Запропоновано формування граматичної структури і мовних конструкцій через представлення знань, що розширює можливості формальних граматичних структур у системах штучного інтелекту

Ключові слова: граматична структура, структура представлення знань, гібридна структура, силогізм, граматичний виведення, формальна мова

Предложено формирование грамматической структуры и языковых конструкций через представление знаний, что расширяет возможности формальных грамматических структур в системах искусственного интеллекта

Ключевые слова: грамматическая структура, структура представления знаний, гибридная структура, силлогизм, грамматический вывод, формальный язык

УДК 004.89:007.001.33

ГРАММАТИЧЕСКИЕ СТРУКТУРЫ С ЛОГИЧЕСКИМ ВЫВОДОМ

В.И. Шинкаренко
Доктор технических наук, профессор, доцент*
Контактный тел.: 063-489-49-15
E-mail: Shinkarenko_vi@ua.fm

В.А. Андрющенко
Кандидат технических наук, доцент*
Контактный тел.: 097-380-87-45
E-mail: andr17102@gmail.com

В.М. Ильман
Кандидат физико-математических наук, доцент*
Контактный тел.: 067-895-18-38

*Кафедра компьютерных информационных технологий
Днепропетровский национальный университет железнодорожного транспорта им. Академика В. Лазаряна
ул. Акад. В. Лазаряна, 2, г. Днепропетровск, Украина, 49000

1. Введение

Формальный грамматический подход и подход на основе логических высказываний относятся к универсальным средствам моделирования информационных систем (ИС). Эти подходы используются при проектировании и представлении языков программирования, алгоритмов, автоматов, системах распознавания образов, проектированиях трансляторов, баз данных и знаний и прочих прикладных систем искусственного интеллекта.

Определенный вклад в развитие теории формальных грамматик внесли Скалозуб В. В., Шинкаренко В. И., Ильман В. М., предложив модель конструктивной грамматической структуры [1]. Грамматическая структура строится на основе формальных грамматик и алгебры операций и понимается как система для формального построения языковых конструкций, т.е. структура рассматривается как грамматическая модель конструирования абстрактных объектов предметной области. Формальная грамматическая система задается тройкой, состоящей из носителя, сигнатуры и конструктивной аксиоматики. Носитель грамматической системы может быть неоднородным и состоять из наборов структур порождающих его элементы. Сигнатура системы также неоднородна и включает различные операции линейных, векторных, матричных и других подстановок, операторы и отношения. В конструктивной аксиоматике определяются правила применения операций сигнатуры над элементами носителя и конструкций, созданных на них, свойства операций, порядок применения правил, инструктивные правила и др. Аксиоматика в системах программной инженерии дополняются семантическими правилами со своими специфическими представлениями и идеологиями реализации.

Классические формальные грамматики и соответствующие грамматические структуры обычно решают прямую задачу выделения части (порожденного языка) из некоторого свободного языка. Процесс порождения языка в грамматике основывается на операциях подстановки, гомоморфизма и прочих. Логические выводы на высказываниях в системах искусственного интеллекта [2, 3] также могут служить инструментами для выделения формальных языков.

Однако, в прикладных системах, в программной инженерии, при формировании языка запросов и других ИС приходится сталкиваться с объектами, для конструирования которых необходимо использовать гибридный подход на основе грамматик и логических выводов. Так объект программирования – абстрактный тип данных (АТД) [4], который определяется математической моделью со встроенными операциями над данными удобно представлять гибридной конструктивной структурой-моделью. Обычно, в прикладном программировании в роли такой «математической» модели выступает некоторая схема-формат [5]. Например, мультимножество beg с операцией приведения к обычному множеству set (удаления одинаковых элементов множества beg) может быть создано по неформальной схеме.

АТД Мультимножество.

1. Данные.

Количество различных элементов множества, кратность элемента: целое число.

Множество типа beg.

Множество типа set.

Элемент множества типа char.

2. Операции.

Инициализация	
Начальное значение:	Присвоение количеству различных элементов множества и кратностям элементов конкретных значений.
Процесс:	Формирование мультимножества M.
Операция 1	
Вход:	Элемент b типа char.
Процесс:	Проверка на принадлежность элемента b множеству M.
Постусловие:	Нет.
Выход:	Возвратить сообщение «Элемент $b \in M$».
Операция 2	
Вход:	Присвоить элементу b начальное значение.
Процесс:	Формирование множества M_1 типа set из мультимножества M.
Постусловие:	Изменить b.
Выход:	Возвратить M_1.
И т.д.	
Конец АТД.	

Анализ АТД-схемы показывает, что формирование мультимножества и раздела операций могут быть заданы формальной моделью на основе гибридной структуры с подстановками и импликациями (формальная модель АТД Мультимножество приводится ниже после введения соответствующего инструментария).

Поэтому, учитывая актуальность конструктивного моделирования в приложениях ИС, в работе рассмотрено развитие формальных грамматических структур с учетом элементов трехзначной логики. Целью работы является разработка гибридной грамматической структуры, которая позволяет обобщить и расширить возможности как грамматических структур, так и интеллектуальных систем.

