

УДК 004.94, 519.171, 519.172

Я.М. Мирзабеков**АЛГОРИТМ ХРАНЕНИЯ ИНФОРМАЦИИ В СЕМАНТИЧЕСКОЙ СЕТИ**

Описаны правила определения предмета семантики – знания. Представлены способы хранения знания в памяти компьютера и алгоритмы воспроизводства знания, хранящегося в памяти. Приведенный в данной работе алгоритм имеет практические приложения для экономного хранения информации, когда важен основной смысл текста, а информация, заключенная в стилистических нюансах, лексических оборотах и выразительных средствах языка, не представляет интереса. Определения и примеры, рассмотренные в предложенной статье, придадут ясность возможности практического применения сети, которую мы назвали семантической сетью, для хранения и воспроизводства знания на компьютере. На этой семантической сети можно сформулировать и решать много интересных задач, которые имитируют языковую способность человека и мышление.

Морфология; синтаксис; семантика; язык; информация; знание; понятие; признак.

Ya.M. Mirzabekov**ALGORITHM FOR STORING INFORMATION IN SEMANTIC NETWORKS**

This paper describes the rules for defining the semantics of the subject - knowledge. We present ways of storing knowledge in memory of the computer algorithms and reproduction of knowledge stored in memory. The example in this paper an algorithm has practical applications for economical storage of information, when important basic meaning of the text and information contained in the stylistic nuances of lexical speed and expressive means of language is of no interest. Definitions and examples discussed in the proposed paper will give a clear practical application of the network, which we called a semantic network, storage and reproduction of knowledge on the computer. On this semantic network can be formulated and solved many interesting problems that mimic the ability of human language and thought.

Morphology; syntax; semantics; language; information; knowledge; concept; sign.

За последние годы, несмотря на свою актуальность, задача компьютерной обработки текста не привлекает к себе должного внимания исключительно по причине отсутствия соответствующей научной парадигмы, в частности формальной теории естественного языка. Камнем преткновения при разработке компьютерной модели языка стала семантика. Формализация семантики оказалась задачей более трудной, чем формализация морфологии или синтаксиса. Предмет семантики не имеет четкого определения ни в лингвистике, ни в психологии.

В данной работе описаны простые правила определения предмета семантики – знания, способы хранения знания в памяти компьютера и алгоритмы просмотра (воспроизводства) знания, хранящегося в памяти. Несмотря на свою простоту, описанный прием представления знания охватывает разные формы знания, которыми оперирует человек.

В статье речь пойдет о представимом знании, точнее говоря, об именованном представлении, с которым оперирует мышление. В такой форме хранятся в памяти знания о предметах и явлениях окружающего нас мира.

Представимое знание более структурировано и поддается формализации. Представление хранится в памяти человека и легко активизируется благодаря имени, которое «привязано» к нему. Именованное представление называется понятием. Имя Name(W) понятия W используется далее как формальный представитель понятия W.

Определение понятия, принятое в данной работе, основано на интуитивном восприятии предметов как совокупности признаков, например, цвет, вес и т.д. Каждый признак проявляет себя в своих значениях. Например, признак цвета имеет

значения: синий, красный и т.д. Таким образом, представление о вещи разбито на три вида знания, образующих иерархию: понятие – признак – значение. Например, роза – цвет – красный.

Пусть понятие F задано признаками: $P_1, P_2 \dots P_m$, а признак P_i в понятии F проявляет себя значениями: $P_{i1}, P_{i2} \dots P_{in}$. Если понятие F , все его признаки и значения признаков отметить точками на плоскости и соединить стрелкой точку F с его признаками, каждый признак – со всеми его значениями, то получится фигура, которая называется корневым деревом с корнем в точке F и далее обозначается через $Tree(F)$. Данное дерево есть модель знания, заключенного в понятии F , или, просто, знание F .

В компьютерной модели (в программе) знание F будет представлено в оперативной памяти компьютера деревом $Tree(F)$. Описываемые ниже алгоритмы хранения и воспроизводства знания могут быть реализованы в отдельном техническом исполнении, где каждая вершина дерева $Tree(F)$ будет представлена отдельным элементом электронной схемы знания. Возможно, такую схему имеет нейронная сеть этого же знания.

Понятие F может оказаться значением некоторого признака Q , которое определяет другое понятие G . Или значение P_{ij} может оказаться полноправным понятием, определенным через свои признаки. Таким образом, различные понятия связаны между собой следующим отношением: одно понятие может быть значением признака другого понятия.

