



OSTIS-2012

(Open Semantic Technologies for Intelligent Systems)

УДК 004.822:514

ЯЗЫКИ И ТЕХНОЛОГИИ ПРОГРАММИРОВАНИЯ, ОРИЕНТИРОВАННЫЕ НА ОБРАБОТКУ СЕМАНТИЧЕСКИХ СЕТЕЙ

Гулякина Н.А., Пивоварчик О.В., Лазуркин Д.А.

*Белорусский государственный университет информатики и радиоэлектроники,
г. Минск, Республика Беларусь*

pivovarchyk@tut.by

В работе описана интеллектуальная help-система для разработчиков программ, ориентированных на обработку семантических сетей. Описаны семантические языки программирования, технологии разработки программ на которых включены в help-систему. Описаны компоненты help-системы. Описан компонент – интеллектуальная help-система по семантической теории программ.

Ключевые слова: интеллектуальная help-система, семантическая теория программ, язык программирования, база знаний.

ВВЕДЕНИЕ

Интеллектуализация является одним из основных направлений развития современных информационных систем. Использование интеллектуальных технологий для разработки таких систем позволяет сократить сроки их проектирования, повысить качество, увеличить количество, повысить эффективность эксплуатации. Для эффективного использования интеллектуальных технологий разработчики должны владеть методами представления знаний, методами моделирования баз знаний, технологиями проектирования программ, обрабатывающих знания. Высокая сложность и уникальность разработки интеллектуальных информационных систем или интеллектуальных компонентов информационных систем ограничивает количество разработчиков. Поэтому важной задачей является повышение качества и сокращение сроков освоения этих технологий. Одним из способов решения задачи является создание программных средств поддержки разработки интеллектуальных систем, которые решают задачи консультантов по проектированию программного обеспечения (преподавателей, опытных программистов, разработчиков моделей предметных областей и др.). Такие программные средства будем называть интеллектуальными help-системами. Интеллектуальные help-системы позволяют повысить уровень интеллектуальной информационной поддержки разработчиков программ. Наличие help-системы в составе инструментального средства проектирования

программ позволит снизить начальные квалификационные требования программистов и повысить количество проектируемых интеллектуальных информационных систем.

В данной статье описана справочная подсистема интеллектуальной help-системы по семантической теории программ. В качестве языка программирования, ориентированного на обработку семантических сетей, в данной статье рассматривается язык SCP (Semantic Code Programming).

1. Языки программирования, ориентированные на обработку семантических сетей

Языки программирования, ориентированные на обработку знаний, существенно отличаются от языков, ориентированных на обработку традиционных структур данных. Особенностью таких языков является:

- наличие мощных языков или систем представления знаний;
- наличие гибких средств управления знаниями или базами знаний;
- возможность программной и (или) аппаратной реализации.

Среди языков, ориентированных на обработку знаний, выделяются языки программирования, ориентированные на обработку семантических сетей. Под семантической сетью понимается графовая структура, вершины и дуги которой обладают определенной семантической

интерпретацией. [Голенков, В.В. и др., 2001a] Переработка знаний, представленных в виде семантических сетей, сводится к изменению конфигурации сети. Языки, которым принадлежат только семантические сети, называются графовыми языками или семантическими графовыми языками. Графовые языки делятся на языки представления знаний и языки обработки знаний.

Существуют множество языков представления знаний для различных предметных областей, например:

- для задач планирования: STRIPS, ADL, PDDL;
- для онтологических систем: OWL, CycL, KL-ONE, KIF, OCML;
- для задач распознавания и генерации текстов: KRL, FRL.

Примерами графовых языков представления знаний являются:

- сетевой язык представления фреймов;
- язык растущих пирамидальных сетей;
- графовый язык представления информации в памяти вегетативной машины;
- графовый язык представления информации в памяти абстрактной сетевой машины;
- графовые логические языки и др.