Формальная система строится в два этапа. На первом этапе конструктивно вводится формальная структура знаний (ФСЗ), как элемент носителя грамматической структуры и на втором – формальная грамматическая система с логическим выводом (ФСЛВ). ФСЗ формируется на основе продукционного подхода представления знаний, как предметная формальная конструктивная система. Система строится на словаре понятий с помощью предложенной семиотики формирования формул знаний.

Множество формул знаний образует бесконечную базу знаний. Построение ФСЛВ осуществлено на ФСЗ с использованием приема выделения формального языка на основе комбинаций логических силлогизмов и операций подстановок. Системы ФСЛВ могут использоваться при проектировании и синтезе автоматов и других объектов искусственного интеллекта. Кроме того, в пределах абстрактной ФСЛВ удается построить предметные ФСЛВ, которые позволяют шире и разнообразнее представить объекты систем искусственного интеллекта.

2. Логический вывод и грамматики

Рассмотрим вначале вопрос вывода языка с неформальной точки зрения. Обычно, формальные грамматики [6] решают прямую задачу выделения части (порожденного языка) из некоторого свободного языка. Процесс порождения языка основывается на элементарной операции подстановки [7], матричной операции подстановок [8], гомоморфизме [9]. Используемые логические выводы (высказывания) в системах искусственного интеллекта также могут служить инструментами для выделения формальных языков. Решение задачи выделения формального языка с использованием высказываний можно привести на основе комбинации логических силлогизмов [2] и операций подстановок [7]. Например, пусть A и $Q_{i,j}$ высказывания в двухзначной логике и высказывание $Q_{i,j}$ определено на продукциях подстановок q_i и q_j, тогда вывод цепочек языка может быть задан на комбинированном силлогизме

$$\text{а) } S_{i,j}: \begin{array}{|c|}\hline A\ \text{условие} \\ \hline A \Rightarrow Q_{i,j} \\ \hline q_{i_1}: x_{i_1} \to y_{i_1}, q_{i_2}: x_{i_2} \to y_{i_2}, \cdots, q_{i_k}: x_{i_k} \to y_{i_k} \\ \hline q_{j_1}: x_{j_1} \to y_{j_1}, q_{j_2}: x_{j_2} \to y_{j_2}, \cdots, q_{j_r}: x_{j_r} \to y_{j_r} \\ \hline \end{array},$$

в котором из отношения $A \mapsto 1$ следует истинность продукций q_i (возможность применения) и из ложности продукций q_i (не возможность применения) следует отношение $A \mapsto 0$ и истинность продукций q_j. При выводе цепочек, последовательности продукций в силлогизме $S_{i,j}$ может быть как упорядоченной, так и не упорядоченной.

Если высказывание $Q_{i,j}$ определено на некоторых элементах y_i и y_j, то порождение цепочек можно выполнить на продукциях подстановок типа:

б) $q_k : x_k \to S_{i,j}$, в которых

$$S_{i,j}: \begin{array}{|c|}\hline A\ \text{условие} \\ \hline A \Rightarrow Q_{i,j} \\ \hline y_{i_1}, y_{i_2}, \cdots, y_{i_n} \\ \hline y_{j_1}, y_{j_2}, \cdots, y_{j_m} \\ \hline \end{array}.$$

Представление подстановок типа б) на введенном силлогизме понимается, как альтернативные множества подстановок $\{q_r : x_k \to y_r\}$ или $\{q_s : x_k \to y_s\}$.

Вывод цепочек может быть получен и на суперпозициях правил а) и б) с применением более сложных подстановок, например, для правил а) – силлогизм $S_{i,j}$ может быть задан на подстановках с силлогизмами $q_k : x_k \to y_k, S_{n,m}$ видов а), б). При этом высказывания в силлогизмах могут быть определены в переменной n-значной логике высказываний. В дальнейшем правила типа а) и б), независимо от их вида обобщения, будем представлять компактной формой $S_{i,j} : A \vec{\Rightarrow} q_i \| q_j$ и $q_k : x_k \xrightarrow{A} y_i \| y_j$.

Правила вида а) формируют формальные конструкции на базе знаний и правилах формальных грамматик. Правила же вида б) являются непосредственно грамматическими и явно определяют условную как взвешенную грамматику. Конструирование языка на этих правилах требует различных подходов в представлении формальных структур. Так, например, в первом случае необходимо создать структуру базы знаний, а во втором – формальную структуру условий. Рассмотрим примеры конструирования языков при этих подходах.

Пусть на терминальном $\{a, b, c, e, d\}$ и нетерминальном $\{\alpha, \beta, \sigma\}$ алфавитах задана схема продукций грамматики:

$$Q = \begin{cases} q_1 : \sigma \to a\alpha b;\ q_2 : \sigma \to a\beta b; \\ q_3 : \alpha \to a\alpha c;\ q_4 : \alpha \to c\alpha; \\ q_5 : \alpha \to d; \quad q_6 : \alpha \to c; \\ q_7 : \beta \to e; \quad q_8 : \beta \to c; \end{cases}$$

в которой продукции q_1, q_2 - начальные для вывода правильных цепочек.

