УДК 004.82 DOI: 10.17587/prin.13.226-238

К. И. Костенко, канд. физ.-мат. наук, доц. кафедры, kostenko@kubsu.ru,
Кубанский государственный университет, Краснодар

Регулярные структуры памяти и домены операций интеллектуальных систем

Определено понятие регулярной области памяти интеллектуальной системы. Основу описаний структур памяти, применяемых в таких системах, определяет универсальный формат представления знаний в формализмах семантических иерархий. Формат всей памяти реализует бесконечное насыщенное бинарное дерево, вершинами которого являются двоичные наборы. Отдельные области памяти задаются с использованием регулярных выражений. Всякое выражение определяет семейство вершин, расширяемое во фрагмент дерева, — регулярную область памяти. Области соответствует класс знаний (домен). Регулярное выражение задает границы структур представлений знаний такого класса. Регулярные области памяти обобщают систему классов доменов морфизмов, разработанную для формализмов семантических иерархий. Исследование регулярных структур памяти способствует разработке и использованию специальных средств моделирования организации памяти и процессов мышления в системах искусственного интеллекта. Связь регулярных выражений с конечными автоматами обеспечивает возможность эффективного моделирования операций обработки и процессов потоков знаний.

Ключевые слова: *формализм знаний, морфизм знаний, домен морфизма, регулярная область памяти, описание структуры памяти*

Введение

Для конструирования математических моделей интеллектуальных систем (ИС) применяют средства формального описания (представления) знаний (формализмы). Они определяются как специальные алгебраические системы [1]. Инварианты формализмов связаны с отражением разных аспектов понятия знания. Общими свойствами формализмов являются конструктивный характер множеств абстрактных знаний, специальных предикатов и преобразований (морфизмов) для таких знаний. В уточнениях структурных и функциональных свойств знаний применяют понятия и конструкции из разных областей математики. Последние соответствуют сущностям моделей, которые связаны с изучением структур памяти и процессов мышления. Универсальная абстрактная модель ИС является многопредметной. Этим обеспечивается полнота и целостность моделирования разных представлений о свойствах интеллектуальности [2].

Элементами модели ИС являются области памяти отдельных компонентов, системы операций и процессов обработки знаний, а также агенты, управляющие жизненными циклами системы. Основу инвариантов модели составляют формализации понятий и принципов таких областей, как когнитивная психология и лингвистика. Управление субъектным существованием ИС в соответствующей области знаний моделируется в инвариантах и принципах кибернетики и общей теории систем. Рассмотренная система инвариантов модели допускает трансформацию в модели прикладных ИС, поддерживающих возможность полнофункционального моделирования двойников интеллектуальных сущностей в отдельных областях знаний [2].

Трансформация абстрактной модели в модели прикладных ИС реализуется морфизмами гомоморфного расширения ее элементов. Конкретные расширения образуют промежуточные модели, адаптированные к особенностям областей знаний. Инструменты трансформации моделей включаются в схемы технологии конструирования формальных моделей ИС разного уровня. Кибернетические инструменты реализации целей систем позволяют формировать инварианты субъектного существования и взаимодействия системы с областями знаний.

Структура организации памяти в компонентах ИС является атрибутом памяти абстрактных моделей. Такая структура связана с инвариантом алгебраической структуры знания в произвольном формализме знаний (семантических иерархий). Структуры знаний трансформируются в системы областей и форматы представления знаний в памяти компонентов ИС для разных этапов существования систем. Структуры памяти отдельных компонентов ИС соответствуют доменам операций и процессов синтеза знаний в таких компонентах и потоках знаний между компонентами. Домены объединяют многообразия знаний близкой структуры, применяемые для моделирования разных типов трансформации структур знаний при реализации целей и операций ИС.

Многообразие доменов морфизмов семантических иерархий включает обширный эмпирически формируемый фрагмент. Ему соответствуют домены морфизмов, адаптированные к форматам исходных данных и результатам содержательных операций, а также алгоритмам реализации таких операций. Спецификации отдельных доменов рассматриваемого многообразия разнородные и слабо связанные между собой. Семейство доменов замкнуто относительно операций прямой суммы и произведения баз [2]. Операциями суммы моделируются схемы интеграции элементов заданных доменов в общую структуру. Произведение доменов состоит в формировании классов знаний, получаемых заменой висячих вершин структур знаний первого домена на структурные представления знаний второго домена произведения.

Цель работы, результаты выполнения которой представлены в статье, — создание обобщающего подхода к описанию класса доменов операций обработки структурированных знаний. Подход основан на описании структур фрагментов бесконечных бинарных деревьев, в которых размещаются элементы конкретных доменов. Элементы таких фрагментов составляют области памяти ИС, предназначенные для формирования и размещения структурированных знаний. Структуры знаний адаптированы к форматам данных моделируемых операций обработки знаний, применяемых для реализации целей и задач ИС. Домены морфизмов включаются в описания структур памяти компонентов ИС. Структуры связаны с процессами размещения, извлечения и преобразования знаний, выполняемыми в соответствующих компонентах. Описания структур памяти задаются специальными формулами. Формируемая система областей памяти может рассматриваться как топология открытых множеств на множестве вершин бесконечного бинарного дерева. Операции и свойства такой топологии допускают интерпретацию, основанную на инвариантах алгебраической и семантической структур знаний, операций и процессов обработки знаний в памяти ИС.

Для описания структур памяти ИС далее будут применяться формулы, составленные из доменов операций в ИС и операций комбинирования доменов. Всякая формула определяет разбиение памяти компонента на подобласти, в которых реализуются определенные операции над знаниями для времени существования ИС. Формулы структур памяти являются инвариантами схем управления памятью компонентов ИС, применяемых в абстрактной модели таких систем. Развитие и изменение структур знаний, составляющих память компонентов ИС, являются элементами такого управления.

1. Формализмы представления знаний

Формализмами представления знаний называются четверки $(M, D_M, \circ, \prec)$. Здесь M (D_M) — алгоритмически перечислимое множество знаний (множество фрагментов знаний) и M разрешимо в D_M. Множество M содержит пустое знание, обозначаемое как Λ. Вычислимая операция $\circ$: $D_M \times D_M \to D_M$ называется композицией, а разрешимое отношение $\prec \subseteq D_M \times D_M$ — вложением фрагментов знаний. Среди формализмов особое положение занимают формализмы семантических иерархий [4]. Отдельные знания в таких формализмах называются конфигурациями. Всякий формализм семантических иерархий включает перечислимые множества конфигураций (M) и разрешимых бинарных отношений между конфигурациями (R). Класс знаний всякого формализма семантических иерархий составляет множество $M \cup R$. Структуры конфигураций формализма определяют вычислимые отображения разложения и связывания конфигураций $\varepsilon : M \to M \times M$ и $\psi : M \to R$. Если $\varepsilon(z) = (\Lambda, \Lambda)$, то $z \in M$ называется элементарной конфигурацией. Если $\varepsilon(z) = (z_1, z_2)$, где z_1 и z_2 — элементарные конфигурации, то z называется простой конфигурацией. Множество элементарных конфигураций обозначается как M_0. Если $z \in M$, $\varepsilon(z) \neq (\Lambda, \Lambda)$ и $\varepsilon(z) = (z_1, z_2)$, то структуру z составляют конфигурации z_1 и z_2. В $z \in M$ эти конфигурации связывает отношение $\psi(z) \in R$.

Фрагментами знаний являются сущности, применяемые для конструирования конфигураций из конфигураций с помощью бинарной операции композиции $\circ$. Если $\varepsilon(z) = (z_1, z_2)$ и $\psi(z) = r$, то конструирование z с помощью $\circ$ выполняется в два этапа. Сначала формируется композиция $z_1 \circ r$, которая применяется для композиции с z_2 в виде $(z_1 \circ r) \circ z_2$.

