



В.П. Дерябкин, А.И. Белоусов

АССОЦИАТИВНЫЙ ПОИСК ПО ОБРАЗЦУ В ФРЕЙМОВОЙ ИНФОРМАЦИОННОЙ КОМПЬЮТЕРНОЙ СРЕДЕ

(Самарский государственный аэрокосмический университет)

В [1-4] были изложены концепция, формальные модели и результаты построения интеллектуальной информационной компьютерной среды (ИКС) с фреймовым способом представления знаний, дополняемой при необходимости правилами-продукциями, которые также могут быть заданы фреймами с унифицированной структурой. Реализация среды выполнена в виде интеллектуальной информационной системы фреймового типа (ИИСФТ).

Фреймы базы знаний могут быть подразделены на группы:

- *активные* в логическом выводе результата, имеющие слоты с исполняемыми присоединенными процедурами или правилами (демонами и/или значениями слотов) F_a – поведение этих фреймов определяет логический вывод;
- *пассивные*, не имеющие слотов с присоединенными процедурами логического вывода F_p (они представляют фреймы-структуры, аналогичные таблицам, изменение значений слотов на время вывода запрещено).

К активным фреймам относятся системные фреймы, например, фрейм приложения и др., имеющие хотя бы одну присоединенную исполняемую процедуру или правило.

Из многочисленных методов решения задач в интеллектуальных системах как с чёткой, так и с нечёткой логикой [5] выделим, на наш взгляд, наиболее перспективные для фреймовых систем:

- a) традиционные, в форме алгоритма инициализации присоединённых процедур или правил по данным пассивных фреймов;
- b) поиск по образцу решения, наиболее близкого к образцу среди пассивных фреймов с использованием той или иной меры близости;
- c) метод распространения возбуждения по сети фреймов.

В системе должен иметься исходный активный фрейм-ситуация «запрос» стандартной структуры, содержащий слоты параметров запроса, а также слот результата. Именно обращение к слоту результата за значением инициализирует процесс решения задачи (логического вывода).

Следует различать базу знаний среды разработки – общую базу знаний, применимую для всех возможных приложений (справочная библиотека) и частную базу знаний (приложение).

База знаний ИКС в случае одного приложения представляет собой множество фреймов $F = F_{sys} + F_{app}$ (объединение фреймов инструментальной системы и фреймов приложения соответственно) [2].

Каждый фрейм $f \in F$ рассматривается как некоторое подмножество слотов из множества S :



$$S = \{ \langle v, u, \{Q_i\}, \{D_j\}, \{C_k\}, \alpha \rangle \}, \quad (1)$$

$v \in \mathbb{T}$ – значение слота; $u \in \mathbb{T}$ – значение слота по умолчанию (значения определены на системе типов \mathbb{T}); $\{Q_i\}$ – упорядоченное множество присоединенных к слотам правил или процедур-демонов поиска значений слота типа IF_NEEDED; $\{D_j\}$ – упорядоченное множество присоединенных к слоту правил или процедур-демонов типа IF_CHANGED, обрабатывающих событие изменения значения слота; $\{C_k\}$ – упорядоченное множество ограничений на значения слота (набор правил или предикатов $C_k \in \mathbb{E}$, \mathbb{E} – множество выражений); α – флаг, используется в контексте вывода для управления выводом, булевского типа. Для доступа к значениям слота используем операцию разыменования: $s.v$, $s.u$; $s.IF_NEEDED_i$; $s.\alpha$ и т.п.

Традиционный метод решения а) использует хорошо известную семантику работы с присоединёнными правилами и процедурами-демонами. Данная семантика подробно рассмотрена в [3,6]. Далее основное внимание будет уделено случаю б). Метод с) на сетях фреймов изложен в [4].

Задача поиска множества фреймов, релевантных фрейму-образцу (эталону), является частным случаем общей задачи ассоциативного поиска релевантных знаний в интеллектуальной системе [7] при условии выполнения следующих ограничений:

- 1) начальные и конечные образы ассоциаций задаются фреймами описанной структуры;
- 2) каждый фрейм f имеет уникальное имя (идентификатор) $id \in I$, $I \subset \mathbb{I}$, (\mathbb{I} – полное множество идентификаторов фреймов и слотов) и содержащий как агрегацию набор слотов из множества S : $S_f \subset S$;
- 3) каждый слот s имеет уникальный в пределах фрейма идентификатор $id_s \in I_s$, $I_s \subset \mathbb{I}$;
- 4) каждый слот s может иметь значения $v \in \mathbb{T}$, $u \in \mathbb{T}$ с учётом наложенных ограничений $\{C_k\}$, включая неизвестное значение NIL;
- 5) пространство признаков фреймов, учитываемых в ассоциациях, может включать все возможные элементы кортежей множества S , как системных (предопределённых), так и пользовательских;
- 6) поиск ассоциативных фреймов ведётся на множестве фреймов, не изменяемых в процессе поиска и не включающем активный начальный фрейм-ситуацию «запрос»;
- 7) фреймы образуют иерархические структуры в виде обобщений с наследованием и агрегаций; для их образования в определение фрейма вводятся специальные системные слоты со значениями ссылочного типа;
- 8) обучение среды в процессе поиска не производится – это тема отдельных исследований.



Будем рассматривать элементарную ассоциацию между двумя фреймами $\alpha(f_0, f_1)$ как некоторое отношение подобия (сходства), определённое с помощью той или иной меры близости. Предполагается, что это отношение симметрично, рефлексивно и транзитивно [7]. Фрейм $f_0 \in F_a$ является начальным образом ассоциации (эталонном), он активен во время поиска ассоциации и является фрейм-ситуацией «запрос», содержащей все необходимые условия поиска. Фрейм $f_1 \in F_p$ является конечным образом ассоциации, он пассивен и не изменяется во время поиска. Успешность поиска фиксируется предикатом $\psi(f_0, f_1, m^a, t^a)$, где m^a – заданная мера подобия, t^a – заданное время поиска (или количество тактов при поиске в дискретном времени). Результатом поиска может быть множество элементарных ассоциаций $A = \{\alpha_i(f_0, f_i)\}$, $i = (1, n)$, для которых предикат истинен.

Стратегия поиска ассоциаций зависит от вида графа, отображающего структуру фреймовой базы знаний. В данном случае все фреймы являются наследниками единственного корневого системного фрейма [2] и образуют иерархию обобщений $f \leftarrow g \Rightarrow \mathbf{f.isa} \mathbf{=} g$, где $\mathbf{f.isa} \mathbf{=} g$ – операция вычисления значения наследуемого слота потомка $f.isa$, g – родитель). Потомок f , как частный пример родителя, может содержать и свои, специфичные для предметной области слоты. Иерархия обобщений отображается корневым ориентированным деревом. Иерархии агрегаций могут не учитываться, так как полный просмотр дерева обобщений гарантирует полный просмотр всех фреймов, составляющих узлы этого дерева.

Известно [5], что основными алгоритмами поиска цели на дереве являются различные варианты поиска с возвратами (backtracking), включающими поиск в глубину, поиск в ширину, поиск от цели, поиск от данных. При разработке алгоритма поиска по образцу во фреймовой среде количество просматриваемых узлов (фреймов) может быть существенно сокращено за счёт следующего:

- 1) все фреймы имеют уникальные идентификаторы (имена), поэтому задание в образце имени фрейма однозначно определяет единственную цель поиска;
- 2) все слоты внутри фрейма имеют уникальные идентификаторы (имена), что должно быть учтено при выборе направления поиска – в первую очередь следует начинать с поиска в ширину от корневого фрейма, просматривая несколько уровней один за другим, а затем при совпадении имён слотов использовать поиск в глубину с возвратами в случае необходимости;
- 3) численная мера подобия фреймов должна быть задана в эталоне вместе с указанием на исполняемую процедуру её определения и порог, ниже которого фреймы-прототипы, частные фреймы (с частично заполненными слотами) и фреймы-экземпляры не могут считаться подобными;
- 4) мера подобия может быть задана лингвистической переменной в эталоне вместе с указанием на исполняемую процедуру определения её значений и условия (предикаты) подобия;
- 5) при задании сложной меры подобия комбинированного типа рекомендуется задавать весовые коэффициенты мер в комбинации с приведением к какой-либо единой метрической шкале.



При сравнении фреймов между собой в поиске ассоциаций целесообразно использовать результаты исследований, приведённые в [10], с учётом формализации фреймовых структур ИИСФТ. При этом *ассоциация по абстракции* соответствует ассоциации между фреймами, определяемой только по именам (идентификаторам) слотов без учёта значений и их типов. *Ассоциация по реализации* определяется с учётом значений слотов и их типов.

Очевидно, все фреймы, входящие в ветви дерева обобщений, находятся между собой в ассоциации по абстракции, так как в силу наследования они имеют один или несколько общих слотов. Наиболее простой мерой подобия, как указано в [10], в этом случае является доля совпадающих идентификаторов слотов от общего количества слотов в ассоциированных фреймах.

