha\beta} =$$
$$= b^{\beta}a^{\alpha\beta}c^{\alpha\beta} = b^{\beta}(c^{\alpha\beta}a^{\alpha\beta}) = b^{\beta}(ac)^{\alpha\beta},$$

т.е. $[(ac)^{\alpha}, b]_{\beta} = e$, значит $ac\in Z(G)$. Здесь и в дальнейшем мы условимся $\alpha\ .\ 1 = 1\ .\ \alpha = \alpha$ при всяком $\alpha\in\Omega$.

Затем, если $a\in Z(G)$, то для любого $\alpha\in\Omega\cup\{1\}$, $(a^{-1})^{\alpha}a^{\alpha} = (aa^{-1})^{\alpha} = e^{\alpha} = e$, т.е. $(a^{-1})^{\alpha} = a^{-\alpha}$. Но тогда для любых $b\in G$, $\alpha\in\Omega$, $\beta\in\Omega\cup\{1\}$, $[(a^{-1})^{\alpha}, b]_{\beta} = [a^{-\alpha}, b]_{\beta} = e$. Потом, по определению центра, для любых $b\in G$, $\beta\in\Omega\cup\{1\}$ имеем

$$e = [a, a^{-1}b]_{\beta} = a^{-\beta}(a^{-1}b)^{-\beta}b^{\beta},\ \text{т.е.}\ b^{-\beta}(a^{-1}b)^{\beta} = a^{-\beta}$$

Но тогда $[a^{-1}, b]_{\beta} = (a^{-1})^{-\beta}b^{-\beta}(a^{-1}b)^{\beta} = a^{\beta}b^{-\beta}(a^{-1}b)^{\beta} = a^{\beta}a^{-\beta} = e$. Итак мы доказали, что $[(a^{-1})^{\alpha},b]_{\beta} = e$ при любых $b\in G$,

$\alpha, \beta \in \Omega \cup \{1\}$, значит $a^{-1} \in Z(G)$.

Так как элементы из $Z(G)$ перестановочны между собой то для любых $a, c \in Z(G)$, $\alpha \in \Omega$ имеем $(ac)^\alpha = c^\alpha a^\alpha = a^\alpha c^\alpha$. Таким образом $Z(G)$ является подквазимодулем квазимодуля G и Ω - модулем. Конечно, если G унитарен, то $Z(G)$ также унитарен.

Наконец, так как для любых $a \in Z(G)$, $b \in G$, $\beta \in \Omega \cup \{1\}$, имеем $[a,b]_\beta = e$, т.е. $b^{-\beta}(ab)^\beta = a^\beta \in Z(G)$ то $Z(G)$ является идеалом квазимодуля G. Теорема доказана.

Обобщая понятия центра квазимодуля мы приходим к понятию централизатора. Пусть даны два подквазимодуля A, B Ω - квазимодуля G. Тогда

$$Z_A(B) =_{df} \{a \in A \mid \forall b \in B\ \forall \alpha, \beta \in \Omega \cup \{1\} = [a^\alpha, b]_\beta = e\}$$

называется <u>централизатором</u> подквазимодуля B в подквазимодуле A.

Очевидно, что $Z_A(B) \leq Z_A^{gr}(B)$, где $Z_A^{gr}(B)$ есть централизатор подгруппы B в подгруппе A. Для рассматривания $Z_A(B)$ необходима следующая лемма.

<u>Лемма 3.</u> Если Ω -квазимодуль G удовлетворяет условию

(1) Аддитивная группа кольца Ω обладает системой Σ обозначающих, которая состоит из эндоморфизмов группы G,