2. Начальные инварианты структур памяти и знаний

Универсальное множество вершин, применяемое для формирования структурных представлений знаний в произвольных формализмах семантических иерархий, составляет бесконечное насыщенное бинарное дерево с правым и левым потомками у каждой вершины. Такие вершины представляются конечными двоичными последовательностями. Множество таких последовательностей (I) включает пустой набор (λ). Соответствие наборов вершинам дерева устанавливают простые правила. Корню дерева соответствует пустой двоичный набор λ. Если выбранной вершине дерева соответствует $\alpha \in I$, то левый и правый потомки вершины α соответствуют наборам $\alpha 0$ и $\alpha 1$.

Полное структурное представление (ПСП) отдельного знания в формализме семантических иерархий задается композицией элементарных знаний формализма, составляющей это знание [1]. Для данного формализма обеспечивается единственность композиции, представляющей произвольное непустое знание (конфигурацию). Каждое такое представление имеет вид нагруженного бинарного дерева. Внутренним вершинам этого дерева приписывается операция композиции фрагментов знаний $\circ$. Висячие вершины всякого дерева размечаются элементарными знаниями. Для формализмов семантических иерархий удобно использовать собственный формат ПСП знаний (конфигураций) нагруженными бинарными деревьями. Он определяется отображениями разложения и связывания конфигураций. Висячие вершины таких деревьев размечены элементарными конфигурациями. Внутренние вершины размечены отношениями, выполняющимися между конфигурациями, представляемыми левым и правым поддеревьями таких вершин.

Для ПСП конфигураций будем использовать специальные обозначения для атрибутов таких структур. Если $z \in M$, то ПСП z обозначается как $\Sigma(z)$. Такая структура единственная для каждой неэлементарной $z \in M$. Обозначения M_i и Σ_i (i = 0, 1, ...) применяются для множеств конфигураций (ПСП конфигураций), представляемых деревьями глубины i. Множество вершин (висячих вершин) ПСП z обозначается как $D(z)$ ($O(z)$). Разметка $\alpha \in D(z)$ ПСП z обозначается как $[z]_\alpha$. Выражение $(z)_\alpha$ обозначает конфигурацию, представимую поддеревом дерева $\sum(z)$ с корнем $\alpha \in D(z)$.

3. Структуры памяти и домены морфизмов в формализмах знаний

Множество I и связанная с ним структура бесконечного бинарного дерева представляют удобный общий унифицированный формат моделирования структур памяти компонентов ИС, имеющий разнообразные приложения. Каждый компонент ИС использует собственный формализм представления знаний (подходящий формализм семантических иерархий). Содержание памяти компонента на разных этапах существования ИС составляют знания в применяемом формализме. Знание z размещается в структуре памяти компонента в форме бинарного дерева ПСП конфигурации. Корнем размещения ПСП назначается некоторая вершина памяти компонента. Прямая сумма ПСП конфигураций, одновременно размещенных в памяти компонента, составляет интегрированное содержание этой памяти в заданный момент времени.

Многообразие действий, используемое для моделирования реализаций целей и задач ИС, связано с морфизмами обработки знаний. Обрабатываемые знания размещаются в нескольких областях памяти ИС. Структуры отдельных областей соответствуют доменам морфизмов обработки знаний, размещаемых в таких областях. Семейство морфизмов формируется при развитии представлений о структурах памяти и процессах мышления, адаптированных к алгебраическим, логическим, топологическим инвариантам моделей различных областей математики, инвариантам других областей знаний. Эмпирически развиваемое семейство доменов морфизмов знаний связано с операциями, применяемыми для реализации процессов конструирования (синтеза) знаний. Процесс синтеза основан на знаниях, составляющих представление содержания области знаний и начальные данные процесса в заданном формализме представления знаний. Унифицированный формат представления содержания областей знаний образуют семейства элементарных и простых знаний в применяемом формализме. Такие семейства являются аналогами онтологий, составленных подмножествами классов элементарных знаний (индивидуалов) и простых знаний (связей между индивидуалами).

Процесс конструирования знаний из элементов представления содержания области знаний (онтологии) называется синтезом. Он основан на конструировании и трансформации структур знаний из элементов онтологии. Операции, используемые процессами синтеза, включают различные структурные, алгебраические и логические преобразования. Примерами доменов операций над знаниями являются окрестности и серии знаний [1]. Окрестности знаний представляют фрагменты онтологии. Они содержат элементарные знания, которые связаны с заданными знаниями произведениями отношений. Серии знаний формируются из фрагментов онтологий, извлекаемых с помощью подходящих предикатов (критериев). Структурные трансформации знаний моделируются операциями вставки, удаления, перестановки и замены фрагментов знаний. Алгебраические и логические трансформации выполняют замену элементов начальных данных алгебраических (логических) операций (правил вывода) на результаты применения таких операций (правил) [1]. Сериями моделируются процессы обработки знаний в ИС, представляемые последовательностями знаний, формируемых на разных шагах (этапах) процессов.

4. Регулярные выражения и множества вершин памяти ИС

Определение доменов морфизмов как алгоритмически перечислимого семейства перечислимых множеств конфигураций, замкнутого относительно операций прямой суммы $(\oplus)$ и произведения $(\otimes)$ доменов, является общим и абстрактным [2]. Оно формирует многообразие доменов морфизмов представлений знаний (конфигураций), обеспечивающее моделирование семейства общих классов конфигураций, обрабатываемых операциями разных типов. Рассмотрим сужение этого многообразия на семейство структур, обобщающих семейство доменов морфизмов, разработанных для моделирования преобразований и процессов обработки абстрактных знаний, адаптирующих функциональные сущности из разных областей математики. Семейство включает приведенные ранее виды доменов операций в формализмах знаний. Унифицированные описания доменов основаны на множествах вершин бинарных деревьев, определяемых с помощью регулярных выражений. Такие вершины применяются в качестве листьев ПСП семантических иерархий, принадлежащих доменам. Класс регулярных выражений для множества двоичных наборов определяют следующие правила:

1) запись всякого набора $\alpha \in I$ является регулярным выражением;

2) если E_1 и E_2 являются регулярными выражениями, то запись $\left(E_1 \cup E_2\right)$ является регулярным выражением;

3) если E_1 и E_2 являются регулярными выражениями, то запись $\left(E_1\right) \circ \left(E_2\right)$ является регулярным выражением;

4) если E является регулярным выражением, то запись $(E)^*$ является регулярным выражением.

Никакие другие записи не являются регулярными выражениями.

Приведенные соотношения соответствуют общему определению регулярного выражения. Такие выражения являются основой конструирования классов регулярных областей памяти и регулярных доменов операций в ИС. Заданные схемы конструирования регулярных выражений называются записью элементарного выражения, объединением и композицией регулярных выражений, а также итерацией регулярного выражения.

Всякому регулярному выражению E соответствует непустое множество $U(E)$ двоичных наборов, представляемых E. Такие множества уточняются отдельно для правил 1—4. Если $E = \alpha$, то $U(E) = \{\alpha\}$. Если $E = \left(E_1 \cup E_2\right)$, то $U(E) = U(E_1) \cup U(E_2)$. Если E_1 и E_2 — регулярные выражения, то $E_1 \circ E_2$ $(E_1 E_2)$ представляет множество двоичных наборов $\left\{\alpha_1\alpha_2 \mid \alpha_1 \in U\left(E_1\right) \& \alpha_2 \in U\left(E_2\right)\right\}$. Наконец, для выражения $(E)^*$ справедливо соотношение $U((E)^*) = (U(E))^*$. Выражение $(E)^*$ представляет множество наборов, являющихся сцеплениями конечных последовательностей элементов $U(E)$. Для объединений регулярных выражений выполняются соотношения ассоциативности и коммутативности:

$$\left(\left(E_1 \cup E_2\right) \cup E_3\right) = \left(E_1 \cup \left(E_2 \cup E_3\right)\right) \text{ и}$$

$$(E_1 \cup E_2) = (E_2 \cup E_1).$$

Поэтому в записях объединений регулярных выражений внутренние скобки можно опускать. Для рассмотренного ранее примера запись регулярного выражения имеет вид $\left(E_1 \cup E_2 \cup E_3\right)$.

