

Проблемы построения картины мира для задач осмысленной обработки информации

Problems of construction of a picture of the world for the purpose of intelligent information processing

Александров Сергей Эдуардович, Фадеев Павел Эдвардович
Компания “Интеллектуальный Партнёр”, г. Минск, Беларусь
<http://www.intelpart.com/>

Аннотация

Доклад продолжает начатое в работе [1] обсуждение и последовательное решение проблем автоматического оперирования смыслами.

Представляются действующие программные продукты, продолжающие линию программы “Грамматический процессор”, продемонстрированной на конференции “Диалог-2003”. Новое поколение программ (“Генератор тезаурусов” и “Генератор Моделей предметных областей”) позволяет автоматически создавать Тезаурусы из текстовых библиотек и создавать Модели предметных областей без участия человека.

Рассматривается решение проблемы увязки различных Моделей предметных областей в целостную систему знаний. Противоречие, лежащее в основе этой проблемы (с одной стороны система знаний безгранична, с другой, чтобы ею управлять, она должна быть ограничена), предлагается разрешать следующим образом: любая информация позиционируется в системе знаний путем соотнесения ее с теми или иными взаимосвязанными целям (средствами) субъекта. Предлагается способ организации и описывается структура “субъектно-ориентированной” (субъектно-подобной) картины мира.

На обсуждение выносятся возможности программ с субъектно-ориентированной картиной мира для решения различных научных и коммерческих задач.

The report presents a continuation of discussion of problems of automatic meanings handling started in the publication [1] and suggests consecutive solution of same. Presenting our new operating software products we continue the “Grammar Processor” line presented at the conference “Dialogue-2003”. New generation of the programs (“Thesaurus Generator” and “Generator of Knowledge Domain Models”) makes it possible to create thesauruses out of text libraries automatically and create Knowledge Domain Models without human participation.

The problem of associating various Knowledge Domain Models into an integrated knowledge system is scrutinized. Resolution of the underlying basic conflict of this problem (on the one hand the knowledge system must be unlimited, on the other hand a controllable system must have limits) is suggested as follows: all information is positioned within the knowledge system via its correlation with certain interrelated aims (means) of the subject. Means of organization of a “subject-oriented” (subject-like) picture of the world is suggested and its structure is described.

Resources offered by programs with subject-oriented picture of the world for pursuing of various research and commercial goals are submitted for discussion.

Введение

Рекурсивный принцип построения тезауруса, когда грамматический процессор с эффективностью 70% создает первичный словарь элементарных смыслов, а сам словарь, в свою очередь, повышает точность уже до 90-95%, позволяет строить словари смыслов (тезаурусы) и, путем организации взаимодействия этих смыслов, автоматически строить конкретные модели предметных областей (МПО) по набору текстовых источников.

Решение проблемы автоматической генерации моделей предметных областей приводит к тому, что мы получаем возможность создать библиотеку моделей предметных областей, каждая из которых ориентирована на решение определённых конкретных задач. Здесь возникает потребность работать не с одной МПО, а со всей библиотекой сразу. Однако при попытке прямого объединения между собой различных МПО возникают противоречия.

- 1) **Противоречие 1**, структурное, онтологическое, присущее и самим частным МПО — это противоречивость самих знаний. Например, каким образом в одной МПО отразить факт дуализма света. С одной стороны свет — это волна, с другой — корпускула. И оба эти представления доказаны и совершенно правомерны. Таким образом, если быть последовательными, то требуется иметь две различные МПО: одна — для свет-волна, другая — для свет-корпускула.
- 2) **Противоречие 2**, временно^е, гносеологическое — противоречие контекстов времени, в которых находится одно и то же знание. Например, слово “фараон” в древние времена служило для обозначения правителя имело один смысл, но сейчас этим словом называют полицейского, а также карточную игру.
- 3) **Противоречие 3** — противоречие между моделями предметных областей. Два факта из разных МПО могут иметь между собой бесконечное число отношений. Например, “в огороде — бузина, а в Киеве — дядька”. Для многих людей этот фразеологический оборот означает, что два факта не имеют между собой абсолютно никакой связи. Но, с другой стороны, обязательно найдутся те, у кого в огороде растёт бузина, а в Киеве живёт дядя. То есть связь всё-таки есть. Возникает вопрос: как данные факты должны быть отображены в МПО? Если исходить из требования непротиворечивости МПО, то мы вынуждены допустить, что прямое объединение двух МПО может породить бесконечное количество МПО-посредников.