то для любых $a \in G$, $\alpha \in \Omega$ имеем

$$(a^{-1})^\alpha = a^{-\alpha}$$

<u>Доказательство.</u> Пусть $a \in G$, $\alpha \in \Omega$. Ввиду условия (1), α должен представиться в виде $\alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_n$ при $\alpha_i \in \pm \Sigma$, $i = 1, 2, \ldots, n$. Где $+\Sigma = \Sigma$, $-\Sigma = \{-\delta \mid \delta \in \Sigma\}$. Но тогда имеем

$$(a^{-1})^{\alpha}=(a^{-1})^{\alpha_1+\alpha_2+\ldots+\alpha_n} = (a^{-1})^{\alpha_1} (a^{-1})^{\alpha_2}\ldots (a^{-1})^{\alpha_n} =$$

$$= (a^{\alpha_1})^{-1} (a^{\alpha_2})^{-1}\ldots (a^{\alpha_n})^{-1} = a^{-\alpha_1} a^{-\alpha_2}\ldots a^{-\alpha_n} =$$

$$= a^{-(\alpha_1+\alpha_2+\ldots+\alpha_n)} = a^{-\alpha} ,$$

что и требовалось доказать.

Теорема 2. Если Ω - квазимодуль G удовлетворяет (1), а A, B - два произвольных подквазимодуля квазимодуля G, то $Z_A(B)$ будет подквазимодулем квазимодуля G.

Доказательство. По определению видно, что из $a\in Z_A(B)$ следует $a^{\gamma}\in Z_A(B)$ при любом $\gamma\in\Omega$.

Пусть $a,\ c\in Z_A(B),\ b\in B,\ \gamma\in\Omega\cup\{1\}$. Ввиду (1), $\gamma = \gamma_1+\gamma_2+\ldots+\gamma_n$, где $\gamma_i\in \pm\Sigma\cup\{1\}$. Мы имеем

$$[(ac)^{\gamma}]^{b} = b^{-1}(ac)^{\gamma}b = b^{-1}(ac)^{\gamma_1+\gamma_2+\ldots+\gamma_n} b =$$

$$= b^{-1}(ac)^{\gamma_1} (ac)^{\gamma_2}\ldots (ac)^{\gamma_n} b .$$

Заметим, что при $\gamma_i\in\Sigma\cup\{1\}$, то $(ac)^{\gamma_i} = a^{\gamma_i}c^{\gamma_i}$; при $(ac)^{\gamma_i}=c^{\gamma_i}a^{\gamma_i}$, то $(ac)^{\gamma_i} =c^{\gamma_i}a^{\gamma_i}$; а элементы $a^{\gamma_i},c^{\gamma_j} (i,j=1,\ldots,n)$ перестановочны между собой и с b. Таким образом

$$[(ac)^{\gamma}]^{b} = (ac)^{\gamma} .$$

Пусть теперь $\alpha,\beta\in\Omega\cup\{1\}$, $\beta=\beta_1+\beta_2+\ldots+\beta_n$ при $\beta_i\in \pm\Sigma\cup\{1\}$. Из только что доказанного и леммы 2 следует

$$[(ac)^{\alpha}]^{b}_{\beta} = [(ac)^{\alpha\beta_1}]^{b^{\beta_1'+\beta_2+\ldots+\beta_n}}\cdot[(ac)^{\alpha\beta_2}]^{b^{\beta_2'+\beta_3+\ldots+\beta_n}}\ldots[(ac)^{\alpha\beta_n}]^{b^{\beta_n'}} =$$

$$= (ac)^{\alpha\beta_1}(ac)^{\alpha\beta_2}\ldots(ac)^{\alpha\beta_n} = (ac)^{\alpha\beta_1+\alpha\beta_2+\ldots+\alpha\beta_n} =$$

$$= (ac)^{\alpha(\beta_1+\beta_2+\ldots+\beta_n)} = (ac)^{\alpha\beta} ,$$

т.е. $[(ac)^{\alpha},b]_{\beta} = e$, значит $ac\in Z_A(B)$.