Примерами графовых языков обработки знаний являются:

- волновые языки, обеспечивающие решение задач навигации на семантических сетях;
- язык СЕТЛ; [Левин Д.Я., 1983]
- графовые языки представления различным образом устроенных продукционных программ, ориентированных на переработку семантических сетей;
- язык передачи маркеров и др.

Для представления и обработки семантических сетей в проекте OSTIS (открытая семантическая технология компонентного проектирования интеллектуальных систем, Open Semantic Technology for Intelligent Systems) [Электронный ресурс] используется семейство графовых языков, текстами которых являются sc-конструкции. Такие языки называются sc-языками. Базовым языком для построения sc-языков является sc-код. [Голенков В.В. и др., 2001a] С помощью sc-кода можно представлять любую информацию, в том числе и нечисловую, описывать сложноструктурируемые предметные области. На базе sc-кода разработано множество языков представления знаний, ориентированных на описание определенных предметных областей.

Для обработки знаний, представленных с помощью sc-языков, существуют языки программирования:

- SCP (Semantic Code Programming) – базовый процедурный язык программирования;

- SCPH (Semantic Code Programming High-level) – процедурный язык программирования высокого уровня;
- функциональный язык программирования;
- логические языки программирования на базе языка SCL.

Данные языки программирования разработаны на базе sc-кода, таким образом, они также являются sc-языками. Следовательно, программы и перерабатываемые ими данные, являются sc-конструкциями. Построение языков представления знаний и языков обработки знаний по одному принципу обеспечивает представление данных и программ в виде sc-конструкций, что позволяет легко интегрировать различные виды информации и хранить в одной базе знаний декларативные и процедурные знания. Процедурные знания называются программами и представляют собой знания о методах решения различных классов задач. Каждой программе соответствует свой класс задач, решаемых с ее помощью. Решение конкретной задачи требует реализации программы над конкретными исходными данными с помощью информационной машины. Каждый конкретный язык программирования представляет собой один из способов записи программ. При этом каждый конкретный язык программирования однозначно задает систему операций абстрактной машины, которая обеспечивает выполнение всех программ заданного языка. Информационная конструкция, хранящаяся в памяти абстрактной машины и преобразуемая в ходе выполнения программы, называется данными программы или декларативными знаниями.

2. Язык программирования SCP

SCP (Semantic Code Programming) [Голенков, В.В. и др., 2001b] – это базовый язык программирования, ориентированный на обработку однородных семантических сетей, имеющих базовую теоретико-множественную интерпретацию. Для представления знаний в виде семантических сетей используется sc-код. Язык SCP является sc-языком, т.е. тексты языка (scp-программы) представляются в виде sc-конструкций. Scp-программы и перерабатываемые ими информационные структуры (sc-конструкции) могут храниться в одной базе знаний. Особенности языка SCP являются:

- ориентация на использование структурно перестраиваемой (графодинамической) ассоциативной памяти;
- хорошая приспособленность к переработке нечисловых структур, что обеспечивает описание сложно структурированных предметных областей;
- ориентация на переработку непосредственно семантических сетей;
- возможность описания механизмов решения задач различного уровня сложности;

- возможность через понятие sc-узла интеграции scr-программ с программами, написанными на других языках программирования и описывающими переработку информационных структур;

- высокий потенциал распараллеливания процессов переработки информации в графовой структурно перестраиваемой ассоциативной памяти. [Голенков, В.В. и др., 2001b]

Язык SCP обладает мощной типологией операторов, позволяющей производить различные виды операций с информационными структурами, представленными в виде семантических сетей. Типология операторов языка SCP:

- операторы генерации sc-конструкций;
- операторы удаления sc-конструкций;
- операторы ассоциативного поиска sc-конструкций;
- операторы выборки;
- операторы проверки условий;
- операторы изменения свойств sc-элемента;
- операторы управления scr-процессом;
- операторы вывода.

Операционная семантика языка программирования SCP задается соответствующей абстрактной машиной. Абстрактная машина задается абстрактной памятью, в которой хранятся перерабатываемые конструкции, и множеством операций. Все вспомогательные данные, необходимые абстрактной scr-машине для реализации хранимых в памяти scr-программ, так же представляются в виде sc-конструкций. Абстрактная память является динамической информационной конструкцией. Структура памяти определяется синтаксисом языка, принципы изменения памяти определяются операциями.