На этом можно было остановиться, основная идея представления и хранения знания высказана. Остается только выделить в языке слова, обозначающие признаки, а далее каждое понятие F определить через признаки и их значения, исходя из собственного понимания структуры знания $Tree(F)$.

Приведенные ниже определения и примеры придадут ясность возможности практического применения данной сети для хранения и воспроизводства знания на компьютере.

Сеть, в которой понятия и признаки связаны между собой указанными выше отношениями, называется семантической сетью. Сеть можно хранить в памяти компьютера и программно перемещаться по ее соседним вершинам. На семантической сети можно сформулировать и решать много интересных задач, имитирующих языковую способность человека и мышление.

Пусть имеется путь $X-A-Y$ из трех понятий (квадратов). Пусть они имеют имена: $X = \text{роза}$, $A = \text{цвет}$ и $Y = \text{красный}$. Этот путь представляет собой знание, которое на русском языке выражается различными сочетаниями слов: роза цвета красного, роза красная и т.д.

Как было сказано выше, словом роза обозначено понятие (представление) о розе и имя розы. В словосочетаниях для определенности будут использованы обозначения для имени – $Name(X)$.

Но роза имеет не только цвет, но и аромат, и форму, и т.д. В этом случае из квадрата X будут исходить несколько путей.

Пусть $B = \text{аромат}$, $C = \text{форма}$, $U = \text{приятный}$, $V = \text{круглый}$. Полученная фигура называется корневым деревом с корнем в X . Это дерево состоит из трех путей: $X-A-Y$, $X-B-U$ и $X-C-V$. Каждый из этих путей образует элементарное знание о понятии X . В знании $X-A-Y$ о понятии X элемент A (закругленный прямоугольник) называется признаком понятия X , а Y – значением признака A .

Понятие определяется набором признаков. У каждого признака может быть несколько значений.

Если $Name(X) = \text{роза}$, $Name(A) = \text{цвета}$, $Name(Y) = \text{красного}$, $Name(Y1) = \text{белого}$, $Name(Y2) = \text{желтого}$, то в данной семантике сосредоточены элементарные знания: роза цвета красного, роза цвета белого, роза цвета желтого. Семантика понятия X задается деревом $Tree(X)$. В упрощенном виде это дерево показано на рис. 1.

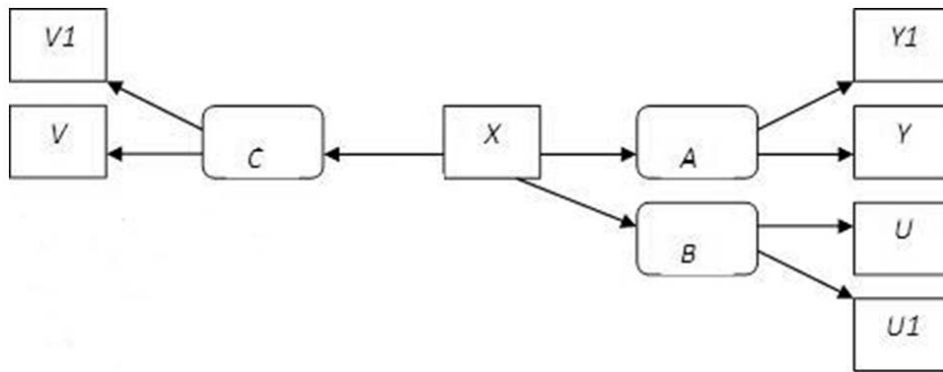


Рис. 1

Семантика понятий намного содержательнее, чем набор изолированных деревьев. Понятие и значение его признака могут меняться местами. Понятие и его значение в сети разделены признаком. Например, семантика сочетания «сок из яблока цвета красного» состоит из двух путей – элементарных знаний (1) и (2):

$$X = \text{сок}, D = \text{из (чего сделан сок)}, Z = \text{яблока}, \quad (1)$$

$$Z = \text{яблоко}, E = \text{цвета}, R = \text{красного}. \quad (2)$$

Упражняясь с простыми примерами семантики понятий и словосочетаниями, мы сталкиваемся с различными проблемами, в частности, вопросами морфологии: какая форма слова будет прикреплена к понятию, т.е. к вершине сети? Можно, конечно, пойти по простому пути, т.е. вершин в сети будет столько, сколько морфологических форм у данного слова. Это означает, что каждая форма слова – самостоятельная лексическая единица. Иначе говоря, у этого языка нет морфологии, каждое слово имеет одну единственную форму.