Пусть $a\in Z_A(B)$, $b\in B$. Прежде всего при $\alpha\in\Omega$, $\beta\in\Omega\cup\{1\}$ имеем по лемме 3

$$[(a^{-1})^{\alpha},\ b]_{\beta} = [a^{-\alpha},\ b]_{\beta} = e\ .$$

Затем, если $\beta \in \Omega$, $\beta = \beta_1+\beta_2+\ldots+\beta_n$, при $\beta_i \in \pm\Sigma$, то, по лемме 2, имеем

$$[a^{-1}]^{b}_{\beta} = [a^{-\beta_1}]^{b^{\beta'_1+\beta_2+\ldots+\beta_n}} [a^{-\beta_2}]^{b^{\beta'_2+\beta_3+\ldots+\beta_n}} \ldots [a^{-\beta_n}]^{b^{\beta'_n}} =$$

$$= a^{-\beta_1} a^{-\beta_2} \ldots a^{-\beta_n} = a^{-(\beta_1+\beta_2+\ldots+\beta_n)} = a^{-\beta}\ ,$$

т.е. $[a^{-1},\ b]_{\beta} = e$. Наконец, видно, что $[a^{-1}, b] = ab^{-1}a^{-1}b = e$. Итак, мы доказали, что $[(a^{-1})^{\alpha}, b]_{\beta} = e$ при любых $b \in B$; $\alpha,\ \beta \in \Omega \cup \{1\}$. Но это значит $a^{-1} \in Z_A(B)$.

Таким образом $Z_A(B)$ является подквазимодулем квазимодуля G . Теорема доказана.

Отношение между $Z_A(B)$ и $Z_A^{gr}(B)$ устанавливается следующей теоремой.

<u>Теорема 3.</u> Если Ω - квазимодуль G удовлетворяет (1) , A , B - два произвольных подквазимодуля квазимодуля G , то

$$\forall a \in G\ (a \in Z_A(B) \leftrightarrow \forall \alpha \in \Sigma \cup \{1\} = a^{\alpha} \in Z_A^{gr}(B))\ .$$

<u>Доказательство.</u> Видно, что если $a \in Z_A(B)$, то $\forall \alpha \in \Sigma \cup \{1\}$: $a^{\alpha} \in Z_A^{gr}(B)$. Обратно, пусть имеет место последнее . Тогда, ввиду (1) , легко видеть, что $\forall \alpha \in \Omega \cup \{1\} = a^{\alpha} \in Z_A^{gr}(B)$, отсюда, в частности, имеем $[a^{\alpha}]^{b} = a^{\alpha}$ для любых $b \in B$, $\alpha \in \Omega \cup \{1\}$. Теперь пусть $\beta \in \Omega$, $\beta = \beta_1+\beta_2+\ldots+\beta_n$ при $\beta_i \in \pm\Sigma$, $b \in B$. Тогда по лемме 2 и только что сказанному

$$[a^{\alpha}]_{\beta}^{b} = [a^{\alpha\beta_1}]^{b^{\beta'_1+\beta_2+\ldots+\beta_n}}[a^{\alpha\beta_2}]^{b^{\beta'_2+\beta_3+\ldots+\beta_n}}\ldots[a^{\alpha\beta_n}]^{b^{\beta'_n}} =$$

$$= a^{\alpha\beta_1}a^{\alpha\beta_2}\ldots a^{\alpha\beta_n}=a^{\alpha\beta_1+\alpha\beta_2+\ldots+\alpha\beta_n} = a^{\alpha(\beta_1+\beta_2+\ldots+\beta_n)}=a^{\alpha\beta}.$$

Итак мы доказали, что $[a^{\alpha},b]_{\beta} = e$ при любых $b\in B$, $\alpha,\beta\in\Omega \cup \{1\}$. Но это значит $a\in Z_A(B)$. Теорема доказана.

Так как $Z(G) = Z_G(G)$, из теоремы 3 следует

<u>Следствие 1</u>. Если Ω - квазимодуль G удовлетворяет (1) , то

$$\forall a\in G\ (a\in Z(G) \leftrightarrow \forall\alpha\in\Sigma\cup\{1\} : a^{\alpha}\in Z^{gr}(G)) .$$

Как увидим в дальнейшем, вообще $Z_A(B)\neq Z_A^{gr}(B)$, даже в том случае, когда G удовлетворяет (1) . Следующее следствие дает одно достаточное уловие для $Z_A(B) = Z_A^{gr}(B)$.