Инструментальное средство проектирования программ базового языка программирования, ориентированного на обработку унифицированных семантических сетей, поддерживают представление scr-программ в графической форме (SCg) и в линейной форме (SCs, M4SCP). Инструментальное средство, используемое в настоящее время, включает: редактор scr-программ, средство сборки репозитория исходных текстов, scr-интерпретатор, scr-отладчик с возможностью визуального отображения процесса отладки scr-программ. Инструментальное средство является интеллектуальной системой и построено с использованием открытой семантической технологии.

3. Интеллектуальная help-система по проектированию программ, ориентированных на обработку семантических сетей

Интеллектуальная help-система является консультантом в области технологии проектирования программ, ориентированных на

обработку семантических сетей. Для проектирования интеллектуальной help-системы используется открытая семантическая технология компонентного проектирования интеллектуальных систем. [Электронный ресурс] В основе семантической технологии проектирования интеллектуальных систем лежит семантическая модель представления знаний, которая использует семантическую сеть с базовой теоретико-множественной интерпретацией и представление знаний в sc-коде. Для представления знаний в базе знаний help-системы используется sc-код. [Голенков В.В. и др., 2001a] Таким образом, help-система является sc-системой.

Help-система состоит из следующих компонентов:

- интеллектуальная help-система по семантической теории программ;
- интеллектуальная help-система по библиотеке программ, ориентированных на обработку семантических сетей;
- интеллектуальная help-система по комплексу инструментальных средств проектирования программ, ориентированных на обработку семантических сетей;
- интеллектуальная help-система по методике проектирования программ, ориентированных на обработку семантических сетей;
- интеллектуальная help-система по методике обучения проектированию программ, ориентированных на обработку семантических сетей.

Каждый компонент содержит знания из соответствующей области технологии проектирования программ. В соответствии с открытой семантической технологией каждый компонент включает:

- справочную подсистему;
- подсистему мониторинга и анализа деятельности разработчика программ;
- подсистему управления обучением.

Каждая из подсистем взаимодействует с другими подсистемами, а также может функционировать автономно. Разработка каждой подсистемы состоит из проектирования:

- базы знаний;
- машины обработки знаний;
- пользовательского интерфейса.

Одним из основных принципов открытой семантической технологии является унификация интеллектуальных систем. Унификация позволяет не только упростить разработку интеллектуальных систем, но и интегрировать как отдельные компоненты, так и целые системы. Это позволяет реализовывать каждый компонент help-системы и каждую подсистему отдельно и легко интегрировать их. В настоящее время разрабатывается справочная подсистема help-системы по семантической теории

программ.

Справочная подсистема является консультантом-экспертом в области семантической теории программ, который может ответить на любой вопрос новичка или опытного пользователя. К функциям справочной подсистемы относятся:

- поиск информации по запросу пользователя, в том числе свободно-конструируемому;
- отображение найденной информации с учетом уровня квалификации пользователя;
- анализ программных текстов и внесение предложений по улучшению их эффективности;
- генерация программных текстов по запросу пользователя.

При проектировании интеллектуальной системы семантическая технология вначале предполагает разработку семантического языка представления вопросов. Разработанный язык определяет способы формулирования запросов к хранящейся в базе знаний информации и формирования ответов по каждому выделенному классу вопросов. Классы вопросов в разрабатываемой справочной системе определяются в соответствии с процедурами (операциями), которые требуются при ответе на них. В справочной подсистеме используются базовые классы семантического языка вопросов, а также выделены специальные классы. Основными классами вопросов являются:

- стандартные вопросы (например, запросы всех известных элементов заданного множества; запросы всех известных множеств, элементом которых является указываемый sc-элемент);
- запросы высказываний, связанных с заданными объектами, обобщенными структурами или высказываниями (например, относится ли scr-оператор genEl к классу scr-операторов поиска);
- запросы, связанные с классификацией множеств (например, какие классы scr-операторов являются подклассами операторов поиска);
- запросы, связанные с отношениями (например, привести отношения, заданные на множестве scr-программ);
- запросы, связанные с внешней идентификацией элементов базы знаний и с внешним представлением различных ее фрагментов (например, какой идентификатор используется для внешнего обозначения объекта, обозначаемого sc-элементом язык программирования C++);
- запросы, связанные с определениями и пояснениями понятий (например, привести определение понятия scr-переменная);
- запросы фрагментов семантических окрестностей элементов баз знаний (например, вывести полную семантическую окрестность понятия scr-оператор);
- запросы, связанные с доказательствами (например, истинным ли является данное высказывание «scr-оператор genEl генерирует одноэлементную sc-конструкции в базе знаний»);

- запросы, связанные с программами (например, каким способом можно посчитать мощность множества);

- запросы фрагментов логических спецификаций формальных теорий (например, как выглядит теоретико-множественная система основных классов объектов семантической теории программ).

При анализе ответов на вопросы выделены ключевые понятия, которые легли в основу базы знаний, и определены отношения, для установления связей между ними. Также, для каждого понятия в базе знаний заданы его теоретико-множественные свойства. К теоретико-множественным свойствам можно отнести: пояснение и определение понятия, разбиение на более частные понятия, включение в более общие понятия, синонимия понятия, пересечение и объединение с другими понятиями, основные утверждения о понятии, примеры понятия и др.

В базе знаний справочной подсистемы интеллектуальной help-системы по семантической теории программ представлена информация о семантической теории программ, описаны синтаксис и денотационная семантика базового языка программирования, ориентированного на обработку унифицированных семантических сетей (языка программирования SCP), абстрактная машина, интерпретирующая базовый язык программирования.

Для описания семантической теории программ в виде семантической сети представлены:

- общие понятия теории программ;
- языки представления знаний;
- языки обработки знаний.

Выделены следующие ключевые понятия и отношения: язык программирования, технология программирования, программа, алгоритм, оператор, операнд, операция, переменная, константа, множество, кортеж, решаемый класс задач*, описание операции* и др.

Для представления синтаксиса и денотационной семантики базового языка программирования в базе знаний используется семантическая сеть. В виде семантической сети представлены:

- структуры программы базового языка программирования, ориентированного на обработку унифицированных семантических сетей
- структуры обрабатываемых информационных конструкций;
- структуры оператора базового языка программирования, ориентированного на обработку унифицированных семантических сетей;
- типов операторов генерации унифицированных семантических сетей;
- типов операторов удаления унифицированных семантических сетей;

- типов операторов ассоциативного поиска унифицированных семантических сетей;
- типов операторов проверки условий;
- типов операторов изменения свойств элементов унифицированных семантических сетей;
- типов операторов управления процессами обработки унифицированных семантических сетей;
- типов операторов вывода на консоль;
- типов операторов ассоциативной выборки;
- примеров программ базового языка программирования, ориентированного на обработку унифицированных семантических сетей.

Выделены следующие ключевые понятия и отношения: scr-программа, scr-константа, scr-переменная, scr-параметр, трехэлементная sc-конструкция, пятиэлементная sc-конструкция, одноэлементная sc-конструкция, scr-оператор, тип scr-оператора, scr-оператор генерации sc-конструкций, scr-оператор ассоциативного поиска sc-конструкций, scr-оператор удаления sc-конструкций, scr-оператор проверки условия, scr-оператор изменения свойств, scr-оператор управления scr-процессом, scr-оператор ассоциативной выборки sc-конструкций, значение*, атрибутивное отношение assign_, атрибутивное отношение fixed_ и др. Также, в базу знаний включены утверждения и правила, описывающие способы проектирования программных текстов и задающие семантику языка. Для каждого понятия выделены его теоретико-множественные свойства. Например, для понятия scr-программа выделены:

- синонимичные идентификаторы;
- пояснение;
- примеры (рисунок 1);
- структура программы.