Дополнительные форматы записи регулярных выражений связаны с комбинациями правил 1—4. Например, обозначение $E_1 \oplus E_2$ соответствует выражению $\left(0E_1 \cup 1E_2 \cup \lambda\right)$. Для последнего выражения справедливо соотношение

$$U\left(E_1 \oplus E_2\right) = \left\{0\alpha \mid \alpha \in U\left(E_1\right)\right\} \cup \left\{1\alpha \mid \alpha \in U(E_2)\right\} \cup \{\lambda\}.$$

Поэтому, если E_1 и E_2 представляют множества вершин конфигураций z_1 и z_2, то $E_1 \oplus E_2$ представляет множество вершин суммы $z_1 \oplus z_2$. Существование подходящих выражений E_1 и E_2, для любых $z_1,\ z_2 \in M$ следует из следующего утверждения.

Для любого конечного непустого множества $B \subseteq I$ $\ (D(z))$ существует регулярное выражение, представляющее B.

Пусть $B = \{\alpha_1, ..., \alpha_k\}$ $\ (D(z) = \{\alpha_1, ..., \alpha_k\})$ — непустое подмножество множества I. Тогда B представляется регулярным выражением $(\alpha_1 \cup \alpha_2 ... \cup \alpha_k)$.

Множество $B \subseteq I$ называется регулярным множеством вершин бесконечного бинарного дерева, если существует такое регулярное выражение E, что $U(E) = B$. Регулярные множества вершин составляют алгоритмически перечислимые и разрешимые области множества I. Множество I является регулярным, поскольку $I = U((0 \cup 1 \cup \lambda)^*)$. Пустое множество двоичных наборов не представляется регулярным выражением и также считается регулярным. Теорема С. Клини об эквивалентности автоматных множеств слов и регулярных множествах слов позволяет применять детерминированные конечные автоматы для исследования свойств регулярных множеств вершин бесконечного полного бинарного дерева и конструирования таких множеств с заданными свойствами.

Если E — регулярное выражение, то множества $U(E)$ может оказаться недостаточно для того, чтобы из элементов этого множества можно было составлять множества вершин ПСП конфигураций. Например, регулярное выражение $E = (0)^*$ представляет множество вершин, составляющих бесконечную левую ветвь бинарного дерева с вершинами из I, начинающуюся в вершине 0. Вершин этой ветви недостаточно для формирования структуры какой-либо конфигурации.

Уточним расширения регулярных множеств вершин в I до множеств, из элементов которых можно конструировать множества вершин ПСП семантических иерархий, которые составляют регулярные домены морфизмов.

Определение. Замыканием $B \subseteq I$ называется множество:

$$[B] = \{\beta \mid \beta \in I \,\&\, \exists\alpha \in B(\beta \subseteq \alpha)\}.$$

Здесь выражение $\beta \subseteq \alpha$ означает, что набор β является началом набора α. В частности, для каждой $z \in M$ справедливо соотношение $[O(z)] = D(z)$.

Теорема 1. Для всякого регулярного выражения E существует такое регулярное выражение E', что $[U(E)] = U(E')$.

Доказательство. Справедливость доказываемого утверждения связана с ограниченностью длины записи всякого регулярного выражения. Для произвольного выражения E имеет место один из случаев: $E = \sigma_1 ... \sigma_k$, $E = (E_1 \cup E_2)$, $E = E_1 \circ E_2$ и $E = (E_1)^*$. Определим оператор $\mathfrak{F}$, преобразующий E в подходящее выражения E'.

Если $E = \sigma_1 ... \sigma_k$, то положим $\mathfrak{F}(E) = (\lambda \cup \sigma_1 \circ \mathfrak{F}(\sigma_2 ... \sigma_k))$. Нетрудно проверить, что $\mathfrak{F}(E) = (\lambda \cup \sigma_1 \cup \sigma_1\sigma_2 \cup ... \cup \sigma_1 ... \sigma_k)$. То есть $U(E') = U(\mathfrak{F}(E))$.

Если $E = (E_1 \cup E_2)$, то положим $\mathfrak{F}(E) = (\mathfrak{F}(E_1) \cup \mathfrak{F}(E_2))$. То есть $U(\mathfrak{F}(E))$ определяется как множество двоичных наборов, представляемых выражениями E_1 и E_2, а также всех начал таких наборов.

Пусть $E = E_1 \circ E_2$. Множество начал двоичных наборов, представляемых E, составляют начала наборов, представляемых E_1, а также всевозможные наборы, начинающиеся с набора из $U(E_1)$ и продолжаемых наборами из $\mathfrak{F}(E_2)$. То есть, для в рассматриваемого случая $\mathfrak{F}(E) = (\lambda, \mathfrak{F}(E_1), E_1 \circ \mathfrak{F}(E_2))$.

Рассмотрим случай $(E = (E_1)^*)$. Началами наборов из $U(E)$ являются наборы, начинающиеся с некоторой (возможно пустой) последовательности наборов из $U(E)$ и продолжающейся началом некоторого набора из $U(E)$. Множество наборов с заданными свойствами представляется выражением $(\lambda \cup E_1)^* \mathfrak{F}(E_1)$.

Приведенные правила позволяют заменить заданное регулярное выражение E на выражение, представляющее множество наборов $[U(E)]$. Процесс построения выражения $\mathfrak{F}(E)$ завершается за конечное число шагов, поскольку длины записей выражений, для которых предполагается дополнительное уточнение значения оператора $\mathfrak{F}$, уменьшаются. *Доказательство окончено.*

Общими свойствами множеств двоичных наборов, представляемых регулярными выражениями, являются их алгоритмическая перечислимость и разрешимость в I.

Определим специальную процедуру пересчета элементов таких множеств. Пусть E — регулярное выражение. Для пересчета элементов $U(E)$ применим конструкцию корневого дерева $\mathfrak{T}(E)$. Схема пересчета наборов основана на рекурсивной обработке записи выражения E. Ему соответствует один из случаев определения регулярного выражения $E = (\sigma_1 ... \sigma_k)$, $E = (E_1 \cup E_2)$, $E = E_1 \circ E_2$ и $E = (E_1)^*$.

Разметкой корня дерева $\mathfrak{T}(E)$ в каждом из приведенных случаев является E. В первом случае корень $\mathfrak{T}(E)$ размечен набором $\sigma_1 ... \sigma_k$. Такая разметка объявляется заключительной для $\mathfrak{T}(E)$. Для случая $E = (E_1 \cup E_2)$ корень $\mathfrak{T}(E)$ имеет два потомка (левый и правый). Эти вершины являются корнями деревьев $\mathfrak{T}(E_1)$ и $\mathfrak{T}(E_2)$. Если разметка некоторой вершины этих деревьев объявлена заключительной в таком дереве, то она объявляется заключительной для $\mathfrak{T}(E)$.

Если $E = E_1 \circ E_2$, то корень дерева имеет одну вершину потомка. Эта вершина является корнем дерева $\mathfrak{T}(E_1)$. Если разметка α некоторой вершины этого дерева в $\mathfrak{T}(E)$ объявлена заключительной для $\mathfrak{T}(E_1)$, то в дереве $\mathfrak{T}(E)$ эта вершина имеет потомка. Последний является корнем дерева $\mathfrak{T}(E_2)$. Всякий набор β, объявленный заключительным в дереве с корнем $\mathfrak{T}(E_2)$, является заключительным и в дереве $\mathfrak{T}(E)$. Это набор имеет вид $\alpha\beta$, где α —

набор, объявленный заключительным для дерева $\mathfrak{T}(E_1)$ в вершине, предшествующей корню рассматриваемого вхождения дерева $\mathfrak{T}(E_2)$ в $\mathfrak{T}(E)$.