<u>Следствие 2</u>. Если Ω - квазимодуль G удовлетворяет (1) и

(2) $\forall a,b\in G\ \forall\alpha\in\Omega = b^{-1}a^{\alpha}b = (b^{-1}a\ b)^{\alpha}$,

то для любых подквазимодулей A, B имеем $Z_A(B) = Z_A^{gr}(B)$. В частности, мы всегда имеем $Z(G) = Z^{gr}(G)$.

<u>Доказательство</u>. Достаточно показать, что $Z_A^{gr}(B) \leq Z_A(B)$. Пусть $a\in Z_A^{gr}(B)$. Тогда для любого $b\in B$

$$b^{-1}\ a\ b = a .$$

Отсюда, ввиду (2) , для любого $\alpha\in\Omega\cup\{1\}$

$$b^{-1}\ a^{\alpha}\ b = (b^{-1}\ a\ b)^{\alpha} = a^{\alpha} ,$$

значит $a^{\alpha}\in Z_A^{gr}(B)$, и следовательно, ввиду теоремы 3, $a\in Z_A(B)$, что и требовалось доказать.

3. Пусть A, B - два подквазимодуля Ω- квазимодуля G . Тогда

$$N_A(B)=_{df}\{a\in A\,|\,\forall b\in B\ \forall\alpha\in\Omega\cup\{\pm 1\}\ \forall\beta\in\Omega\cup\{1\}:[b]_{\beta}^{a^{\alpha}}\in B\}$$

называется нормализатором подквазимодуля B в подквазимодуле A. В частности, $N_G(B)$ называется нормализатором подквазимодуля B и обозначается $N(B)$.

Очевидно, что $N_A(B) \le N_A^{gr}(B)$, где $N_A^{gr}(B)$ - нормализатор подгруппы B в подготовке A.

Теорема 4. Если Ω - квазимодуль G удовлетворяет (1); A, B - два произвольных подквазимодуля, то $N_A(B)$ будет подквазимодулем квазимодуля G.

Доказательство. Прежде всего видно, что если $a \in N_A(B)$, то $a^{\gamma} \in N_A(B)$ при любом $\gamma \in \Omega$.

Затем, ввиду леммы 3, если $a \in N_A(B)$, то для любых $b \in B$, $\alpha \in \Omega u\{\pm 1\}$, $\beta \in \Omega u\{1\}$ имеем

$$[b]_{\beta}^{(a^{-1})^{\alpha}} = [b]_{\beta}^{a^{-\alpha}} \in B,$$

значит $a^{-1} \in N_A(B)$.

Пусть теперь $a, c \in N_A(B)$. Тогда при любом $b \in B$, если $\gamma \in \sum u\{1\}$, то

$$[b]^{(ac)^{\gamma}} = (ac)^{-\gamma} b (ac)^{\gamma} = c^{-\gamma} a^{-\gamma} b a^{\gamma} c^{\gamma} = [[b]^{a^{\gamma}}]^{c^{\gamma}} \in B;$$

а если $\gamma \in -\sum u\{-1\}$, то

$$[b]^{(ac)^{\gamma}} = (ac)^{-\gamma} b (ac)^{\gamma} = a^{-\gamma} c^{-\gamma} b c^{\gamma} a^{\gamma} = [[b]^{c^{\gamma}}]^{a^{\gamma}} \in B.$$

Затем, если $\alpha = \alpha_1 + \alpha_2 + \ldots + \alpha_n$, $\alpha_i \in \pm \sum u\{\pm 1\}$, то, ввиду только что сказанного, имеем при любом $b \in B$

$$[b]^{(ac)^{\alpha}} = (ac)^{-(\alpha_1+\alpha_2+\ldots+\alpha_n)} b (ac)^{\alpha_1+\alpha_2+\ldots+\alpha_n}$$
$$= (ac)^{-\alpha_1} (ac)^{-\alpha_2} \ldots (ac)^{-\alpha_n} b (ac)^{\alpha_1} (ac)^{\alpha_2} \ldots (ac)^{\alpha_n}$$
$$= [\ldots] [b]^{(ac)^{\alpha_1}} [^{(ac)^{\alpha_2}} \ldots]^{(ac)^{\alpha_n}} \in B.$$