Таким образом, возникает задача создания принципов и механизмов, позволяющих оперировать взаимодействием различных моделей предметных областей, подобно тому, как происходит оперирование взаимодействием различных модулей элементарных смыслов [1, 2].

В докладе предлагается решение этой задачи в рамках структурного (онтологического) контекста [5]. Описываются принципы и механизмы, позволившие на технологическом, программном уровне решить указанную задачу при создании коммерческих программных продуктов: “Дельта-Персонал”, “Приёмы менеджмента”, “Генератор тезаурусов” и “Генератор моделей предметных областей”.

1. Элементарный структурный контекст. Представление контекста

Обработка информации на основе модуля элементарного смысла [1, 2] не учитывает то, что понятия приобретают разный смысл в зависимости от контекста, в котором они находятся.

Чтобы уйти от ситуации бесконечного количества контекстов, в которых может находиться любой элементарный смысл, удобно ограничиться *субъектно-ориентированным элементарным контекстом (модулем элементарного смысла в контексте)*.

Для того чтобы оперировать элементарным смыслом в информационном пространстве с учетом элементарных контекстов необходимо иметь:

- Ключевые контексты, свойственные пользователю (контексты, которые задаёт пользователь).
- Принципы связи и взаимодействия контекстов между собой.
- Структуру представления контекстов.

Представить понятие с учетом контекста можно различными способами, но удобно использовать его минимальное (элементарное) представление. Для этого *определим контекст как элементарный момент изменения элементарного смысла*. Момент изменения может быть как структурным (изменение связи с другими понятиями), так и временным (элементарное изменение самого понятия во времени). В этом случае, чтобы отобразить смысл в контексте требуется:

- 1) Указать название понятия.
- 2) Указать имманентные (онтологические) границы понятия. “ОТ” чего начинается изменение понятия, “ЧЕРЕЗ” что оно изменяется, “К” чему оно приходит (системный контекст).
- 3) Указать внутреннее элементарное движения понятия — подсистемный контекст понятия: удобен и требуется для того, чтобы в дальнейшем иметь возможность представлять внутреннюю структуру контекстов.
- 4) Указать внешнее элементарное движение понятия — надсистемный контекст понятия: удобен и требуется для того, чтобы в дальнейшем иметь возможность представлять внешнюю структуру контекстов.

Тогда модуль элементарного контекста выглядит следующим образом (рис. 1).

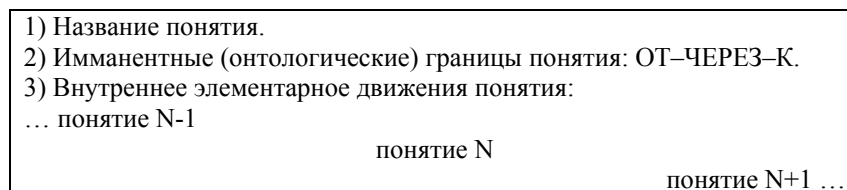


Рисунок 1. Модуль элементарного контекста понятия.

Сам модуль элементарного контекста является одной из фаз внешнего элементарного движения понятия (рис. 2).

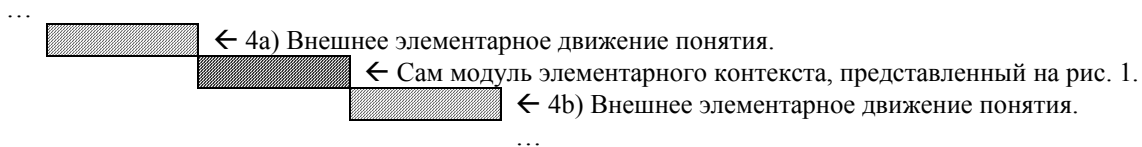


Рисунок 2. Фазы внешнего элементарного движения понятия.