Но есть и второй путь решения этого вопроса: к вершине будет приписано имя в исходной форме. К каждому признаку привязан определенный морфологический оператор [5], который определяет форму членов в соответствующем элементарном знании. Функции языка (морфологии и синтаксиса) заключаются в генерации фразы «роза аромата приятного» из слов: роза, аромат и приятный.

Описанная здесь модель языка еще раз подчеркивает важность идеи о разграничении полномочий морфологии, синтаксиса и семантики, которая была высказана в [1]. Такой подход к моделированию языка дает широкие возможности для импровизации и реализации на компьютере моделей языка различной сложности морфологии, синтаксиса и семантики.

Используя первый путь, можно быстро написать программу, которая будет «рассказывать» (печатать) сказку или рассказ, каждый раз используя новые слова и предложения.

Одновременно возникает вопрос об обратной задаче: по тексту рассказа построить соответствующее семантическое дерево. Иначе говоря, рассказ – исходная информация, а сетевое представление сюжета строится программно, а далее, при необходимости, этот рассказ будет пересказываться многократно по данному сюжету.

Данная задача может иметь практические приложения для экономного хранения информации, когда важен основной смысл текста, а информация, заключенная в стилистических нюансах, лексических оборотах и выразительных средствах языка, не представляет интереса.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Ельмслев Л. Прологомены к теории языка // Новое в лингвистике. – М., 1960. – Вып. 1.
2. Мирзабеков Я.М., Шихиев Ш.Б. Компьютерная модель самодостаточного синтаксиса // НТВ СПбГПУ. – 2009. – № 2 (76). – С. 172-177.
3. Мельчук И.А. Опыт исследования лингвистических моделей «Смысл-Текст». – М.: Наука, 1974.
4. Слобин Д., Грин Дж. Психолингвистика. – М.: Прогресс, 1976.
5. Тузов В.А. Морфологический анализатор русского языка // Вестник СПбГУ. Сер. 1. – 1996. – Вып. 3, № 15.
6. Харари Ф. Теория графов. – М.: Мир, 1974.
7. Шихиев Ф.Ш. Формализация и сетевая формулировка задачи синтаксического анализа: Дис. ... канд. физ.-мат. наук. – СПб.: СПбГУ, 2006.

Статью рекомендовал к опубликованию д.т.н., профессор Е.А. Башков.

Мирзабеков Яхья Мирзабекович – Дагестанский государственный университет; e-mail: yash831@mail.ru; 367030, респ. Дагестан, г. Махачкала, пр. им. Шамиля, 39, кв. 76; тел.: 88722675817; старший преподаватель.

Mirzabekov Yah'ya Mirzabekovich – Dagestan State University; e-mail: yash831@mail.ru; 39, Imam Shamil avenue, ap. 76, Makhachkala, Dagestan, 367030, Russia; phone: +78722675817; senior lecturer.

УДК 681.325.36

О.В. Тужилкин, Н.С. Ульянин

МЕТОДЫ ОЦЕНКИ ЭФФЕКТИВНОСТИ РАБОТЫ БЕСПРОВОДНОЙ СЕНСОРНОЙ СЕТИ

Рассмотрены методы оценки эффективности работы беспроводных сенсорных сетей. Приведены основные параметры, характеризующие беспроводные системы. Описаны критерии оценки работы сети: критерий суммарной средней задержки, критерий максимальной задержки, критерий средней задержки; описаны типы систем, для которых использование оценки по этим критериям наиболее предпочтительно. Приведены недостатки методик оценки эффективности, использующих единичные характеристики. Обоснована необходимость комплексной оценки работы беспроводной сети. Предложена методика многокритериальной оценки эффективности работы датчиковой сети. Описан способ тестирования системы по отдельным критериям, ориентированный на работу в неблагоприятных условиях. Приведены формула для расчёта эффективности работы сети и методика выбора весовых коэффициентов. Описаны преимущества предложенной методики.

Беспроводные сенсорные сети; методика оценки; критерии оценки; комплексная оценка.

O.V. Tuzhilkin, N.S. Ulianin

METHODS OF EVALUATING THE PERFORMANCE OF WIRELESS SENSOR NETWORK

Methods for evaluating the performance of wireless sensor networks are considered. The main parameters of wireless system are described. We describe the evaluation criteria of sensor networks. The method of integrated assessment of the effectiveness of wireless sensor networks is suggested. The advantages of the proposed method are described. Necessity of a complex assessment of works of a wireless network is proved. The technique a lot of criteria overall performance