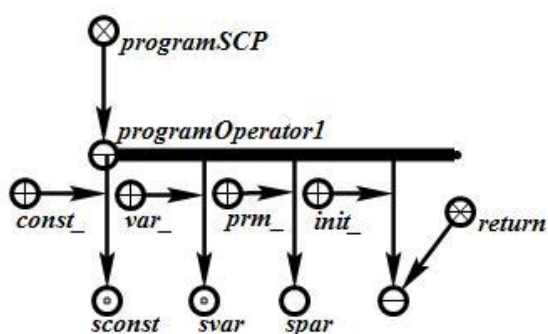


Рисунок 1 – Scr-программа с одним scr-оператором

Для понятия тип scr-оператора выделены следующие свойства:

- синонимичные идентификаторы;
- определение;
- пояснение;
- разбиения по различным признакам (например, разбиение соответствующее типу выполняемого действия);
- структура (обобщенная для всех операторов);
- примеры.

Для понятия genElStr3-оператор выделены следующие свойства:

- синонимичные идентификаторы;
- пояснение;
- принадлежность;
- включение;
- компоненты;
- семантика;
- примеры с семантическими комментариями (рисунок 2).

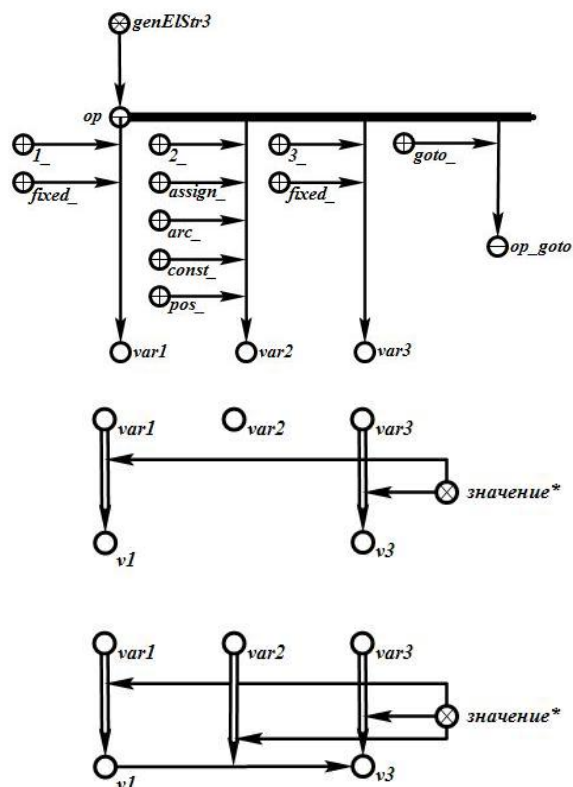


Рисунок 2 – Пример записи scr-оператора genElStr3, состояние памяти до и после выполнения оператора

При работе с системой пользователь может посмотреть синтаксис в виде семантической сети, в виде, представленном традиционно применяемыми средствами (форма Бэкуса-Наура и др.), в виде естественно-языковых текстов.

Для описания абстрактной машины, интерпретирующей базовый язык программирования, в базу знаний были включены понятия необходимые для представления в виде семантической сети:

- языковых средств, описывающих текущее состояние процесса интерпретации scr-программ;
- операций машины, отражающих специфику языка SCP (операционных семантик scr-операторов разных типов);
- используемых общих подпрограмм;
- модели sc-памяти;
- способов реализаций машины.