В последнем случае $\left(E = (E_1)^*\right)$ дерево $\mathfrak{T}(E)$ конструируется так, чтобы заключительные наборы для этого дерева формировались как конечные последовательности слов из $U(E_1)$. Корень дерева $\mathfrak{T}(E)$ размечен выражением E. Этот корень имеет одну вершину потомка, которая размечена выражением E_1. Данная вершина является корнем дерева $\mathfrak{T}(E_1)$. Всякий набор, объявленный заключительным для некоторой вершины рассматриваемого дерева $\mathfrak{T}(E_1)$ объявляется заключительным для $\mathfrak{T}(E)$. В $\mathfrak{T}(E)$ вершина, объявленная заключительной, получает дополнительного потомка, размеченного выражением E_1. Этот потомок является корнем еще одного вхождения $\mathfrak{T}(E_1)$ в $\mathfrak{T}(E)$.

Если разметка $\beta \in I$ некоторой вершины $\mathfrak{T}(E_1)$ объявляется заключительной, то вхождение этой вершины в $\mathfrak{T}(E)$ размечается набором $\alpha\beta \in U(E)$ и эта разметка объявляется заключительной. Здесь $\alpha \in I$ — это разметка, объявленная заключительной для вершины предка корневой вершины рассматриваемого вхождения дерева $\mathfrak{T}(E_1)$ в $\mathfrak{T}(E)$. В дереве $\mathfrak{T}(E)$ вершина ν имеет одного потомка. Последняя вершина размечена выражением E_1 и является корнем следующего вхождения дерева $\mathfrak{T}(E_1)$ в $\mathfrak{T}(E)$. Если некоторая вершина ν' этого вхождения размечена набором $\gamma \in U(E)$, который объявлен заключительным в этом вхождении $\mathfrak{T}(E_1)$ в $\mathfrak{T}(E)$, то эта вершина размечается набором $\alpha\beta\gamma \in U(E)$, который объявляется заключительным для $\mathfrak{T}(E)$.

Процесс конструирования $\mathfrak{T}(E)$ для рассматриваемого случая является бесконечным. Всякий раз, когда в некоторое вхождение $\mathfrak{T}(E_1)$ в $\mathfrak{T}(E)$ добавляется вершина, для которой некоторое $\beta \in I$ объявляется заключительным в этом дереве, такая вершина размечается набором $\alpha_1 ... \alpha_k \beta \in U(E)$, $k \geqslant 0$, который объявляется заключительным в $\mathfrak{T}(E)$. Здесь $\alpha_1 ... \alpha_k$ — набор из $U(E)$, объявленный заключительным для $\mathfrak{T}(E)$ в вершине — предке корня последнего вхождения $\mathfrak{T}(E_1)$ в $\mathfrak{T}(E)$. Общая схема формирования слова $\alpha_1 ... \alpha_k \beta \in U(E)$ для случая $k > 0$ приведена на рис. 1.

Естественный процесс построения дерева $\mathfrak{T}(E)$ для произвольного регулярного выражения E основан на обходе этого дерева (например, в ширину). При этом можно реализовать пересчет элементов множества $U(E)$.

Определение. Границей $U(E)$ называется множество

$$\mathfrak{B}(E) = \left\{\alpha \mid \alpha \in U(E) \,\&\, \alpha 0 \notin U(E) \,\&\, \alpha 1 \notin U(E)\right\}.$$

В общем случае $\mathfrak{B}(E) \subseteq U(E)$. Для всякого регулярного выражения E множество $\mathfrak{B}(E)$ — алгоритмически разрешимо относительно I. Разрешимыми оказываются и многие другие свойства границ множеств, связанных с регулярными выражениями.

5. Регулярные домены морфизмов

Каждое регулярное выражение определяет область памяти $U(E) \subseteq I$. Этой области соответствует множество конфигураций, ПСП которых составлены с использованием вершин из множества $U(E)$.

Определение. ПСП конфигурации $z \in M$ соответствует регулярному выражению E, если для всякого $\alpha \in O(z)$ выполняется одно из условий:

1) $\alpha \in U(E)$;

2) $\alpha \in \left\{\beta_i \mid i \in N \,\&\, \forall i(\beta_i \subset \beta_{i+1})\right\} \subseteq \left[U(E)\right]$;

3) $\alpha = \beta\sigma$, где $\sigma \in \{0, 1\}$, $\beta, \beta\overline{\sigma} \in | U(E) |$, $\beta\overline{\sigma} \in$ $\in | U(E) |$.

То есть, если конфигурация $z \in M$ соответствует регулярному выражению E, то висячая вершина ПСП z:

- является элементом $U(E)$;
- является элементом последовательности вложенных двоичных наборов из $|U(E)|$;
- не принадлежит $|U(E)|$, но является соседним с элементом этого множества.

Проверка условия второго случая приведенного определения реализуема алгоритмически. Пусть $\left(\{0, 1\}, Q, \varphi, q_0, D\right)$ — конечный автомат, распознающий слова множества $U(E)$. Здесь $\{0, 1\}$ — входной алфавит; Q — множество состояний; q_0 — начальное состояние; φ — функция изменения состояний; D — множество распознающих состояний автомата. Проверка рассматриваемого условия по диаграмме переходов автомата сводится к проверке существования бесконечного пути в такой диаграмме, начинающегося из состояния $\varphi(\alpha, q_0)$, содержащего бесконечное множество вхождений состояний, из которых ведут пути в состояния из D.

Если $z \in M$ соответствует регулярному выражению E, то справедливо следующее свойство границы: $\alpha \in \mathfrak{B}(E) \cap D(z) \to \alpha \in O(z)$.

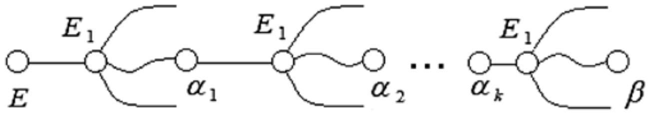

Рис. 1. Фрагмент дерева $\mathfrak{T}(E)$

Определение. Множество $B \subseteq M$ называется регулярным доменом, если существует такое регулярное выражение *E*, что $B = \{z \mid z$ соответствует $E\}$.

Всякий регулярный домен является алгоритмически перечислимым. Регулярный домен знаний в заданном формализме семантических иерархий, соответствующий выражению *E*, будем обозначать как $\mathfrak{D}(E)$. Если домен морфизмов пространств семантических иерархий (формализмов знаний) является регулярным, то формальной спецификацией этого домена является регулярное выражение. Регулярными оказываются домены морфизмов, применявшихся при моделировании операций в формализмах семантических иерархий, а также процессов синтеза реализаций когнитивных целей моделей в ИС [3].

Рассмотрим примеры регулярных доменов морфизмов семантических иерархий.

1. Множество $M(\Sigma)$ всех знаний (ПСП знаний) произвольного формализма семантических иерархий является регулярным доменом. Рассмотрим регулярное выражение $E = (\lambda \cup (0, 1)^*)$. Для этого выражения $U(E) = I$ и поэтому всякая $z \in M$ принадлежит $\mathfrak{D}(E)$. Поскольку $\mathfrak{D}(\delta) = \varnothing$ (δ — пустое выражение), то пустое множество конфигураций также является регулярным.

2. Множество конфигураций M_k (ПСП конфигураций Σ_k), ПСП которых образуют полные бинарные деревья глубины *k*, $k \in \{0, 1, \ldots\}$, также является регулярным. Ему соответствует выражение, составленное как объединение всех выражений, представляющих все двоичные наборы длины *k*. В частности, множества элементарных (M_0) и простых (M_1) знаний — это регулярные домены.

3. Множество серий элементарных знаний (записывается как *S* или $S \otimes M_0$) конфигураций образует регулярный домен. Это множество определяется регулярным выражением $(\lambda \cup (0)^*1)$. Множество вершин бесконечного бинарного дерева (*I*), применяемых в ПСП серий элементарных конфигураций, составляют вершины бесконечной левой ветви этого дерева, а также правые потомки вершин данной ветви. Множества серий знаний заданной глубины $k \in N$ далее будут обозначаться как $S \otimes \Sigma_k$ $(S \otimes M_0)$. Они также являются регулярными доменами.

