

Текстовая визуализация баз знаний

Кушмар С.Е.

Кафедра интеллектуальных информационных технологий
Белорусский государственный университет информатики и радиоэлектроники
Минск, Республика Беларусь
e-mail: s.kushmar@gmail.com

Аннотация—Рассматриваются способы визуализации баз знаний в виде текстов в контексте семантической технологии компонентного проектирования интеллектуальных систем.

Ключевые слова: интеллектуальная система; база знани; пользовательский интерфейс; визуализация

I. ВВЕДЕНИЕ

Для создания качественного пользовательского интерфейса интеллектуальных систем важно использовать эффективные способы представления знаний. Одним из направлений визуализации баз знаний является текстовая визуализация, под которой понимается представления информации с помощью текстов на различных языках (в том числе формальных).

В данной статье рассматриваются языки текстовой визуализации баз знаний в рамках семантической технологии компонентного проектирования интеллектуальных систем различного назначения.

II. Языки текстовой визуализации баз знаний

Наряду с графическими языками содержимое базы знаний может быть представлено с помощью текстовых языков. К языкам текстовой визуализации относятся SCs (Semantic Code string), SCn (Semantic Code natural), естественный язык.

SCs-код является линейной (символьной) модификацией языка SC (Semantic Code) [1]. SCs обеспечивает представление унифицированных абстрактных семантических сетей в виде, близком к традиционным текстам [2]. В алфавит SCs-кода входят символы для представления:

- инцидентности;
- принадлежности / непринадлежности;
- нечеткой принадлежности;
- атрибута и компонента кортежа;
- ориентированной / неориентированной пары;
- равенства / неравенства значений;
- нечеткого равенства значений;
- бинарных отношений над множествами;
- бинарных отношений над числами;
- унарных функций над числами;
- бинарных функций над множествами;
- бинарных функций над числами;
- комментария;
- содержимого sc-элемента.

Важно, что SCs позволяет задать не только константы, но и переменные, а также метаварьиемые.

Ввиду того, что SCs является формальным языком, он может использоваться для оформления исходных текстов баз знаний. Этот язык, однако, уступает

графическому языку SCg (Semantic Code graphical) в удобстве для человеческого восприятия. Соответственно, SCs-код целесообразнее использовать в качестве "промежуточного" языка представления базы знаний, чем в пользовательском интерфейсе конечного пользователя.

Другим формальным языком, пригодным для визуализации баз знаний, является SCn-код. SCn – это язык гипертекстового представления в псевдоестественном виде. Он предназначен для упрощения записи и чтения человеком исходных текстов баз знаний [3].

Тексты языка SCn представляют собой набор sc.n-статей. Каждая sc.n-статья описывает некоторый sc-элемент и состоит из идентификатора этого элемента и набора sc.n-полей, каждое из которых описывает одну или несколько связей отношения. В SCn выделяют поля общего и специального вида. Sc.n-поле может описывать связи как ориентированного, так и неориентированного отношения. По признаку арности отношения выделяют:

- sc.n-поле, описывающее связи бинарного отношения,
- sc.n-поле, описывающее связи ролевого отношения,
- sc.n-поле, описывающее связи неролевого бинарного отношения,
- sc.n-поле, описывающее связи небинарного отношения.

С помощью полей специального вида представляются отношения: синонимии, пояснения, подмножества, надмножества, принадлежности, семантической эквивалентности и другие.

В семантической технологии проектирования баз знаний sc-систем SCn-код используется в качестве одного из основных способов представления знаний. Такой язык понятен пользователю интеллектуальной системы и его использование существенно повышает качество восприятия текстов баз знаний [4].

Пример sc.n-статьи о конференции OSTIS-2012 представлен на рисунке 1.

Тексты базы знаний также могут быть транслированы в тексты на естественном языке. Для естественно-языкового пользовательского интерфейса возможно использование ограниченного набора лексики и грамматики языка без серьезного ущерба функциональности системы [5]. Применение текстов на естественном языке при общении пользователя с системой является оправданным, но в то же время текст на естественном языке в своем первоначальном виде не может быть использован в качестве исходного текста базы знаний. Для этого необходимо применять процедуру извлечения знаний в отличие от формальных языков SCs и SCn.

При разработке вопросно-ответных интеллектуальных систем важной задачей является определение оптимального объема информации, наиболее релевантной запросу пользователя. Данная проблема четко проявляется при вопросах вида "что такое", когда пользователю должна быть представлена некая семантическая окрестность заданного понятия. Как при текстовой, так и при графической визуализации баз знаний важно выделить границы такой семантической окрестности.

III. ЗАКЛЮЧЕНИЕ

На данном этапе развития семантической технологии компонентного проектирования интеллектуальных систем в качестве языка текстовой визуализации целесообразно использовать язык SCn. С одной стороны, этот язык является формальным, что позволяет использовать его в качестве исходных текстов баз знаний. С другой стороны, приближенность SCn-кода к естественному языку и структурированность текстов на данном языке делает его удобным для восприятия.

- [1] Представление и обработка знаний в графодинамических ассоциативных машинах / В.В. Голенков [и др.]; под ред. В.В. Голенкова. – Минск : БГУИР, 2001.
- [2] Голенков В.В. Графодинамические модели параллельной обработки знаний: принципы построения, реализации и проектирования / В.В. Голенков, Н.А. Гулякина // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2012): материалы Междунар. научн.-техн. конф. Минск, 16-18 февраля 2012 г.) – Минск: БГУИР, 2012.
- [3] Ивашенко В.П. Семантические модели и средства интеграции и отладки баз знаний / Ивашенко В.П. // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2012): материалы Междунар. научн.-техн. конф. Минск, 16-18 февраля 2012 г.) – Минск: БГУИР, 2012.
- [4] Колб Д.Г. Семантическая модель сайта конференции / Колб Д.Г. // Информационные технологии и системы 2011 (ИТС 2011): материалы Междунар. научн. конф. Минск, 26 октября 2011 г.) – Минск: БГУИР, 2011.
- [5] Гецевич Ю.С. Семантическая технология проектирования белорусско- и русскоязычных ея-интерфейсов вопросно-ответных систем / С.А. Гецевич, Ю.С. Гецевич, О.Е. Елисеева, В.А. Житко, А.А. Кузьмин // Открытые семантические технологии проектирования интеллектуальных систем (OSTIS-2012): материалы Междунар. научн.-техн. конф. Минск, 16-18 февраля 2012 г.) – Минск: БГУИР, 2012.

OSTIS-2012

⇒ Конференция OSTIS-2012

- = международная конференция
- = научно-техническая конференция
- = Конференции OSTIS
- = OSTIS-2011

§ Время проведения: с_ :2012.02.16; по_ :2012.02.18.

§ Информационное письмо:

- **русскоязычный вариант_:**

- **англоязычный вариант_:**


§ Программный комитет: Гаврилова Т. А.; Голенков В. В.; Гулякина Н. А.; Головки В. А.; Грибова В. В.; Ефименко И. В.; Заболева-Зотова А. В.; Загоруйко Ю. А.

Рис. 1. Пример sc.n-статьи