Наконец, если $\beta = \beta_1 + \beta_2 + \ldots + \beta_n$, $\beta_i \in \pm \sum u\{1\}$, то, из леммы 2 и только что доказанного следует, что при любых $b \in B$, $\alpha \in \Omega u\{\pm 1\}$

$$[b]_{\beta}^{(ac)^{\alpha}} = [b^{\beta_1}]^{(ac)^{\alpha(\beta_1'+\beta_2+\ldots+\beta_n)}} \cdot [b^{\beta_2}]^{(ac)^{\alpha(\beta_2'+\beta_3+\ldots+\beta_n)}} \ldots [b^{\beta_n}]^{(ac)^{\alpha\beta_n'}} \in B \ .$$

Но это значит $ac \in N_A(B)$. Теорема доказана.

<u>Теорема 5</u>. Если Ω - квазимодуль G удовлетворяет (1) , B - произвольный подквазимодуль, то $N(B)$ будет наибольшим подквазимодулем квазимодуля G , принимающим B в качестве идеала.

<u>Доказательство</u>. По теореме 4, $N(B)$ является подквазимодулем квазимодуля G . Затем, при любых $b\in B$, $a\in N(B)$, $\beta\in\Omega\cup\{1\}$ имеем

$$a^{-\beta}\,(ba)^{\beta} = [b]_{\beta}^{a} \in B \ .$$

Но это значит, что B является идеалом $N(B)$.

Пусть теперь A - произвольный подквазимодуль квазимодуля G , принимающий B в качестве идеала. Тогда, если $a\in A$, то, при любых $b\in B$, $\alpha\in\Omega\cup\{\pm 1\}$, $\beta\in\Omega\cup\{1\}$, имеем

$$[b]_{\beta}^{a^{\alpha}} = (a^{\alpha})^{-\beta}\,(ba^{\alpha})^{\beta} \in B \ ,$$

значит $a\in N(B)$. Итак $A \le N(B)$. Теорема доказана.

<u>Лемма 4.</u> Пусть Ω - квазимодуль G удовлетворяет (1) , и A - произвольный подквазимодуль G . Тогда A будет идеалом квазимодуля G тогда и только тогда, когда он является нормальным делителем группы G .

<u>Доказательство</u>. Необходимость следует из определения идеала. Докажем достатосность. Пусть A - нормальный делитель группы. Тогда, при любых $a\in A$, $b\in G$ имеем $[a]^{b} \in A$. Затем, если $\beta = \beta_1+\beta_2+\ldots+\beta_n$, $\beta_i \in \pm\Sigma\cup\{1\}$, то, из леммы 2 и только что сказанного следует, что для любых $a\in A$, $b\in G$

$$[a]_{\beta}^{b} = [a^{\beta_1}]^{b^{\beta_1'+\beta_2+\ldots+\beta_n}}\,[a^{\beta_2}]^{b^{\beta_2'+\beta_3+\ldots+\beta_n}} \ldots [a^{\beta_n}]^{b^{\beta_n'}} \in A \ ,$$

Значит A является идеалом квазимодуля G, что и требовалось доказать.

Теорема 6. Если Ω - квазимодуль G удовлетворяет (1), а A, B - два произвольных подквазимодуля квазимодуля G, то $Z_A(B)$ будет идеалом Ω - квазимодуля $N_A(B)$.

Доказательство. По теоремам 2 и 4, $Z_A(B)$ и $N_A(B)$ являются подквазимодулями квазимодуля G. Пусть $a \in Z_A(B)$. Тогда, при любых $b \in B$, $\alpha \in \Omega \cup \{\pm 1\}$, $\beta \in \Omega \cup \{1\}$ имеем

$$b^{-\beta}[b]_{\beta}^{a^{\alpha}} = b^{-\beta}a^{-\alpha\beta}(ba^{\alpha})^{\beta} = a^{-\alpha\beta}b^{-\beta}(a^{\alpha}b)^{\beta} = [a^{\alpha}, b]_{\beta} = e ,$$

т.е. $[b]_{\beta}^{a^{\alpha}} \in B$. Но это значит $a \in N_A(B)$, и следовательно $Z_A(B)$ является подквазимодулем квазимодуля $N_A(B)$.