Схема Q порождает язык

$L(Q) = \{(a^k c^s)^u b\} \cup \{(a^m c^r)^p dc^v b\} \cup \{aeb\}$,

$r, v \in \{0, 1, 2, \ldots\}$;

$k, s, u, m, p \in \mathbb{N}$.

Предположим, что на нетерминалах, рассмотренной грамматики сформирована база знаний БЗ= $\{\alpha, \beta, \sigma, \alpha \wedge \beta, \alpha \vee \beta\}$, которая определена на двузначной логике так, что:

$$\varphi : \begin{cases} \alpha \mapsto 1, \\ \beta \mapsto 0, \\ \sigma \mapsto 1. \end{cases}$$

И пусть определены высказывания $A_1 \equiv \sigma$, $A_2 \equiv \alpha \wedge \beta$, $A_3 \equiv \alpha \vee \beta$, и $A_4 \equiv \alpha$, на которых задана схема логических выводов:

$$S = \begin{cases} S_{1,2} : A_1 \Rrightarrow q_1 \parallel q_2;\ S_{3,4} : A_2 \Rrightarrow q_3 \parallel q_4; \\ S_{5,6} : A_3 \Rrightarrow q_5 \parallel q_6;\ S_{7,8} : A_4 \Rrightarrow q_7 \parallel q_8. \end{cases}$$

Если $S_{1,2}$ начальный силлогизм, то в результате логических выводов по схеме S получим язык $L(S) = \{ac^n db;\ n = 0, 1, 2, \ldots\}$.

Такой же язык можно получить в результате применения грамматических выводов по схеме продукций вида б):

$$Q_1 = \begin{cases} q_1 : \sigma \xrightarrow{A_1} a\alpha b \parallel a\beta b;\ q_2 : \alpha \xrightarrow{A_2} a\alpha c \parallel c\alpha; \\ q_3 : \alpha \xrightarrow{A_3} d \parallel c; \qquad q_4 : \beta \xrightarrow{A_4} e \parallel c. \end{cases}$$

Схема логических выводов может быть более сложной, если использовать конструкции силлогизмов

$$\alpha \vee \beta \Rrightarrow q_i \mid q_j \parallel q_k \mid q_n \ \text{ и } \ \alpha \wedge \beta \Rrightarrow q_i \,\&\, q_j \parallel q_k \,\&\, q_n,$$

где символы | и & соответствуют безусловным связкам «или» и «и» соответственно. Так, по схеме

$$S_1 = \begin{cases} S_{1,2} : A_1 \Rrightarrow q_1 \parallel q_2;\ S_{3\&5,4\&6} : A_2 \Rrightarrow q_3 \,\&\, q_5 \parallel q_4 \,\&\, q_6; \\ S_{3|5,4|6} : A_3 \Rrightarrow q_3 \mid q_5 \parallel q_4 \mid q_6;\ S_{7,8} : A_4 \Rrightarrow q_7 \parallel q_8; \end{cases}$$

выводится язык

$$L(S_1) = \{a^{k+1} dc^k b;\ k = 0, 1, 2, \ldots\} \cup \{a^n c^{n+1} b;\ n = 1, 2, \ldots\}.$$

Эквивалентная к схеме S_1 по порождаемому языку может быть построена и схема продукций грамматики вида б):

$$Q_2 = \begin{cases} q_1 : \sigma \xrightarrow{A_1} a\alpha b \parallel a\beta b;\ q_2 : \alpha \xrightarrow{A_2} a\alpha c \,\&\, d \parallel c\alpha \,\&\, c; \\ q_3 : \alpha \xrightarrow{A_3} a\alpha c \mid d \parallel c\alpha \mid c; \qquad q_4 : \beta \xrightarrow{A_4} e \parallel c. \end{cases}$$

Структуры со схемами S и Q_1 будем называть конструкциями первого уровня, а со схемами S_1 и Q_2 – второго уровня.

В дальнейшем, ради простоты изложения, рассмотрим формализацию только для схем логического вывода первого уровня типа а). Для этого нам понадобится ввести формальную систему базы знаний ФСЗ [1].

3. Формальная структура высказываний (ФСВ) и представления БЗ

Как известно, существуют различные подходы к построению моделей представления знаний. Порождающая ФСВ, например, может строиться на алгебраических моделях Мальцева [9] и др. Воспользуемся схемой, предложенной в работе [3]: «предметная область» → «поле знаний» → «формализация» → «модель знаний». Предположим, что уже произошел переход к этапу «поле знаний». Выделим в поле знаний «объекты-факты», обозначив их символами φ_i алфавита $\Phi = \{\varphi_i\}_{i=0}^{n}$, в котором φ_0 соответствует пустому символу ε. Обратим внимание на то, что символ ε, в нашем случае, может выступать в роли пустого символа для множества-алфавита, нейтрального символа в алгебре операции конкатенации и неопределенного значения при высказываниях. Пусть формализацию «характеристик» фактов Φ выполняют элементы алфавита $B = \{\beta_i\}_{i=0}^{m}$, здесь β_0 - также соответствует пустому символу ε. Введем еще два алфавита: алфавит специальных символов $G = \{,\ ,\ ),\ (,\ ;,\ :,\ =,\ \varepsilon\}$ и алфавит предикатных значений $T = \{\text{true}, \text{false}, \varepsilon\}$, в котором символ ε соответствует - неопределенности. Тогда модель структуры системы знаний ФСЗ на основе продукционного метода представления можно записать как упорядоченную четверку [1]:

$$C_Z = \langle M_Z, C_A, \Sigma_Z, \Lambda_Z \rangle, \qquad (1)$$

где M_Z – носитель системы, составленный из алфавитов Φ, B, G и T; C_A – алгоритмическая структура формирования (генерирования) объектов знаний, сигнатура Σ_Z – совокупность множеств отношений O_1 и операций O_2; Λ_Z – семиотика системы.