Выделены ключевые понятия: scr-процесс, тип состояния scr-процесса, then-состояние scr-

процесса, else-состояние scp-процесса, repeat-состояние scp-процесса, run-состояние scp-процесса, error-состояние scp-процесса, dead-состояние scp-процесса, sc-память, сегмент, операционная семантика genEl-оператора, операционная семантика genElStr3-оператора и др. Для понятий выделены теоретико-множественные свойства. Например, для понятия scp-процесс определены синонимичные идентификаторы, пояснение, компоненты, примеры. Понятия для описания общих подпрограмм представляют собой имена программ, использующихся для реализации операционных семантик scp-операторов. Для каждой программы задана спецификация: пояснение, условие применения, класс задач, решаемых с помощью этой программы, алгоритм, примеры выполнения программы. Для описания понятий операционная семантика определенного scp-оператора выделены свойства: пояснение, алгоритм, примеры выполнения заданного оператора с семантическими комментариями.

Семантическая технология содержит множество типовых операций, которые можно использовать для реализации машины обработки знаний интеллектуальной help-системы по семантической теории программ. Множество операций машины обработки знаний help-системы было выделено на основании анализа спецификации вопросов. Каждому классу вопросов соответствует операция (типовая операция или специальная операция).

ЗАКЛЮЧЕНИЕ

При разработке интеллектуальной help-системы для разработчиков программ, ориентированных на обработку семантических систем, используется та же технология, что и использовалась для проектирования инструментальных средств. Это позволяет легко интегрировать систему с инструментальным средством и обеспечить информационное обслуживание разработчика scp-программ как при обучении, так и в процессе проектирования программного обеспечения. Язык программирования и язык представления знаний разработаны на базе той же технологии. Это позволяет унифицировать и интегрировать все необходимые знания о семантической теории программ.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- [Голенков В.В. и др., 2001a] Голенков, В.В. Представление и обработка знаний в графодинамических ассоциативных машинах. Монография / В.В. Голенков, О.Е. Елисеєва, В.П. Ивашенко и др. Под ред. В.В. Голенкова. – Мн.: БГУИР, 2001. – 412 с.
- [Левин Д.Я., 1983] Левин, Д.Я. Язык сверхвысокого уровня СЕТЛ и его реализация (для ЭВМ БЭСМ-6). Монография / Д.Я. Левин – Новосибирск: Наука, 1983. – 158 с.
- [Электронный ресурс] Открытая семантическая технология компонентного проектирования интеллектуальных систем [Электронный ресурс]. – Электронные данные. – Режим доступа: <http://ostis.net/mediawiki/index.php/>.

[Голенков В.В. и др., 2001b] Программирование в ассоциативных машинах / В.В.Голенков, Г.С.Осипов, Н.А.Гулякина и др. – Мн.: БГУИР, 2001. – 276 с.

INTELLECTUAL SYSTEM FOR INFORMATION SERVICE OF SCP-PROGRAMMES' DEVELOPERS

Guliakina N.A., Pivovarchyk O.V.,
Lazurkin D.A.

*Belarusian State University of Informatics and
Radioelectronics, Minsk, Republic of Belarus*

pivovarchyk@tut.by

The article is devoted to the intellectual help-system for programmes developers which are oriented on handling semantic net. Semantic programming languages and programming technologies are described. The components of help-system are presented in the article. The intellectual help-system of semantic programming theory is described in details.

INTRODUCTION

Programmers use programming technologies which are oriented on handling knowledge base for engineering intellectual information systems. The use of this technologies requires special knowledge and a high level qualification. The use of intellectual help-systems will permit to decrease the basic requirements to the programmer's qualification and to increase the quantity of the engineered intellectual information systems.

MAIN PART

The article is devoted to the programming technology which are oriented on handling semantic nets and to the intellectual help-system. Sc-code is a knowledge representation language. SCP is a programming language.

The intellectual help-system is composed of components. Components contain formal models of technology's areas: a semantic programming theory, programmes' libraries, compute-aided software engineering, a methodology of software engineering, a training methodology of software engineering.

Now we engineer intellectual help-system of semantic programming theory. For help-system's development we use the open semantic technology for intelligent systems (OSTIS).

CONCLUSION

The knowledge representation language, the programming language, the compute-aided software engineering, the intellectual help-system develop using one technology. It will permit to integrate the help-system and the compute-aided software engineering, to make sure information service of scp-programmes' developer in the time of learning and in the time of engineering