Так как свойства (1) наследственно, то $N_A(B)$ также удовлетворяет (1). Но тогда для показания того, что $Z_A(B)$ является идеалом $N_A(B)$, ввиду леммы 4, только надо показать, что $Z_A(B)$ является нормальным делителем группы $N_A(B)$, т.е. показать, что для любых $a \in Z_A(B)$, $c \in N_A(B)$ имеет место $c^{-1}ac \in Z_A(B)$. Для этого, по теореме 3, достаточно показать, что

$$\forall \alpha \in \Sigma \cup \{1\} : (c^{-1}ac)^{\alpha} \in Z_A^{gr}(B) .$$

Но последний факт верен потому, что при любом $b \in B$ имеем

$$b^{-1}(c^{-1}ac)^{\alpha}b = b^{-1}c^{-\alpha}a^{\alpha}c^{\alpha}b = (c^{\alpha}b)^{-1}a^{\alpha}(c^{\alpha}b) =$$

$$= (b_1c^{\alpha})^{-1}a^{\alpha}(b_1c^{\alpha}) = c^{-\alpha}b_1^{-1}a^{\alpha}b_1c^{\alpha} =$$

$$= c^{-\alpha}a^{\alpha}c^{\alpha} = (c^{-1}ac)^{\alpha} ,$$

где b_1 - надлежащий элемент из B. Теорема доказана.

4. Пусть G произвольный Ω - квазимодуль и A, B - два его подквазимодуля. По определению взаимным коммутантом подква-

зимодулей A и B , обозначаемым $[A, B]$, мы будем называть идеал подквазимодуля (A, B) , порожденный множеством всех элементов вида $[a,b]_\alpha$, где $a\in A$, $b\in B$, $\alpha\in\Omega\cup\{1\}$. Здесь и в дальнейшем квазимодуль, порожденный множеством M обозначаем через (M) .

<u>Теорема 7</u>. Если Ω - квазимодуль G удовлетворяет (1) и (2); а A, B - два произвольных идеала G , то

$$H = ([a, b]_\alpha \mid a\in A, b\in B, \alpha\in\Omega\cup\{1\})$$

будет идеалом квазимодулей G.

<u>Доказательство</u>. Прежде всего при любых $a, b, c\in G$; $\alpha\in\Omega\cup\{1\}$ имеем

$$c^{-1}[a, b]_\alpha c = [c^{-1}ac, c^{-1}bc]_\alpha \quad .$$

В самом деле, используя (2) получаем

$$c^{-1}[a, b]_\alpha c = c^{-1}a^{-\alpha}b^{-\alpha}(ab)^\alpha c = c^{-1}a^{-\alpha}c\ c^{-1}b^{-\alpha}c\ c^{-1}(ab)^\alpha c =$$

$$= (c^{-1}ac)^{-\alpha}\ (c^{-1}bc)^{-\alpha}\ (c^{-1}abc)^\alpha =$$

$$= (c^{-1}ac)^{-\alpha}\ (c^{-1}bc)^{-\alpha}\ (c^{-1}ac\ c^{-1}bc)^\alpha = [c^{-1}ac, c^{-1}bc]_\alpha$$

Теперь, из только что доказанного ясно, что при любом образующем h из H имеем $c^{-1}hc \in H$ для всякого c из G . Затем, если h_1, h_2 - два произвольных элемента из H таких, что $c^{-1}h_ic \in H$, $i = 1, 2$, то имеем

$$c^{-1}h_1^{-1}c = (c^{-1}h_1c)^{-1} \in H \quad ;$$

$$c^{-1}h_1h_2c = (c^{-1}h_1c)(c^{-1}h_2c) \in H \quad ;$$

$$c^{-1}h_1^\alpha c = (c^{-1}h_1c)^\alpha \in H$$

где α - любой элемент из Ω . Итак мы доказали, что для любых $h\in H$, $c\in G$, $c^{-1}hc\in H$. Но это значит, по лемме 4, что H должен быть идеалом квазимодуля G , что и требовалось доказать.