Множество O_1 есть набор двуместных и одноместных отношений двухзначной логики $\{\wedge^2, \vee^2, \ni^2, \neg^1, \equiv^2, \Rightarrow^2\}$ и нестандартного отношения $\{\xrightarrow{p}{}^2\}$, а множество операций $O_2 = \{\underline{\cup}^2, \bar{\cup}^2, \cdot^2, \bullet^2\}$. Правила применения операций их свойства и пр. задаются в семиотике Λ_Z структуры (1).

Рассмотрим формальную семиотику формирования БЗ в этой системе.

Определение 1. Процесс формирования некоторого элемента в структуре C_A или системе C_Z называется его генерацией, а сам элемент – объектом генерации.

Формирование знаний в структуре (1) осуществляется на БД с помощью правил вывода семиотики системы C_Z.

1. Множество $\Phi \bigcup B$ образует БД:

– элементы предметных алфавитов Φ и B реализуются в структуре C_A и алфавиты Φ, B являются объектами генерации в системе C_Z;

– e - элементарное высказывание, если $\forall \varphi_i \in \Phi$ и $\beta_j \in B$, $e: \varphi_i \ni \beta_j$ или $e: \beta_j \ni \varphi_i$, или $e: \varphi_i \ni \varphi_i$, или $e: \beta_j \ni \beta_j$; множество высказываний $\{e_k\} = E$;

– высказывания $e_1, e_2 \in E$ – смежные, если $e_1: \varphi_i \ni \beta_j$ и $e_2: \varphi_i \ni \beta_k$ или $e_1: \beta_j \ni \varphi_i$ и $e_2: \beta_j \ni \varphi_m$;

– элементарные высказывания множества E генерируются в алгебраической структуре C_A;

– множество E – объект генерации;

– высказывание v – сложное, если $\forall e_i, e_j \in E$ имеет место $v: e_i \vee e_j$ или $v: e_i \wedge e_j$, или $v: \neg e_i$, или получено при помощи операции суперпозиции $(\bullet)$ и отношений $(\vee, \wedge, \neg)$ на других сложных высказываниях; множество сложных высказываний $\{v_k\} = V$;

– $\Phi \bigcup B \subset E \subset V$;

– элементы множества V формируются в алгоритмической структуре C_A и являются объектами генерации в структуре C_Z.

2. Формирование правил вывода в структуре C_Z осуществляется с помощью высказывания $(v_i \Rightarrow v_j)$:

1) $\forall v_i, v_j \in V; (v_i \equiv true) \wedge (v_i \Rightarrow v_j)$, тогда $(v_j = true)$ и правило вывода p представляется так $p: (v_i \Rightarrow v_j) \xrightarrow{p} v_j$ или в сокращенной форме

$$p: (v_i, v_j) = p(v_i, v_j) = (v_i, v_j); \qquad (2)$$

– бинарное правило p (отношение ($\xrightarrow{p}$)) – рефлексивное,

– правило вывода p – транзитивное;

2) транзитивность правила вывода p определяет операцию умножения правил $(\cdot)$, если $p_1: (v_1, v_2)$ и $p_2: (v_2, v_3)$, то $p_1 \cdot p_2 = p_3(v_1, v_3)$;

– операция $(\cdot)$ не коммутативна,

– операция умножения частично порождает рефлексию отношений, так для выводов $p_1: (v_1, v_2)$ и $p_2: (v_2, v_1)$ имеем $p_1 \cdot p_2 = p_3(v_1, v_1)$;

3) над правилами вида (2) частично можно выполнять операции левого $\underset{-}{\cup}$ и правого $\bar{\cup}$ объединений;

– отправляясь от предположения, что в структуру знаний введена алгоритмическая структура C_A, считаем операции $\underset{-}{\cup}$ и $\bar{\cup}$ допустимыми в структуре C_Z,

– $\forall p_1, p_2$ таких правил, что $p_1: (v_1, v_3)$ и $p_2: (v_2, v_3)$, имеет место, $p_1 \underset{-}{\cup} p_2 = \begin{cases} (v_1 \vee v_2, v_3), \\ (v_1 \wedge v_2, v_3); \end{cases}$ причем результат операции $p: (v_1 \vee v_2, v_3)$ возникает тогда, когда высказывания v_1 и v_2 смежные,

– $\forall p_1, p_2$ таких, что $p_1: (v_1, v_2)$ и $p_2: (v_1, v_3)$ применима $p_1 \bar{\cup} p_2 = (v_1, v_2 \vee v_3)$,

– операции левого $\underset{-}{\cup}$ и правого $\bar{\cup}$ объединений коммутативны и ассоциативны;