4. Окрестности элементарных знаний заданной глубины $k \in N$ (O^k) являются доменами морфизмов. Они составляют фрагменты содержания областей знаний, извлекаемых из онтологий таких областей [4]. Регулярность домена окрестностей элементарных знаний глубины 1 (O^1) доказывается с помощью выражения $E = \left(0 \cup \left(1\left(\lambda \cup 1 \cup (0)^*1\right)\right)\right)$. Границу множества двоичных наборов для последнего выражения составляют наборы 0 и 11, 101, 1001, 10001, ... Вершине $0 \in U(E)$ приписываются элементарные конфигурации, для которых составляются окрестности. Остальные наборы границы окрестностей составляют вершины, представляемые выражением $1(1 \cup (0)^*1)$. Вершина 1 является листом в окрестности элементарного знания, если окрестность не содержит элементов.

Регулярность множества окрестностей элементарных знаний произвольной глубины доказывается аналогично. Применяемые для этого регулярные выражения задают схемы конструирования окрестностей. Так, при построении окрестностей глубины 2 вершины границы (серии) в окрестности элементарного знания глубины 1 заменяются на окрестности глубины 1. Рассматриваемый случай реализует регулярное выражение $(0 \cup (1(\lambda \cup ((1 \cup (0)^*1)(0 \cup 1(\lambda \cup (0)^*1))))))$. Такое выражение формируется как продолжение выражения для окрестности глубины 1, продолжаемых выражением для еще одной окрестности глубины 1. При конструировании выражений, представляющих окрестности возрастающей глубины, используется инвариантная запись $(0 \cup 1(\lambda \cup (0)^*1))$. Эта запись определяет множество наборов, которые составляют окрестности глубины 1, замещающие вершины границ при конструировании окрестностей возрастающей глубины.

Для обозначения множеств окрестностей элементарных знаний глубины $k \in N$ далее используются символы O^k. Объединение всех таких множеств будет обозначается как *O*. Комбинации серий и окрестностей позволяют расширить возможности описания структур знаний, синтезируемых процессами в ИС. Основными средствами комбинирования доменов являются операции прямой суммы и произведения. В моделях абстрактных многомерных ИС комбинации доменов составляют основу форматов описания структур памяти компонентов [2]. Форматы определяют структуры знаний, которые размещаются и обрабатываются в памяти ИС.

Теорема 2. Если множества конфигураций B_1 и B_2 являются регулярными, то произведение (сумма) таких множеств является регулярным множеством конфигураций.

Доказательство. В случае, когда хотя бы одно из множеств B_1 и B_2 является пустым — произведение и сумма этих множеств равны пустому множеству и являются регулярными. Рассмотрим случай, когда B_1 и B_2 — это непустые регулярные множества, которые определяются с помощью регулярных выражений E_1 и E_2. Тогда сумма $B_1 \oplus B_2$ представляется регулярным выражением

$0E_1 \cup 1E_2$. Произведение $B_1 \otimes B_2$ представляется выражением $(E_1) \circ (E_2)$. *Доказательство окончено.*

Комбинации регулярных областей памяти определяются специальными формулами. Они позволяют задавать структуры организации памяти компонентов ИС. Каждая формула определяет поддерево дерева I с корнем λ. Такое дерево может быть бесконечным.

Пусть $\mathfrak{S}$ — произвольное алгебраическое выражение (композиция) серий и окрестностей элементарных знаний. Множество двоичных наборов, используемых в качестве вершин — листьев конфигураций, представляемых этим выражением, обозначим как $D(\mathfrak{S})$.

Следующая теорема связана с задачей исследования выразительных возможностей описания структур памяти с помощью регулярных выражений, соответствующих комбинациям сумм и произведений окрестностей и серий.

Теорема 3. Для любой алгебраической комбинации окрестностей и серий элементарных конфигураций $\mathfrak{S}$ найдется конфигурация $z \in M$, для которой $D(z) \cap D(\mathfrak{S}) = \varnothing$.

Доказательство. Для серий (S^1) и окрестностей (O^1) элементарных знаний всякий ярус с номером $k > 2$ бесконечного насыщенного бинарного дерева содержит вершину, не принадлежащую таким сериям (окрестностям). Это так для вершин из соответствующих ярусов, принадлежащих самой правой ветви дерева. Такими вершинами являются 11 и 111.

Если $\mathfrak{S}$ содержит не менее двух вхождений выражений для серий или окрестностей, то рассмотрим последнюю комбинацию операций, примененную при конструировании $\mathfrak{S}$. Для таких вхождений серий (окрестностей) в рассматриваемые выражения возможен один из случаев — $C_1 \oplus C_2$ или $C_1 \otimes C_2$.

Пусть рассматривается конструкция $C_1 \oplus C_2$. Найдется вершина бинарного дерева $\alpha \in I$ для C_i, $i \in \{1, 2\}$, которая не используется в ПСП конфигураций, представимых с помощью C_i. Тогда в конфигурациях, представляемых композицией $C_1 \oplus C_2$ не используется вершина 0α, если $i = 1$, и 1α, если $i = 2$.

Рассмотрим случай композиции $C_1 \otimes C_2$. Пусть $\alpha \in I$ — вершина, не используемая в ПСП конфигураций, представляемых композицией C_2. Пусть $\beta = 1...1$ — самая длинная последовательность, представляющая вершину в ПСП конфигураций, соответствующих C_1. Тогда вершина $\beta\alpha \in I$ не принадлежит ПСП конфигураций, соответствующих $C_1 \otimes C_2$.

При конструировании двоичных последовательностей, не являющихся вершинами конфигураций, представимых композициями окрестностей и серий элементарных конфигураций, длины подходящих последовательностей β возрастают на не более чем 2 для каждой суммы или произведения таких комбинаций. Поэтому, если глубина некоторой алгебраической комбинации C равна k, то длина двоичного набора, соответствующего вершине из I, не принадлежащей ПСП конфигураций, соответствующих C, не превосходит $2k + 1$. *Доказательство окончено.*

Последнее свойство комбинаций прямых сумм и произведений окрестностей и серий элементарных конфигураций позволяет утверждать, что класс регулярных доменов операций шире класса баз операций, формируемых из доменов окрестностей и серий элементарных конфигураций.

6. Конструирование регулярных структур памяти

Полное структурное представление элементов регулярных доменов всякого формализма семантических иерархий составляют вершины из специальных подмножеств множества I. Такие множества составляют области I, в которых размещаются ПСП конфигураций из таких доменов. Многообразие непустых областей I, представляемых регулярными выражениями, ограничено поддеревьями множества I с корнем λ. Расширим это многообразие с помощью переноса корней поддеревьев в произвольные $\alpha \in I$. Пусть E — произвольное регулярное выражение.

Определение. Множество $B \subseteq I$ называется регулярной областью памяти (I), если существует такое регулярное выражение E, что $\exists \alpha \in I(B = \{\alpha\beta \mid \beta \in [U(E)]\})$.

Набор α из последнего определения является корнем поддерева дерева I. Для обозначения рассматриваемой регулярной области I далее будем использовать выражение $\mathfrak{D}(E, \alpha)$. Обозначим многообразие всех регулярных областей в I как $\mathfrak{R}$. Его элементами являются множества I и $\varnothing$. Это многообразие составляют области, применяемые далее для структур памяти компонентов многомерных ИС.

Операции объединения и пересечения регулярных областей (принадлежащих множеству $\mathfrak{R}$) моделируют интеграцию и сужение доменов морфизмов в формализмах семантических иерархий. Если объединяемые области имеют непустое пересечение, то они формируют область памяти, расширяющую множество ПСП конфигураций, которые могут размещаться в области как элементы одного домена и трансформироваться в этой области в знания из другого домена морфизмов знаний.

Пусть $\mathfrak{D}(E, \alpha) \cap \mathfrak{D}(E, \beta) \neq \varnothing$, где $\alpha \neq \beta$ и $\beta = \alpha\gamma$. Тогда для области $\mathfrak{D}(E, \alpha) \cup \mathfrak{D}(E, \beta)$ и домена конфигураций, представимых в этой области, допускаются конфигурации, получаемые заменой фрагментов $(z)_\gamma$ конфигураций, размещаемых в $\mathfrak{D}(E, \alpha)$, на конфигурации, размещаемые в области $\mathfrak{D}(E, \beta)$. В частности, если $\gamma \in \mathfrak{B}(E)$, то такая замена выполняется для $\gamma \in O(z)$.