Механика оперирования контекстами заключается в том, что для представления понятия в некоем контексте необходимо:

- 1) Взять модуль элементарного смысла рассматриваемого понятия [1].
- 2) Построить модуль элементарного контекста для рассматриваемого понятия.
- 3) Выбрать в созданном модуле элементарного контекста тот контекст, который интересует в данном конкретном случае и который, кстати, является понятием, а значит, также представим в виде модуля элементарного смысла.
- 4) Наложить выбранный контекст на модуль элементарного смысла рассматриваемого понятия и по процедурам, описанным в [1], уточнить рассматриваемое понятие в рамках выбранного контекста. Иными словами, происходит объединение рассматриваемого понятия и контекста.

2. Оперирование элементарным структурным контекстом. Взаимодействие понятий в контекстах. Структурная картина мира

Описанный подход к оперированию контекстами имеет недостатки, поскольку не учитывает, что:

- 1) У любого понятия может быть неограниченное количество контекстов.
- 2) Отношения между контекстами многозначны. Например, контекст А в контексте В, при наличии контекста С, через контекст D...
- 3) Если информация в информационном пространстве представлена частями и в разных местах, то это не позволяет оперировать ею как единым смысловым целым.

Для того чтобы разрешить указанные проблемы, необходимо:

- Представить понятия и контексты в виде некоей ограниченной системы отношений.
- Иметь возможность изменять представление отношений понятий и контекстов при их онтологических изменениях.
- Оценивать и формировать отношения понятий и контекстов в соответствии с контекстами целей информационной системы.

Информационную систему, удовлетворяющую перечисленным выше требованиям, на практике удобно строить исходя из следующих принципов:

- Задаться некоторым максимально широким контекстом (неким предельно полным явлением), с которым соотносятся любые другие понятия и контексты. Например, принцип позиционирования в космическом пространстве, когда координаты объекта определяются по далеким звёздам, положение которых считается константным (хотя на самом деле это не так). Таким предельно полным явлением, для целей информационных систем, может служить такое понятие, как “человек” и основные этапы его жизни. В этом случае, например, такие явления как “математика” и “физика” получают конкретные отношения друг с другом либо в контексте начальной школы, либо в контексте вузовской программы, либо в контексте производственной или научной деятельности и так далее.
- Структура отношений понятий и контекстов разворачивается в соответствии с процессом (генезисом) формирования системы контекстов и отношений субъекта.

Конкретный модуль элементарного контекста понятия “человек” можно рассматривать и использовать как предельно широкую, но конкретную (операбельную) информационную структуру, где *любая* информация имеет своё конкретное место внутри такого модуля. Поэтому модуль элементарного контекста понятия “человек” можно рассматривать как *субъектно-ориентированную картину мира*.

Использование субъектно-ориентированной картины мира позволяет разрешить указанные во введении противоречия. Так противоречивость знаний разрешается за счёт указания и позиционирования условий (контекстов), в которых проявляются одинаковые явления. Аналогично разрешаются противоречия временны́е и противоречия между различными МПО. Кроме того, разрешение противоречий между различными МПО позволяет получать информацию, которая явно в МПО не содержится.

3. Практические особенности объединения различных МПО. Восстановление информации, которая явно в картине мира не содержится

Особенности объединения различных МПО рассмотрим на примере коммерческой программы “Дельта-Персонал”. Программа оперирует 4096 элементарными моделями предметных областей, каждая из которых содержит проявления и особенности типовых характеров в контексте деловых качеств сотрудника.

Например, есть сотрудник, обладающий характером — “уравновешенный” и проявляющий себя на работе следующим образом — “уклоняется от работы”. Соответствующая указанным деловым качествам МПО содержит свойственные характеру сотрудника ошибки и управляющие действия, позволяющие эти ошибки устранить. Однако одновременно руководитель видит в сотруднике и многие другие черты. Например, “консервативность” и “надменность”. Возникает задача не только объединения деловых качеств: “уравновешенный” + “консервативный” + “надменный”, но и объединения соответствующих МПО, поскольку именно такой подход позволяет адекватно выявить наиболее эффективные способы управления подчинённым.

Результатом объединения этих трёх МПО в программе “Дельта-Персонал” является интегральная модель предметной области, не сводимая ни к одной исходной, а именно — “нежелание слабого сотрудника (эффективно работать)”. Действительно, если проанализировать полученный результат, то приходим к выводу, что именно “слабость” подчинённого порождает такие качества, как “консервативность”, “надменность” и, в определенной степени, “уравновешенность”. В итоге для данной ситуации руководитель получает ошибки и управляющие действия, свойственные именно этому случаю, а не какому-нибудь другому (рис. 3, рис. 4).