4) формирование языка знаний в структуре C_Z:

– P – множество формул сконструированных на правилах вида (2) и правилах, полученных с помощью множества операций $\{\underset{-}{\cup}^2, \bar{\cup}^2, \cdot^2, \bullet^2\}$ над этими выводами,

– $\forall p(v_i, v_j) \in P$, определено в структуре знаний порождение v_j объектом порождения $W_{v_j} = p(v_i, v_j)$,

– W_{v_i} – объект генерации на некоторой начальной конфигурации $\varpi_0 \subset \$ \bigcup B$,

– порождение знаний в структуре C_Z происходит по конечной или бесконечной схеме вывода $\varpi_0 \xrightarrow{p_{i_1}} W_{v_1} \xrightarrow{p_{i_2}} W_{v_2} \cdots \xrightarrow{p_{i_k}} W_{v_k}$,

– язык знаний в структуре C_Z образует множество W подмножеств $W_{\varpi_0} = \{W_{v_k}\}_k$ образованных на различных начальных конфигурациях,

– пустой (нейтральный) символ ε - элемент множества W,

– множество W перечислимо,

– язык W свободный по отношению $\xrightarrow{p}$.

Прагматически система ФСЗ может быть использована и для генерации более сложных логических высказываний на основе исчисления предикатов.

Очевидно, предложенная система ФСЗ является одной из множества общих формальных систем формирования знаний.

4. Формальная грамматическая структура ФСЛВ

Для задания формальной структуры необходимо корректно определить ее объекты. Во-первых, объекты грамматической структуры должны быть конечными [1]. В рассматриваемом случае ФСЛВ необходимо определять на бесконечном множестве W системы знаний. Поэтому в носитель ФСЛВ следует включить конечную структуру ФСЗ, которая формирует бесконечный язык знаний. Во-вторых, объекты носителя ФСЛВ должны быть свободными в том отношении, что их можно вставлять в конструкцию цепочек языка грамматики в любое место, предусмотренное правилами грамматической структуры. Исходя из этого свойства при формировании формального языка, постулируем следующее:

– любой алфавит носителя структуры ФСЛВ свободный;

– конструктивные элементы, построенные ФСЗ и используемые в структуре ФСЛВ свободные.

Определим модель ФСЛВ в виде грамматической структуры [1], т.е.:

$$C_{GL} = \langle M, \Sigma, \Lambda \rangle. \qquad (3)$$

В модели (3) носитель M состоит из системы формирования знаний C_Z, терминального и не терминального алфавитов $A = \{a_i\}_{i=0}^{s}$, $a_0 = \varepsilon$ и $N = \{\alpha_i\}_{i=1}^{p}$, и алфавита связок $\{|, \|, \&\}$ со смыслом «или», «альтернатива», «и» соответственно. Причем алфавиты N и Φ могут, как пересекаться так и не пересекаться, поэтому элемент $\alpha \in N \bigcap \Phi$ может иметь смысловую нагрузку «объекта-факта» из поля знаний. Сигнатура $\Sigma = \{\otimes^2, \rightarrow^2, \underset{q}{\Rightarrow}^2, \underset{C}{\Rightarrow}^k, \vec{\Rightarrow}^2, \underset{s}{\vec{\Rightarrow}}^2, \underset{C}{\vec{\Rightarrow}}^k\}$, в которой символ $\otimes$ обозначает операцию конкатенации, символы $\rightarrow$ и $\underset{q}{\Rightarrow}$ соответствуют отношениям обычной подстановки и непосредственного вывода цепочек, а – $\vec{\Rightarrow}$ и $\underset{s}{\vec{\Rightarrow}}$ отношения силлогизма и силлогизма непосредственного вывода цепочек. Отношения выводов цепочек языка обозначены в сигнатуре структуры C_{GL} символами $\underset{C}{\Rightarrow}$ и $\underset{C}{\vec{\Rightarrow}}$. Под операцией конкатенации, в общем случае, может пониматься совокупность ее модификаций $\otimes = \{\otimes_i\}$, однако, в дальнейшем эти модификации не будем рассматривать. Конструктивная аксиоматика Λ структуры (3) сложнее аксиоматики классической грамматической структуры. Рассмотрим правила формирования грамматических конструкций в составных частях аксиоматики Λ.

Аксиоматика формирования свободных языков

Свободные языки включают все возможные конструкции, полученные с помощью операции обладающей свойствами свободной алгебраической полугруппы.