Следствие 3. Если Ω - квазимодуль G удовлетворяет (1) и (2), а A, B - два произвольных идеала квазимодуля G, то

$$[A, B] = ([a, b]_\alpha \mid a\in A,\ b\in B,\ \alpha\in\Omega\cup\{1\}) .$$

В частности, когда $A = B = G$, мы имеем

$$[G, G] = ([a, b]_\alpha \mid a, b\in G;\ \alpha\in\Omega\cup\{1\}) .$$

Замечание. Для унитарных Z - квазимодулей, понятия подквазимодулей и идеала квазимодуля соответственно сводятся к понятиям подгруппы и нормального делителя группы; понятия центра, централизатора, нормализатора, взаимного коммутанта - к одноименным понятиям в теории групп; а условия (1) и (2) выполнены автоматически. Таким образом, применяя выше сформулированные утверждения к с лучаю унитарных Z - квазимодулей, мы получаем известные факты в теории групп. Например, для теоремы 1 имеем: центр любой группы G будет абелевой группой и будет нормальным делителем группы G; Для теоремы 5 имеем : нормализатор произвольной подгруппы B группы G будет наибольшей подгруппой группы G, принимающей B в качестве нормального делителя; или для теоремы 6 имеем : если A, B - две произвольные подгруппы группы G, то централизатор подгруппы B в подгруппе A будет нормальным делителем нормализатора подгуппы B в подгруппе A.

5. В утверждениях, сформулированных в этой заметке мы часто наложили на квазимодулях ограничения (1) и (2). Как мы уже отметили выше, унитарные Z - квазимодули обладают свойствами (1) и (2). Этими свойствами, очевидно, обладают и модули. В следующем примере покажем, что существуют квазимодули, обладающие свойствами (1) и (2), но не являющиеся ни унитарными Z - квазимодулями, ни модулями. Затем, так как свойства (1) и (2) определяются тождествами, то они инвариантны относительно операций взятия подквазимодулей, взятия гомоморфных образов и взятия декартова произведения. Это показывает, что класс квазимодулей со свойствами (1) и (2) довольно

широк, и следовательно, утверждения, сформулированные здесь достаточно сильны.

Пример 1. Рассмотрим мультипликативную группу G всех невырожденных матриц второго порядка с рациональными элементами. Если a – такая матрица, то ее определитель будет отличным от нуля рациональным числом и поэтому может быть написан в виде $\frac{\lambda}{t}\, 2^{n(a)}$, где числа λ, t нечетны, а число $n(a)$ целое. Из того, что определить произведения матриц равен произведению определителей, вытекает, что

$$n(ab) = n(a) + n(b)$$

для любых $a, b \in G$.

Определяем отображение $\alpha = G \to G$ следующим образом

$$a^{\alpha} =_{df} \begin{pmatrix} 1 & n(a) \\ 0 & 1 \end{pmatrix} .$$

Равенства

$$(ab)^{\alpha} = \begin{pmatrix} 1 & n(ab) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n(a)+n(b) \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & n(a) \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & n(b) \\ 0 & 1 \end{pmatrix} = a^{\alpha}.b^{\beta}$$

показывает, что α будет эндоморфизмом группы G.

Видно, что

$$z^{gr}(G) = \left\{ \begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \;\middle|\; r \text{ – отличное от нуля рациональное число} \right\}$$

Теперь, если на множестве

$$\Omega = \{(\alpha, n) \mid n \text{ целое}\}$$

определяем операции сложения и умножения следующим образом :

$$(\alpha, n) + (\alpha, m) = (\alpha, n+m) \quad ;$$

$$(\alpha, n) . (\alpha, m) = (\alpha, 0) ,$$

то легко проверить, что Ω будет коммутативным кольцом, аддитивная группа которого порождается элементом $(\alpha, 1)$.