- | |
|--|
| <ol style="list-style-type: none">1) НЕЛЬЗЯ! Категорически настаивать на получении сотрудником результатов в нужной форме и нужными методами при этом, даже формально не считаясь с его мнением и точкой зрения, что препятствует объективному анализу сотрудником своих ошибок.2) НЕЛЬЗЯ! Третировать и шантажировать сотрудника, используя информацию о его неудачах или ошибках, что спровоцирует зависть сотрудника к преуспевающим коллегам и стремление всеми силами помешать их эффективной работе.3) НЕЛЬЗЯ! Компенсировать отсутствие результатов в работе сотрудника введением новых сотрудников, что не дает сотруднику возможности найти в себе, сформировать и использовать сильные качества. |
|--|

Рисунок 3. Рекомендации программы “Дельта-Персонал”: ошибки управления сотрудником.

- 1) СЛЕДУЕТ. Определить и показать выгоду сотрудника, в случае успешного достижения требуемых результатов; отдавать приоритеты таким заданиям для сотрудника, которые ему понятны и интересны.
- 2) СЛЕДУЕТ. Поручить сотруднику работу, выполняя которую он может стабильно получать необходимые результаты; привлечь к взаимодействию единомышленников сотрудника, предоставив сотруднику возможность самостоятельно подбирать людей для реализации тех или иных задач.
- 3) СЛЕДУЕТ. Сформулировать какой конкретно результат, в каком виде и в какие сроки должен быть получен сотрудником; устранить возможности сотрудника лавировать и манипулировать обстоятельствами выполнения выданных ему заданий.

Рисунок 4. Рекомендации программы “Дельта-Персонал”: способы управления сотрудником.

Однако, оперирование локальными и замкнутыми МПО не столь интересно по сравнению с возможностью организации саморазвивающейся структуры модуля элементарного контекста “человек” и превращения его в субъектно-ориентированную картину мира (СКМ). Рассмотрим возможную организацию такой картины мира. Здесь следует напомнить, что отдельные части модуля элементарного контекста легко представляются в виде модуля элементарного смысла, который уже имеет программную реализацию [1, 2]. Это позволяет перейти к программной реализации субъектно-ориентированной картины мира.

Рассмотрим механизмы и процедуры оперирования модулем элементарного контекста “человек”. Если представить понятие “человек” в виде явлений, соответствующих основным функциональным возрастам человека (основным качественным этапам), то внутренняя структура модуля элементарного контекста для понятия “человек” приобретает следующий вид (рис. 5).

1.1 “младенец”	2.1	3.1	4.1	5.1	6.1	7.1	8.1	9.1
1.2	2.2 “ребёнок”	3.2	4.2	5.2	6.2	7.2	8.2	9.2
1.3	2.3	3.3 “подросток”	4.3	5.3	6.3	7.3	8.3	9.3
1.4	2.4	3.4	4.4 “родитель”	5.4	6.4	7.4	8.4	9.4
1.5	2.5	3.5	4.5	5.5 “семьянин”	6.5	7.5	8.5	9.5
1.6	2.6	3.6	4.6	5.6	6.6 “воспитатель”	7.6	8.6	9.6
1.7	2.7	3.7	4.7	5.7	6.7	7.7 “родитель родителя”	8.7	9.7
1.8	2.8	3.8	4.8	4.8	6.8	7.8	8.8 “родитель семьи”	9.8
1.9	2.9	3.9	4.9	4.9	6.9	7.9	8.9	9.9 “родитель воспитателя”

Рисунок 5. Внутренняя структура модуля элементарного контекста для понятия “человек”.

Каждая из клеток сама является модулем элементарного контекста и имеет аналогичную разворачивающуюся внутрь структуру.

Пример возможной развёртки этапа “младенец” (рис. 6).