Определение 2. Цепочки, полученные в результате операции конкатенации элементов носителя, образуют свободные языки A^*, L^* и H^*:

1) $A \bigcap N = \varnothing$;

$\forall a \in A \,|\, N \,|\, W,\ a \otimes \varepsilon = \varepsilon \otimes a = a$;

$\forall a, b \in A \,|\, N \,|\, W,\ a \otimes b \neq b \otimes a,\ a \otimes (b \otimes c) = (a \otimes b) \otimes c$;

$A^* = \{l_i\} : l_i = \varepsilon \,|\, l_i = l_j \otimes a_k, \quad a_k \in A,\ l_j \in A^*$;

2) $L^* = \{l_i\} : l_i = \varepsilon \,|\, l_i = l_j \otimes a_k \,|\, l_i = l_j \otimes \alpha_k$,

$a_k \in A,\ \ \alpha_k \in N,\ \ l_j \in L^* = (A \bigcup N)^*$;

3) $H^* = \{l_i\} : l_i = \varepsilon \,|\, l_i = l_j \otimes a_k \,|\, l_i = l_j \otimes \alpha_k \,|\, l_i = l_j \otimes v_k$,

$a_k \in A,\ \ \alpha_k \in N,\ \ v_k \in W$,

$l_j \in H^* = (A \bigcup N \bigcup W)^*$,

$\forall l_j \in H^*, l_j = l_1 v_i l_2,\ v_i \equiv \delta \in T,\ l_1, l_2 \in L^*$;

4) $A^* \setminus \{\varepsilon\} = A^+$, $L^* \setminus \{\varepsilon\} = L^+$, $H^* \setminus \{\varepsilon\} = H^+$.

Из этой части аксиоматики имеют место цепи по включению $A^+ \subset L^+ \subset H^+$ и $A^+ \subset A^* \subset L^* \subset H^*$.

Аксиоматика операций подстановки

Рассмотрим вначале отношение $(\rightarrow)$

$q_i : l_j \rightarrow l_k,\ l_j \in L^+,\ l_k \in L^*$,

$0 < |l_j| < k_1,\ 0 \leq |l_k| < k_2,\ k_1, k_2 \in \mathbb{N}$,

где $|l|$ – длина цепочки l;

множество продукций

$\{q_i\}_{i=1}^n = Q,\ q_i \neq q_j,\ i \neq j$;

если

$q_i : l_j \rightarrow l_k$ и $q_r : l_s \rightarrow l_m$, и $i \neq r$,

то справедливо $l_j \neq l_s \,|\, l_j = l_s$;

отношение $(\rightarrow)$ не является отношением эквивалентности.

Отношение ($\vec{\Rightarrow}$) определяется на декартовом произведении $H^+ \times (Q \times W)^r$ так, что

$s_i : l_j \vec{\Rightarrow} b_k \,\|\, b_m$, $l_j \in H^+$,

где b_i - выражения, составленные из пар отношений $q_i \in Q$, $v_m \in W$ и связок $(|, \&)$;

альтернативное отношение $(\|)$ определяет подстановки:

– если $v_i \equiv \text{true}$, то $l_j \vec{\Rightarrow} b_k$, $v_i \subset l_j$,

– если $v_i \equiv \text{false}$, то $l_j \vec{\Rightarrow} b_m$;

продукционное множество $\{s_i\}_{i=1}^m = S$.

Это отношение обладает теми же свойствами, что и отношение $(\rightarrow)$.

Аксиоматика отношений непосредственного вывода

Определение 3. Продукция $q_i \in Q$ - допустимая к цепочке $l \in H^+$ и задает непосредственный вывод цепочки $l_i \in H^*$, если:

– $\left(l = l_1 l_j l_2\right) \in H^+$ и $l_j \in L^+$,

– существует продукция $\left(q_i : l_j \rightarrow l_k\right) \in Q$ такая, что $l_1 l_j l_2 \underset{q_i}{\Rightarrow} l_1 l_k l_2 = l_i$.

Определение 4. Продукция силлогизма $s_i \in S$ - допустимая к цепочке $l \in H^+$ и задает непосредственный вывод $l \underset{s_i}{\vec{\Rightarrow}} l_i$ цепочки $l_i \in H^*$, когда:

1) $\left(l = l_1 l_j l_2\right) \in H^+$ и $l_j \in L^+$,

– в выражении b_k продукции $s_i : l_r \vec{\Rightarrow} b_k \,\|\, b_m$, $l_r \in H^+$, $l_r = l_1 v_r l_2$ при значении $v_r \equiv \text{true}$ существует хотя бы одна допустимая продукция $\left(q_i : l_j \rightarrow l_k\right) \in Q$,

– или в выражении b_m продукции s_i при значении $v_r \equiv \text{false}$ найдется, по крайней мере, одна допустимая продукция $\left(q_p : l_j \rightarrow l_n\right) \in Q$;

2) $\left(l = l_i l_r l_j\right) \in H^+$ и $l_r \in H^+$, $l_r = l_1 v_r l_2$,

– продукция силлогизма s_i такая, что $s_i : l_r \vec{\Rightarrow} b_k \,\|\, b_m$ и имеет место $l \underset{s_i}{\vec{\Rightarrow}} l_i (b_k \,\|\, b_m) l_j$ с допустимыми продукциями из множества Q к цепочкам l_i или l_j.

Аксиоматика начальных символов и цепочек

Для организации процесса отбора языка из некоторого свободного множества с помощью допустимых продукций q_i и s_k зададим множество начальных символов $U \subset N$ и цепочек $X \subset H^+$:

– если $q_i : \alpha \rightarrow l$, где $l \in L^*$ и $\alpha \in U$, то q_i - начальная допустимая продукция;

s_i - начальная допустимая продукция силлогизма

– если $s_i : l_r \vec{\Rightarrow} b_k \,\|\, b_m$ и в выражениях b_k или b_m, или в обеих – присутствует хотя бы одна начальная допустимая продукция вида $q_i : \alpha \rightarrow l$,

– если $s_i : l_r \vec{\Rightarrow} b_k \,\|\, b_m$ и $l_r \in X$.