Вводя ту или иную меру подобия, можно оценить силу ассоциации по реализации между фреймами как количественно, так и качественно (слабая, сильная и т.п.). Для количественных оценок представляется перспективным использование методов нечёткой логики, что является предметом дальнейших исследований.

Литература

1. Дерябкин, В.П. Фреймовая база знаний информационной компьютерной среды [Текст] / В.П. Дерябкин, А.И. Белоусов // Перспективные информационные технологии для авиации и космоса. Избранные труды международной конференции с элементами научной школы для молодёжи. – Самара, СГАУ, 2010. – С. 61-64.
2. Дерябкин, В.П. Модель базы знаний интеллектуальной фреймовой среды [Текст] / В.П. Дерябкин // Перспективные информационные технологии в научных исследованиях, проектировании и обучении (ПИТ 2012). – Самара: Издательство Самарского научного центра РАН, 2012. – С. 164-168.
3. Белоусов, А.И. Семантика языка представления знаний в инструментальной среде фреймового типа [Текст] / А.И. Белоусов // Избранные труды международной конференции с элементами научной школы для молодёжи. – Самара, СГАУ, 2010. – С. 158-162.
4. Дерябкин, В.П. Поиск решения методом распространения возбуждения в сети фреймов интеллектуальной компьютерной среды [Текст] / В.П. Дерябкин // Перспективные информационные технологии (ПИТ 2013). – Самара: Издательство Самарского научного центра РАН, 2013. – С. 105-108.
5. Джарратано, Д. Экспертные системы: принципы разработки и программирования, 4-е издание [Текст] / Д. Джарратано, Г. Райли. – М.: ООО «И.Д. Вильямс», 2007. -1152 с.
6. Сошников Д.В. Логический вывод на основе удалённого вызова и включения в системах с распределённой фреймовой иерархией [Текст] / Д.В. Сошников // Под ред. В.Е. Зайцева. - М.: «Вузовская книга», 2002. – 48 с.
7. Гаврилов, А.В.. Гибридные интеллектуальные системы [Текст] / А.В.Гаврилов. – Новосибирск: Изд-во НГТУ, 2002. – 142 с.



8. Любченко, В.В. Метод поиска фрейма по шаблону на основе ассоциаций [Текст]/ В.В. Любченко, В.А. Крисиллов // Труды Одесского политехнического университета, спецвыпуск. – Одесса, Издательство Одесского политехнического университета, 2006. – С. 60 – 63.

Д.В. Еленев, А.О. Линник

АВТОМАТИЗИРОВАННОЕ СНИЖЕНИЕ ДУБЛИРОВАНИЯ ИНФОРМАЦИИ В СИСТЕМЕ МОНИТОРИНГА ДЕЯТЕЛЬНОСТИ ПОДРАЗДЕЛЕНИЙ УНИВЕРСИТЕТА

(Самарский государственный аэрокосмический университет имени академика С.П.Королева (национальный исследовательский университет))

Задача сбора и анализа данных о результативности деятельности подразделений является насущной для всех крупных научно-образовательных организаций. Построение внутренних рейтингов подразделений и работников, оценки результатов научной деятельности работников и студентов, подготовка внешних отчетов требует полноту, достоверность и своевременность поступления данных.

В Самарском государственном аэрокосмическом университете имени академика С.П.Королева (национальном исследовательском университете) задача сбора сведений о достижениях работников и подразделений решается при помощи информационно-аналитической системы мониторинга деятельности подразделений и количественной оценки качества работы университета [1]. В состав системы мониторинга входит комплекс автоматизированных рабочих мест (АРМ): «Ректор», «Проректор», «Дирекция Программы развития национального исследовательского университета», «Кафедра», «Управление образовательных программ», «Интеллектуальная собственность», «Научно-исследовательская работа студентов», «Управление обеспечения инновационной деятельности», «Отдел управления качеством» и др. Основная нагрузка по сбору сведений о достижениях работников и подразделений ложится на АРМ «Кафедра».

Для оперативного контроля деятельности подразделений в системе мониторинга используется построение отчетов на основании данных, введенных в течение текущих контрольных периодов. Текущий контрольный период установлен равным декаде, а текущие контрольные точки – 10, 20 и последние числа каждого месяца. Данные, собранные и обобщенные пользователем АРМ за текущий контрольный период, заносятся пользователями АРМ «Кафедра» не позднее даты текущей контрольной точки.

В число предоставляемых пользователями АРМ «Кафедра» данных входят сведения об опубликованных научных работах, участии студентов в конкурсах, участии работников в конкурсах, проведенных научных мероприятиях,