“младенец” → “зародыш”

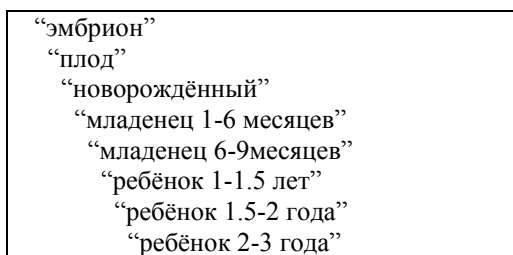


Рисунок 6. Пример возможной развёртки этапа “младенец”.

Примечание. Правильность и точность развёртки той или иной фазы модуля элементарного контекста не является критичным, поскольку в ходе оперирования субъектно-ориентированной картиной мира постоянно происходит её корректировка и уточнение.

Кроме того, отдельные фрагменты картины мира взаимодействуют между собой. Например, фаза “родитель” непосредственно взаимодействует (резонирует) с фазой “младенец” (рис. 7):

1.1 “младенец”	4.4 “родитель”	7.7 “родитель родителя”
2.2 “ребёнок”	5.5 “семьянин”	8.8 “родитель семьянина”
3.3 “подросток”	6.6 “воспитатель”	9.9 “родитель воспитателя”

Рисунок 7. Взаимодействие фрагментов картины мира.

Или в более полном виде:

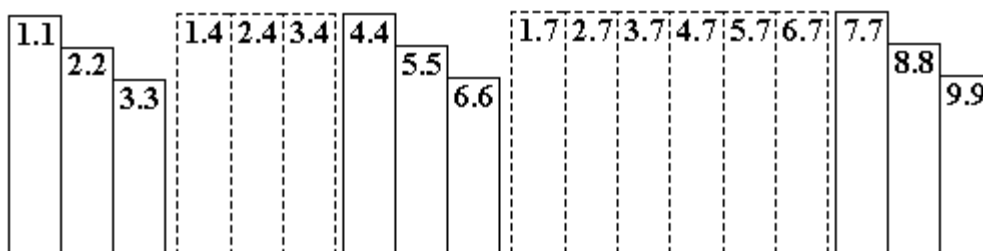


Рисунок 8. Взаимодействие фрагментов картины мира.

Такое представление позволяет восстанавливать информацию, которая формально в картине мира не содержится. Так, например, картина мира подростка (этап 3.3) будет присутствовать в картине мира воспитателя (этап 6.6) и так далее.

Заключение

Субъектно-ориентированный принцип построения картины мира позволяет оперировать структурным контекстом знаний. Однако при этом не учитывается, что явления, понятия и контексты изменяются во времени. В общем случае, каждый конкретный (актуальный) момент времени задает конкретные (актуальные) контексты. Другими словами, структурные контексты сами структурируются во времени.

Задачу оперирования временными контекстами можно решить, если принять субъектно-ориентированную картину мира за некий начальный, нулевой момент (такт) времени. Тогда следующий момент (такт) времени (t+) будет соответствовать другой картине мира, которая связана с предыдущей определёнными процедурами. И так далее.

В настоящий момент мы заняты решением задач реализации принципов и механизмов, позволяющих оперировать временным взаимодействием картин мира, подобно тому, как можно оперировать взаимодействием различных моделей предметных областей.

Литература

- 1) Александров С.Э., Фаддев П.Э. Осмысленная обработка информации. Семантические и категориальные аспекты // Компьютерная лингвистика и интеллектуальные технологии. Труды международной конференции “Диалог-2003”. М: Наука, 2003. С. 141-144.
- 2) Аляксандраў С., Фадзееў П. Шматкантэкставая аўтаматычная апрацоўка вялікіх аб’ёмаў інфармацыі. Прынцыпы, мэтады, зыходныя альгарытмы. Прага: МК Tisk, 2000.
- 3) Карпов В.А. Язык как система. Мн: Вышэйшая школа, 1992; 2-е изд. М: УРСС, 2003.
- 4) Мартынов В.В. Основы семантического кодирования. Опыт представления и преобразования знаний. Мн: ЕГУ, 2001.
- 5) Нариньяни А.С. ТЕОН-2: От тезауруса к онтологии и обратно // Компьютерная лингвистика и интеллектуальные технологии. Труды международного семинара “Диалог-2002”. Т. 1. Теоретические проблемы. М: Наука, 2002. С. 307-313.
- 6) Новое в русской лексике. Словарные материалы-81. М: Русский язык, 1986.