Аксиоматика вывода правильной цепочки

Определение 5. Последовательность непосредственных выводов $l_j \underset{q_i}{\Rightarrow} l_k$ и $l_r \underset{s_i}{\vec{\Rightarrow}} l_m$ из некоторой цепочки $l_1 \in H^+$ с помощью допустимых продукций q_i и s_i образует вывод $l_1 \underset{C}{\Rightarrow} l_2$ или $l_1 \underset{C}{\vec{\Rightarrow}} l_2$, или $l_1 (\underset{C}{\Rightarrow}, \underset{C}{\vec{\Rightarrow}}) l_2$ цепочки $l_2 \in H^*$ в структуре C_{GL}.

Каждый непосредственный вывод силлогизма, применяемый при выводе цепочки, выполняется на всех допустимых продукциях q_i выражений соответствующих истинностным знаниям.

Определение 6. Цепочка l_m называется правильной, если найдутся такие допустимые мультимножества продукций $\{q_i\}$ и $\{s_i\}$, которые образуют ее вывод в структуре ФСЛВ причем:

– вывод начинается с начальных символов множества U,

– $l_m \in A^*$.

Множество $\{l_m\} = L(C_{GL})$ образует формальный язык, выводимый в структуре ФСЛВ.

В заключение приведем несколько общих результатов устанавливающих зависимости между языками.

Утверждение 1. Из всякого мультисимвольного языка L(C) грамматической структуры C всегда можно выделить язык, выводимый в структуре ФСЛВ.

Так как продукции программной грамматической структуры являются частными случаями структуры ФСЛВ, то можно утверждать, что.

Утверждение 2. Множество программных языков является подмножеством ФСЛВ языков.

Как было указано во вступлении, абстрактные типы данных представляются математическими моделями с последующей программной реализацией на выбранном языке. Например, в структуре ФСЛВ построим формальную модель рассмотренного выше АТД – мультимножество. Для этого введем некоторые допущения, обозначения и уточнения задачи.

Дополним ФСЛВ новыми понятиями. Если x имя (идентификатор) элемента, индекса и пр., то $|x|$ его значение и введем номинативное отношение $''\mapsto''$, действующее по правилу $|x|\mapsto m, m \in K$. Рассмотрим отношения сравнения $''=\!=''$ элементов по именам $''\le''$ по значениям и операции $''+''$ сложения и вычитания $''-''$ значений. Пусть $E = \{a_i\}$ некоторый упорядоченный по индексу универсум символов-имен. Мультимножество – множество с повторяющимися элементами определенной кратности. Теперь уточним поставленную задачу. Необходимо построить систему продукций предметной грамматической структуры, которые формируют на универсуме E мультимножество M с n различными элементами a_i кратности k_i. Очевидно, мощность мультимножества есть $\#M = \sum_{i=1}^{n} k_i$.

Система продукций решающую задачу формирования множества M следующая:

$$s_1 : (\sigma v) \vec{\Rightarrow} (\sigma \to \{\alpha\} \& |i| \mapsto 1) \| (\sigma \to \{\varepsilon\});$$

$$s_2 : (|i| \le |n|) \vec{\Rightarrow} \big(|j| \mapsto 1 \& (|j| \le k_i) \vec{\Rightarrow}$$
$$\vec{\Rightarrow} (\alpha \to a_i, \alpha \& |j| \mapsto |j|+1)\big) \& |i| \mapsto |i|+1 \| (,\alpha \to \varepsilon).$$

Модель операции построения множества типа set по мультимножеству M также проста:

$$s_3 : (\sigma v) \vec{\Rightarrow} (\sigma \to \{\alpha\} \& |i| \mapsto 1 \& |j| \mapsto 1) \| (\sigma \to \{\varepsilon\});$$

$$s_4 : (|i| \le |n|) \vec{\Rightarrow} \big((a_i =\!= a_j) \vec{\Rightarrow}$$
$$\vec{\Rightarrow} (\alpha \to a_i, \alpha \& |j| \mapsto |j| - 1 + |k_j|)\big) \& |i| \mapsto |i|+1 \| (,\alpha \to \varepsilon).$$

Анализ схемы продукций позывает, что модель АТД можно упростить (оптимизировать), если мультимножество представлять набором пар $\{(a_i, k_i)\}_{i=1}^{n}$.

5. Выводы

Предложена сложная гибридная структура, учитывающая формальную структуру знаний и обычную грамматическую структуру. Введенная структура ФСЛВ разнообразнее, чем классические грамматические структуры и позволяет конструировать цепочки, как на основе обычных подстановок, так и с помощью подстановок силлогизма.

Структуры ФСЛВ шире и разнообразнее возможностей грамматических структур программных языков.

Введенная ФСЛВ структура допускает обобщение на случай исчисления предикатов многозначной логики.