Объединения непересекающихся областей из $\mathfrak{R}$ таким свойством не обладают. Конфигурации, формируемые в отдельных областях объединения, составляют пары семантических иерархий, размещенных в объединяемых областях, которые могут рассматриваться как прямые суммы конфигураций.

Произвольные композиции регулярных доменов морфизмов определяют структуры памяти компонентов многомерных ИС, допускающие разные варианты использования отдельных подобластей памяти операциями обработки знаний, размещаемых в этих областях. Система областей проектируется для времени существования ИС и поддерживает реализацию жизненных циклов таких систем. Развертывание системы описаний системы областей памяти компонентов является частью технологии управления моделями ИС. Уточнение описаний структур областей на основе описания регулярных доменов способствует большей специализации и эффективности алгоритмов реализации операций, учитывающих структурные особенности обрабатываемых знаний.

Рассмотрим пример моделирования памяти для компонента алгоритмического уровня полностью структурированного представления содержания области знаний. В двумерной архитектуре ИС с измерениями абстрактности (*поверхностный* — А, *алгоритмический* — В, *когнитивный* — *С*) и структурированности знаний (*целостные образы* — 1, *частично структурированные* — 2 и *полностью структурированные* или *атомарные* — 3) этот компонент обозначается как В-3 [2]. Соседи этого компонента — компоненты А-3, В-2, С-3 (рис. 2).

На рис. 2 изображен дополнительный компонент А-0. Он соответствует внешнему окружению ИС или области деятельности, к которой относится ИС. Содержание памяти А-0 составляют разнообразные теоретические и эмпирические знания, представленные в форматах, принятых в области знаний. Взаимодействие ИС с А-0 осуществляется потоками знаний, реализуемыми между компонентами А-0 и А-1. В последнем выполняется

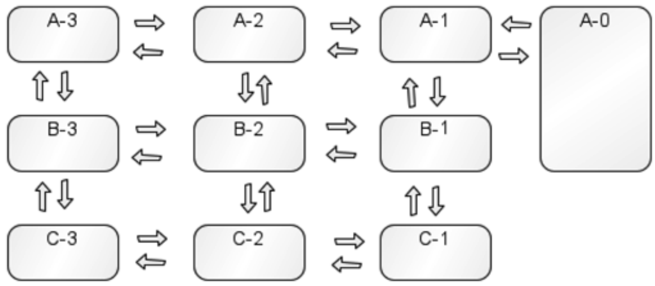

Рис. 2. Компоненты и связи компонентов двумерной ИС

начальная обработка знаний, поступающих в ИС. Обратный перенос знаний из А-1 в А-0 реализует обмен результатами реализаций целей ИС с А-0.

Потоки знаний в рассматриваемой архитектуре основаны на переносе фрагментов содержания областей памяти между соседними компонентами. Варианты межкомпонентных потоков знаний изображаются стрелками. При переносе знаний выполняется трансформация формата формализма компонента передаваемого знания в формат формализма знаний следующего компонента [2, 5].

Модель неструктурированной области памяти для компонента В-3 определяет ее с помощью выражения $\mathfrak{D}((\lambda \cup (0,1)^*), \lambda)$. Такая область совпадает с множеством *I*. Операции и процессы в В-3 реализуют цели извлечения новых знаний из входного потока структурированных данных, поступающего из А-3, распознавания и формирования целей и задач, предполагающих решение в ИС. Такие данные извлекаются из описаний ситуаций в А-0, поступающих в А-1.

Структуру памяти В-3 задают следующие виды знаний:

1) серия серий простых знаний, представляющих начальные данные, последовательно обрабатываемых в А-1, А-2 и А-3, которые переносятся в В-3;

2) серия простых знаний (онтология предметной области в формате семантических иерархий компонента В-3);

3) серии серий знаний, содержащие исходные данные и фрагменты онтологии области знаний, переносимые в В-2 в целях реализации процессов решения задач.

Обозначим эти области как *D*1, *D*2, *D*3. Тогда структура памяти В-3 может быть определена с помощью формулы $(D1 \oplus D2) \oplus D3$.

Структура каждой области уточняет форматы представления знаний, учитывающие операции формирования соответствующих знаний в В-3. Поток знаний, поступающих в В-3 из А-3, составляют фрагменты внешней или *E*-онтологии текущей ситуации в области знаний. Они представляются сериями простых знаний, получаемых

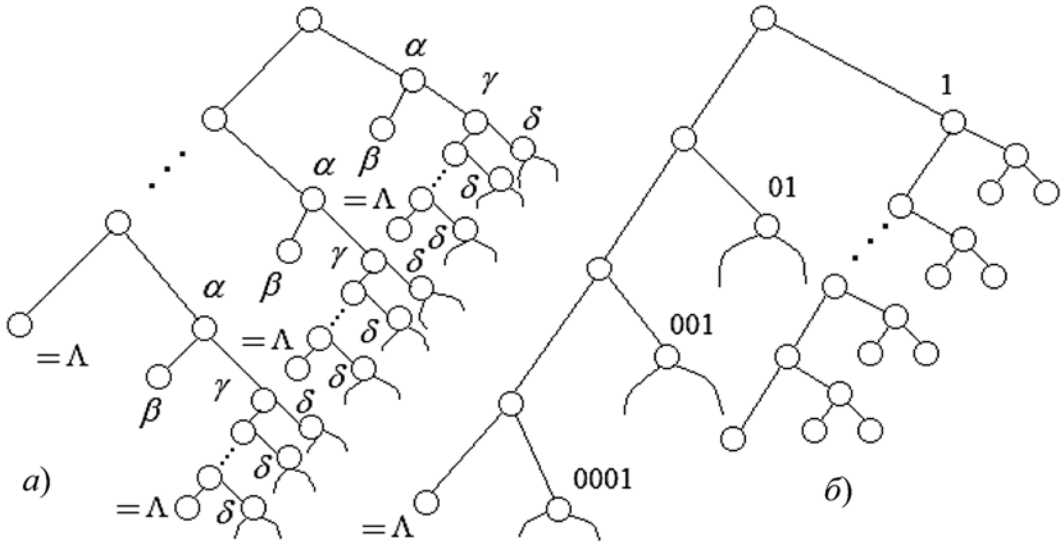

Рис. 3. Структуры областей памяти компонента В-3

декомпозицией отдельных ситуаций. Всякая серия дополняется описанием свойств серии, распознанных при обработке начальных данных в А-1, А-2 и А-3. Из отдельных серий составляется серия серий. Структуру области памяти *D*1 для последней серии (изображена на рис. 3, *а*) определяет выражение

$$\mathfrak{D}(\lambda \cup ((\lambda \cup (0)^*) \circ 1 \circ (0 \cup 1 \circ (1 \cup (0) * 1))), 00).$$

Символ α на рис. 3 обозначает корневые вершины отдельных серий, символ β — вершины, размеченные идентификаторами отдельных групп. Символы γ и δ обозначают соответственно корни серий простых знаний и отдельных простых знаний в таких сериях. Структура описаний групп в вершинах β может быть усложнена. Например, если заменить листья семантических иерархий потока из А-3 в В-3 на серии элементарных знаний, составляющих представление содержания знаний, приписываемых листьям.