Затем, для любых $a \in G$, $(\alpha, n) \in \Omega$ положим

$$a^{(\alpha, n)} = (a^{\alpha})^{n}$$

и заметим, что $a^{\alpha^2} = e$, где e - единичная матрица, то легко видеть, что G будет Ω - квазимодулем.

Так как

$$a^{(\alpha, 1)} = a^{\alpha},$$

то Ω - квазимодуль G удовлетворяет (1) при $\sum = \{(\alpha, 1)\}$.

По следствию 1, $a \in Z(G)$ тогда и только тогда, когда $a \in Z^{gr}(G)$ и $a^{(\alpha,1)} \in Z^{gr}(G)$, т.е. тогда и только тогда, когда $a \in Z^{gr}(G)$ и $n(a) = 0$. Таким образом имеем

$$Z(G) = \left\{ \begin{pmatrix} \frac{\lambda}{t} & 0 \\ 0 & \frac{\lambda}{t} \end{pmatrix} \;\middle|\; \lambda,\ t \text{ - нечетные числа} \right\},$$

значит $Z(G) \neq Z^{gr}(G)$, и следовательно, по следствию 2, Ω - квазимодуль G не удовлетворяет (2).

Одним словом, построенный нами Ω - квазимодуль G удовлетворяет (1), не удовлетворяет (2), не является ни унитарным Z - квазимодулем, ни модулем.

<u>Пример 2.</u> Пусть A, B - две произвольные некоммутативные группы. Положим

$$G = A \times B.$$

Затем, на множестве

$$\Omega \quad \{(p, g) \mid p, g \text{ - целые}\}$$

определяем операции сложения и умножения следующим образом:

$$(p,g) + (r, \lambda) = (p + r,\ g + \lambda)\ ,$$

$$(p,g) \,.\, (r, \lambda) = (pr,\ gr + p\lambda + g\lambda)\ .$$

Легко проверить, что Ω будет кольцом с единицей $(1, 0)$, а его аддитивная группа имеет систему образующих $\sum = \{(1,0),(0,1)\}$

Если для любых $(a,b)\in G$, $(p,g)\in\Omega$ положим

$$(a,b)^{(p,g)} = (a^{p+g},\ b^{p})\ ,$$

то, прямой проверкой, легко видеть, что будет унитарным Ω - квазимодулем.

Из

$$(a,\ b)^{(1,0)} = (a,\ b)$$

$$(a,\ b)^{(0,1)} = (a,\ e)$$

следует, что Ω - квазимодуль G удовлетворяет (1) .

Затем имеем

$$(c,d)^{-1}(a,b)^{(p,g)}(c,d) = (c^{-1}a^{p+g}c,\ d^{-1}b^{p}d)=((c^{-1}ac\ ^{p+g},\ d^{-1}bd)^{p})=$$

$$=(c^{-1}ac,d^{-1}bd)^{(p,g)}=((c,d)^{-1}(a,b)(c,d)^{(p,g)},$$

значит Ω - квазимодуль G также удовлетворяет (2) .

Одним словом, Ω - квазимодуль G удовлетворяет (1), (2) , не является ни унитарным Z - квазимодулем, ни модулем.

Цитированная литература

|1| До Лонг Ван, Нгуен Куок Тоан. Квазимодули 1.

Ö s s z e f o g l a l ó

Kvázimodulok II.

Do Long Van – Nguen Kuok Toan

A szerző az előző cikkében definiált kvázimodulok tulajdonságait vizsgálja. Bevezeti a rész kvázi modulok speciális osztályait, és megmutatja, hogy a kvázimodulok elméletében azok ugyanazt a szerepet töltik be, mint a csoportelméletben a megfelelő részcsoportosztályok.

S u m m a r y

Quasimodules II.

Do Long Van – Nguen Kuok Toan

The author deals with properties of quasimodules defined in a previous paper. He introduces special classes of subquasimodules and demonstrates their roles in the theory of quasimodules being the same as those of the corrresponding subgroups in group theory.

SZÁMOK Repró. 72208203/203. A/4. 430 pld.