Литература

1. Ільман, В. М. Формальні структури та їх застосування [Текст] / В. М. Ільман, В. В. Скалозуб, В. І. Шинкаренко. – Д.: Вид-во Дніпропет. нац. ун-ту залізн. трансп. ім.. акад. В. Лазаряна, 2009. – 205 с.
2. Представление и использование знаний [Текст]. Пер. с япон. / Под ред. Х. Уэно, М. Исидзука – М.: Мир, 1989. – 220 с.
3. Гаврилова, Т. А. Базы знаний интеллектуальных систем [Текст] / Т. А. Гаврилова, В. Ф. Хорошевский. – СПб.: Питер, 2000. – 384 с.
4. Ахо, А. В. Структуры данных и алгоритмы [Текст] / А. В. Ахо, Д. Э. Хопкрофт, Д. Д. Ульман. – М.: Издательский дом «Вильямс», 2003. – 384 с.
5. Топп, У. Структуры данных в C^{++} [Текст] / У. Топп, У. Форд. – М.: ЗАО «Из-во БИНОМ», 1999. – 816 с.
6. Фу, К. Структурные методы распознавания образов [Текст] / К. Фу. – М.: Мир, 1977. – 318 с.
7. Гладкий, А.В. Формальные грамматики и языки [Текст] / А.В. Гладкий. – М.: Наука, 1973. – 368 с.
8. Андон, Ф. И. Алгебро-алгоритмические модели и методы параллельного программирования [Текст] / Ф. И. Андон, А. Е. Дорошенко, Г. Е. Цейтлин, Е. А. Яценко. – К.: Академпериодика, 2007. – 634 с.
9. Саломаа, А. Жемчужины теории формальных языков [Текст] / А. Саломаа. – М.: Мир, 1986. – 159 с.
10. Мальцев, А. И. Алгебраические системы [Текст] / А. И. Мальцев. – М.: Наука, 1970. – 391 с.

Abstract

The formal grammatical approach and the approach based on logical statements are general-purpose simulation means of information systems. These approaches are used in designing and representation of programming *languages, algorithms, state machines, systems of pattern recognition, designing of compilers, bases of data and knowledge, and other applications of artificial intelligence.*

The goal of this research is to construct a hybrid grammatical structure which allows to generalize and extend both grammatical structures with operations of substitution, matrix substitution, homomorphism and intelligent systems with the operations of substitution and logical syllogisms.

The formal hybrid system is constructed in two stages. At the first stage, the formal structure of knowledge is determined as an element of the grammatical structure carrier. And at the second stage, the formal grammatical system is determined with the logical conclusion. The formal structure of knowledge is based on the production approach of knowledge representation as a subjective formal constructive system. The system is constructed on the dictionary of notions using the proposed semiotics of formation formulas of knowledge. The set of formulas of knowledge forms an infinite knowledge base. The construction of the formal grammatical system with the conclusion is implemented on the formal structure of knowledge using the technique for the identification of the formal language based on a combination of logical syllogisms and operations of substitutions.

The formal grammatical system with logical conclusion can be used in the designing and synthesis of state machines and other objects of artificial intelligence

Keywords: *grammatical structure, knowledge representation, hybrid structure, syllogism, grammatical conclusion, formal language*

УДК 51-76

ИНФОРМАЦИОННАЯ ТЕХНОЛОГИЯ КОМПЛЕКСНОЙ ОЦЕНКИ ПРОЦЕССОВ СТАРЕНИЯ БИОЛОГИЧЕСКИХ ОРГАНИЗМОВ

Н.Д. Гернет
Старший научный сотрудник*
E-mail: nadezhdadg@yandex.ru
А.И. Божков
Доктор биологических наук, профессор
Кафедра молекулярной биологии и биотехнологии
Директор*
Контактный тел.: (057) 707-53-40
E-mail: bozhkov@univer.kharkov.ua
*НИИ биологии
Харьковский национальный университет им. В.Н. Каразина
пл. Свободы, 4, г. Харьков, Украина, 61022

Розглядається інформаційна технологія комплексної оцінки процесів старіння біологічних організмів (біологічних систем) на основі ідей методу аналізу ієрархій. Інформаційна технологія дозволяє визначати найбільш інформативні кількісні характеристики старіння класу біологічних організмів, що розглядається, і на їх основі - інтегральні характеристики системного старіння біологічних організмів, комплексну оцінку процесів старіння організму в цілому, їх динамічні особливості

Ключові слова: біологічний організм, старіння, інформаційна технологія, метод аналізу ієрархій, експерт, комплексна оцінка, інтегральні характеристики

Рассматривается информационная технология комплексной оценки процессов старения биологических организмов (биологических систем), основанная на идеях метода анализа иерархий. Информационная технология позволяет определять наиболее информативные количественные характеристики старения рассматриваемого класса биологических организмов и на их основе - интегральные характеристики системного старения биологических организмов, комплексную оценку процессов старения организма в целом, их динамические особенности

Ключевые слова: биологический организм, старение, информационная технология, метод анализа иерархий, эксперт, комплексная оценка, интегральные характеристики

1. Введение

В настоящее время одной из наиболее значимых мировых проблем является проблема постарения человеческой популяции, состоящая в резком увеличении в популяции доли лиц с существенно сниженными возможностями осуществления своих биологических и социальных функций. С этим связан ряд острых