Стандартная структура области памяти для онтологии содержания моделируемой области знаний (*D*2) соответствует серии простых знаний. Такие знания формируют описания классов элементарных знаний и отношений между классами, а также разделов онтологии. Например, это могут быть разделы представления содержания *понятий*, *законов*, *задач* и *методов* в соответствующей области знаний. Отношения между классами допускают принадлежность классов разным разделам онтологии. Содержание отдельного раздела онтологии задается серией простых знаний. Такие знания представляют элементы классов и отношения между классами. Для случая, когда классы относятся к разным областям онтологии, отношения между классами представляются в области, к которой относится первый из классов, связываемых отношением. Если $E = (\lambda \cup (1 \circ (1 \circ (0 \cup 1) \cup (0)^* 1 \circ (0 \cup 1))))$ — регулярное выражение, представляющее серию простых знаний, то четыре такие серии, представленные областями $\mathfrak{D}(E, 1)$, $\mathfrak{D}(E, 01)$, $\mathfrak{D}(E, 001)$ и $\mathfrak{D}(E, 0001)$, интегрируются в описание области памяти $\mathfrak{D}(1 \circ E \cup 01 \circ E \cup 001 \circ E \cup 0001E, \lambda)$. Последняя область изображена на рис. 3, *б*. Она соответствует домену морфизмов знаний. Структура памяти этого домена определяется выражением $(((SM_1 \oplus SM_1) \oplus SM_1) \oplus SM_1)$.

На рис. 3, *б* вершины 1, 01, 001 и 0001 являются корнями четырех серий, которые имеют одинаковую структуру. В сериях размещается описание содержания отдельных разделов онтологии. Первая из серий изображена детально и состоит из простых знаний.

Рассмотрим описание области памяти, в которой синтезируются структуры знаний, переносимых затем в В-2 для реализации процессов конструирования реализаций целей ИС. Формат этой структуры соответствует сумме серий простых знаний.

Процесс формирования семейств простых знаний, подготавливаемых для переноса в В-2, реализуется в памяти компонента В-3, обозначенной как *D*3. Он включает прохождение нескольких этапов обработки отдельных серий простых знаний (описаний ситуаций), переносимых в В-3 из А-3.

Конструируемые в *D*3 элементы серий знаний интегрируют семейства простых знаний, достаточные для синтеза реализаций целей ИС. Такие знания содержат серии начальных данных реализуемых целей и необходимые фрагменты содержания (онтологии) области знаний. При этом фрагмент

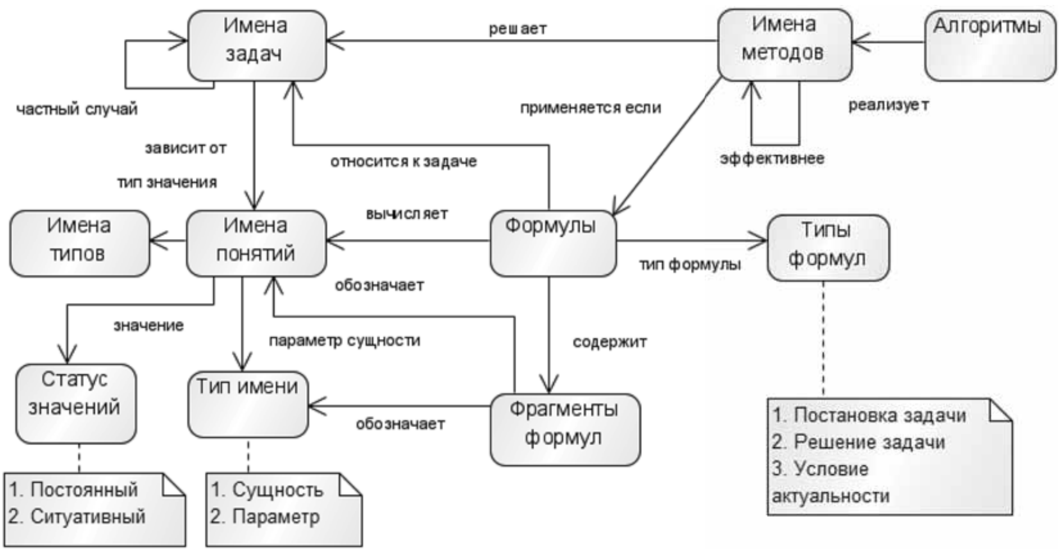

Рис. 4. Фрагмент онтологии профессиональных знаний

онтологии извлекается из $D2$, а семейства начальных данных целей извлекается из $D1$.

Рассмотрим пример карты знаний, представляющей общий фрагмент онтологии задач. Классы знаний, составляющие данную карту, группируются в области *понятий*, *формул*, *задач* и *методов*. Пример фрагмента карты знаний для моделирования процесса формирования семантической структуры, подготавливаемой к передаче в В-2, приведен на рис. 4. Классы и отношения, составляющие данную карту знаний, позволяют формировать фрагменты онтологии содержания области знаний, относящиеся к целям, активируемым сериями начальных данных. Такие фрагменты включают описания методов идентификации и реализации целей (задач) ИС, соответствующих сериям простых знаний, составляющих результат декомпозиции описаний отдельных ситуаций.

Приведенная карта знаний применима для разных предметных областей. Элементарные знания для фрагмента онтологии задаются именами и формулами. Связная структура таких знаний определяет формализованное содержание предметной области. Математические формулы разного типа и алгоритмы реализации методов реализации целей обрабатываются механизмами вывода. Последние моделируют схемы мышления в структурах памяти.

Процесс формирования структур знаний, передаваемых в В-2, включает распознавание (идентификацию) целей (задач), нахождение метода (алгоритма) ее решения. Для этого применяются элементы классов "Имена задач", "Имена понятий" и "Формулы". Серии начальных данных задают ограничения для параметров области знаний, представленных в классе "Имена понятий". Ограничения применяются для нахождения значений формул распознавания задач и методов их решения. Весь процесс реализуется в области памяти $D3$ компонента В-3.

Вариант такого процесса составляют следующие действия:

1) извлечение серии формул, являющихся условиями активации отдельных задач (извлекаются из класса "Формулы" с помощью подходящего морфизма селекции);

2) замена каждой такой формулы на окрестность параметров формулы (извлекаются из класса "Имена понятий");

3) поиск значений параметров формул условий активации профессиональных задач (извлекаются из серии начальных данных ситуации, содержащейся в $D1$);

4) вычисление значений недостающих параметров формул активации целей (задач) с помощью элементов классов "Формулы" и "Имена понятий";

5) вычисление значений формул, для которых достаточно начальных данных, и идентификация задач, которые требуют решения (по выражениям из класса "Формулы");

6) выбор методов (алгоритмов) решения распознанных задач (с помощью классов "Имена задач" и "Имена методов");

7) построение серий значений параметров для отдельных отобранных методов (алгоритмов) и нахождение (вычисление) значений недостающих понятий с помощью начальных данных и формул, вычисления значений понятий (аналогично действию 4 с помощью классов "Формулы" и "Имена понятий").

Формат области памяти, в которой составляется отдельное знание, переносимое в В-2 для синтеза решения цели (задачи), представленной таким знанием, имеет вид $(M_0 \oplus M_0) \oplus (S \otimes M_1)$. Здесь $M_0 \oplus M_0$ — имена задачи и выбранного метода ее решения; $S \otimes M_1$ — серия начальных данных метода, составляющих значения и условия на параметры метода.

Формула структуры памяти области $D3$ включает фрагмент, определяющий структуру области памяти, в которой вычисляются значения параметров условий актуальности задач и параметров выбранных методов решения. Обозначим такой фрагмент как A. Тогда формула области памяти в $D3$, используемой для синтеза знаний, передаваемых в В-2, представляется как $(M_0 \oplus M_0) \oplus (S \otimes M_1) \oplus A$. Фрагмент $S \otimes M_1$ последнего выражения — это серия простых знаний, являющихся исходными данными цели (задачи). Серия переносятся из $D1$ после идентификации цели (задачи). Конструкция A представляет унифицированное описание структуры памяти, применяемое общей схемой обработки фрагмента онтологии, приводящего к нахождению значений параметров, передаваемых в В-2 и необходимых для реализации цели (задачи).

Приемлемую структуру памяти области A задает выражение $B_1 \oplus B_2$. Здесь первое слагаемое определяет область памяти, в которой моделируются действия 1) — 5), а второе — области памяти, в которой реализуются действия 6) — 7). В B_1 реализуется рекурсивная схема нахождения значений параметров формул, условий активации профессиональных задач. Процесс в этой области связан с проверкой выполнимости условий активации отдельных задач. Для этого всякой цели ставится в соответствие окрестность глубины 1 (O^1), составленная формулами активации. Формулы дополняются значениями (M_1) или структурами, позволяющими найти эти значения, с использованием значений параметров формул. Списки параметров извлекаются из класса "Имена понятий". Значения параметров извлекаются из серий начальных данных. Для этого формулы трансформируются в окрестности формул глубины 1, границы которых образуют имена параметров формул. Далее параметры либо заменяются на их значения (извлекаемые из начальных данных), либо вычисляются с использованием элементов класса "Формулы" и серий начальных данных целей (задач).

Параметры, значения которых не представлены начальными данными (M_1), дополняются сериями окрестностей параметров глубины 2 (O^2). Окрестности каждой серии составляют формулы вычисления значений параметров, а затем серии параметров каждой такой формулы. Серии формул и параметров формул извлекаются из классов "Имена параметров" и "Формулы".

Затем параметрам ставятся в соответствие их значения, извлекаемые из серий начальных данных. Для параметров, значения которых отсутствуют, повторяются приведенные действия. Для этого применяется область памяти, формат которой определяет выражение $M_1 \cup S \otimes O^2$. Если для некоторой формулы всем параметрам этой формулы поставлены в соответствие значения, то выполняется вычисление значения формулы и замена развернутой серии окрестностей формул, вычисляющих заданное понятие, на простое знание о значении понятия (M_1).

Приведенный процесс соответствует схеме, формальное описание которой содержится в работе [4]. Глубина формируемого дерева структурного представления знания ограничивается условием запрета на продолжения построения, если для вычисления значения некоторого понятия делается попытка вычисления этого же понятия. Формула памяти для рассмотренного процесса имеет вид

$$S \otimes (O^1 \otimes (M_1 \cup (S \otimes (O^1 \otimes (M_1 \cup S \otimes O^2)^*)))).$$

Структуру области памяти B_2 определяет аналогичная формула. Она имеет вид

$$S \otimes (O^2 \otimes (M_1 \cup (S \otimes (O^2 \otimes (M_1 \cup S \otimes O^2)^*)))).$$

В этой области формируется серия окрестностей глубины 2. Окрестности составляются для целей (задач), необходимость активации которых установлена в B_1. Задачам соответствуют методы решения, а методам — параметры, значения которых необходимы для реализации методов. Параметры заменяются на значения, которые либо содержатся в серии начальных данных (M_1), либо вычисляются аналогично рассмотренному процессу в B_1. Для этого применяется структура памяти $S \otimes \left(O^2 \otimes \left(M_1 \cup S \otimes O^2\right)^*\right)$.

Заключение

Класс регулярных доменов морфизмов знаний позволяет разрабатывать описания структур памяти, используемых семействами процессов синтеза знаний в компонентах ИС. Этого класса достаточно для высокоуровневого моделирования диаграмм процессов, представляемых диаграммами морфизмов и доменов морфизмов формализмов знаний для аналогов функциональных сущностей из разных областей математики, развивающих принципы подхода теории категорий к моделированию процессов [6]. Многообразие регулярных областей определяет топологическую структуру памяти ИС. Это позволяет использовать понятия и методы общей топологии для исследования и проектирования схем управления памятью в таких системах. Модели иерархических структур памяти для знаний в формате семантических иерархий допускают использование в качестве элемента формализации и моделирования концепции соотношения семантической памяти и сознания, предложенной в работе [7].

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 20-01-00289, а также РФФИ и администрации Краснодарского края в рамках научного проекта 19-41-230008 р_а.

Список литературы

1. **Костенко К. И.** Операции когнитивного синтеза формализованных знаний // Программная инженерия. 2018. Том 9, № 4. С. 174—184.
2. **Kostenko K.** Knowledge flows processes at multidimensional intelligent systems // Russian Advances in Artificial Intelligence: selected contributions to the Russian Conference on Artificial intelligence (RCAI 2020). October 10—16, 2020, Moscow. 2020. Vol. 2648. P. 74—84.
3. **Stanovich K. E.** Rationality and the reactive mind. Oxford. Univ. Press, 2010. 344 p.
4. **Костенко К. И.** Диаграммы процессов и правила синтеза знаний из элементов онтологий // VIII Международная научная конференция "Знания—Онтологии—Теории", 8—12 ноября 2021. Новосибирск, 2021. С. 131—140.
5. **Burgin M.** Theory of Knowledge: Structures and Processes. World Scientific, 2017. 948 p.
6. **Ковалёв С. П.** Теоретико-категорный подход к проектированию программных систем // Фундаментальная и прикладная математика. 2014. Том 19, № 3. С. 111—170.
7. **Анохин К. В.** Когнитом: в поисках фундаментальной нейронаучной теории сознания // Журнал высшей нервной деятельности им. И. П. Павлова. 2021. Том 71, № 1. С. 39—71.

Regular Memory Structures and Domains Descriptions for Operations in Intelligent Systems

K. I. Kostenko, kostenko@kubsu.ru, Kuban State University, Krasnodar, 350040, Russian Federation

Corresponding author:

Kostenko Konstantin I., Associatet Professor, Kuban State University, Krasnodar, 350040, Russian Federation,
E-mail: kostenko@kubsu.ru

Received on March 01, 2022
Accepted on March 13, 2022

The concept of a regular memory area of an artificial intelligence system defined. Semantic hierarchies are considered as a uniform and universal format for structural descriptions of knowledge representations. An infinite saturated binary tree, the vertices of which are binary sequences, implements a single memory format. Finite saturated binary trees with labeled vertices represent individual knowledge in infinite memory. Regular expressions are used as a specification for the memory areas in question. Regular memory domains generalize the system of morphism domain class developed for semantic hierarchy formalisms. The study of regular memory structures and the choice of special tools for modeling memory and thought processes in artificial intelligence systems carried out. The structure of regular memory is associated with classes of knowledge processing operations, as well as the goals of classifying, storing and applying knowledge of a given structure in such systems. Applications of regular memory structures creates conditions for a deep formalization of the tools of constructing memory structures and modeling thought processes in intelligent systems. Associating regular expressions with state machines allows using them as a parameter for effective modeling of the intelligent processes.

Keywords: *knowledge formalisms, knowledge morphism, morphism domain, memory regular structure, memory area, memory structure formulae*

Acknowledgements: *This work was funded by RFBR and administration of Krasnodar territory grant project number № 19-41-230008 and by RFBR grant project number № 20-01-00289.*

For citation:

Kostenko K. I. Regular Memory Structures and Domain Descriptions for Operations in Intelligence Systems, *Programmnaya Ingeneria,* 2022, vol. 13, no. 5, pp. 226—238.

DOI: 10.17587/prin.13.226-238

References

1. **Kostenko K. I.** Operations of formalized knowledge cognitive synthesis, *Programmnaya Ingeneria*, 2018, vol. 9, no. 4, pp. 174—184 (in Russian).
2. **Kostenko K.** Knowledge flows processes at multidimensional intelligent systems, *Russian Advances in Artificial Intelligence: selected contributions to the Russian Conference on Artificial intelligence (RCAI 2020)*, October 10—16, 2020, Moscow, 2020, vol. 2648, pp. 74—84.
3. **Stanovich K. E.** *Rationality and the reactive mind*, Oxford Univ. Press, 2010, 344 p.
4. **Kostenko K. I.** Processes diagrams for knowledge synthesis based on ontologies, *VIII International Scientific conference "Knowledge—Ontologies—Theories"*, 8—12 November, 2021, Novosibirsk, 2021, pp. 131—140 (in Russian).
5. **Burgin M.** *Theory of Knowledge: Structures and Processes*, World Scientific, 2017, 948 p.
6. **Kovalyov S. P.** Category-theoretic approach to software systems design, *Fundamentalnaya i prikladnaya matematika*, 2014, vol. 19, no. 3, pp. 111— 170 (in Russian).
7. **Anokhin K. V.** Cognitome: in search of fundamental neuroscience theory of consciousness, *Zhurnal vysshej nervnoj dejatel'nosti im. I. P. Pavlova,* 2021, vol. 71, no. 1, pp. 39—71 (in Russian).