

РАСПРЕДЕЛЕННАЯ СИСТЕМА БАЗ ЗНАНИЙ В СРЕДЕ INTERNET

А.В. Барсуков, А.В. Гаврилов
Кафедра вычислительной техники
Новосибирского государственного технического университета
630087 Новосибирск 87, ул. Немировича-Данченко, 136
Россия
E-mail: avg@osilab.cs.nstu.ru
Тел.: (3832)-460-219

Ключевые слова: поиск информации, экспертная система, Internet, сеть фреймов.

Abstract

Inexperienced users typically obtain one of three possible outcomes when they search for online information: they are buried under an information avalanches, they are unable to locate any useful information at all, or they find what they need in roughly the amount they need. Unfortunately, the latter outcome is most rare. Far from being the answer to everyone's information dreams, distributed sources of online information, i.e., WWW, compound the problem and may often turn into an information nightmare. To address this problem, the 'Ephemeride' system is being developed. This paper describes the architecture and algorithms of this system.

1. Введение

В настоящее время человечество переживает очередную техническую революцию, а общество постепенно превращается в информационное. Растет количество информации, хранимой в электронном виде. Компьютерные сети представляют пользователям огромные массивы информации, причем со временем экспоненциально растет как количество этой информации, так и число людей, получивших к ней доступ, благодаря сети Internet. Появилось также огромное количество поисковых систем, облегчающих доступ к ней. Впрочем, как правило, они не отличаются особой интеллектуальностью и используют ту или иную модификацию поиска по ключевым словам. К тому же большое количество информации сейчас хранится в реляционных таблицах различных типов, к которым поисковые системы доступа не имеют. В любом случае пользователь должен проявить определенное упорство и знание предметной области для того, чтобы получить нужную информацию.

Целью данной работы является создание системы, способной предоставлять доступ к информации, доступной с помощью WWW (World Wide Web), и хранящейся в виде текстовых или гипертекстовых документов, баз данных, а также баз знаний о различных предметных областях. Очевидно, что подобная система должна быть распределенной, поскольку один сервер физически не в состоянии хранить всю собранную человеком информацию.

В данной статье описываются принципы построения распределенной системы,

хранящей знания о данных, понимающей вопросы на языке, близком к естественному, и обеспечивающей поиск ответов на вопросы, которые могут не присутствовать в явном виде в текстах или гипертекстах на серверах сети. Такую систему можно назвать распределенной экспертной системой. Данная система получила название “Эфемериды”. Представляемая система является развитием системы “Socrat” [1].

В мире ведется множество разработок в этом направлении, объединяемых обычно парадигмой “Искусственные информационные агенты” [2], но все они пока носят сугубо теоретический характер. Информационные агенты описываются как индивидуальные сущности, совместно работающие над решением общей проблемы с целью разбиения ее на частные подзадачи. Каждый агент способен выделять из задачи свою подзадачу, решать ее и представлять решение как часть решения общей задачи.

2. Система представления знаний.

В качестве системы представления знаний была выбрана сеть фреймов. Такая структура допускает разбиение знаний по предметным областям, что позволяет сделать систему распределенной. Для совместной работы множества узлов распределенной базы знаний необходимы специальные средства, такие как сетевой протокол и язык передачи информации. В системе “Эфемериды” в качестве сетевого протокола используется ТСР/ІР, а данные (фреймы), представляющие собой запросы и ответы, кодируются в двоичном виде. Информация о структуре распределенной системы хранится в базе знаний каждого узла в виде фрейма, описывающего понятие, и инструмента (см. ниже), который отвечает за работу с сетью.

Кроме того, выбор сети фреймов в качестве системы представления знаний обусловлен ее схожестью с реляционной моделью данных. Структура реляционной таблицы однозначно описывается при помощи фрейма. Каждый столбец таблицы описывается с помощью соответствующего слота. Связь двух реляционных таблиц по ключу может быть представлена фреймом, который является значением слота другого фрейма.

В основу любой системы, работающей с данными, заложены так называемые бизнес-правила – информация о структуре хранилища данных и способах их записи и извлечения. В сложившейся ситуации пользователи не могут получить информацию из произвольного хранилища, не зная его специфики (бизнес-правил) и не имея специальных программ. Соответственно, описываемая система должна уметь по некоторым общим признакам породить бизнес-правила для конкретной системы.

3. Инструменты.

Для реализации бизнес-правил необходимо уметь хранить не только информацию о структуре таблиц, но и некоторые специфические знания, имеющие производственный характер (служебные записи в таблицах, специальные значения полей и пр.). Существует несколько способов реализации правил-продукций с помощью сети фреймов. Одним из них является объектно-ориентированный подход, когда объект-фрейм содержит методы-процедуры. В данной статье предлагается иной способ.

Введем понятие инструмента. Это особый тип фрейма, который может заставлять систему выполнять некоторые действия. Другими словами, инструмент – это исполняемый фрейм. Он состоит из двух частей – фреймовой и собственно исполняемой. Когда в процессе логического вывода система встречает инструмент, то исполняемая часть получает из фреймовой некоторые значения и выполняет действия. Каждый исполняемый фрейм должен быть описан в специальной системной таблице инструментов. В ней хранится идентификатор фрейма, имя вызываемой внешней функции, а также информация о входных и выходных параметрах.

Если снабдить систему некоторым набором инструментов, которым соответствуют понятия-действия, то можно будет описывать сложные действия, состоящие из нескольких простых. Таким образом, система способна создавать и модифицировать своеобразные программы. Это открывает возможности для самообучения.

Система для доступа к реляционной базе данных должна содержать как минимум один исполняемый фрейм. Его исполняемая часть содержит вызов системной функции, осуществляющей запрос к базе данных, например, SQL-запрос. Если системе в процессе поиска решения встречается такой фрейм, то SQL-запрос выполняется, и во фреймовую часть прописываются возвращенные значения. Другими словами, система осуществляет некоторые "подсознательные" действия.

В качестве другого примера использования инструментов можно привести работу распределенной системы. В такой системе каждая локальная база знаний должна содержать инструменты, соответствующие понятиям, хранимых в других узлах. Когда такое понятие встречается в запросе, при попытке сопоставить ему значение формируется подзапрос на удаленный сервер, отклик которого и используется в процессе дальнейшего логического вывода.

4. Алгоритм логического вывода

Логический вывод в системе "Эфемериды" основан на порождении и отсеивании гипотез. На этапе ввода/распознавания запроса в некотором специально отведенном месте (будем называть его пространством гипотез) формируются гипотезы о смысле запроса, каждая из которых представляет собой фрейм, которому сопоставлен коэффициент достоверности. В качестве гипотез выбираются фреймы базы знаний, совпадающие с фреймами запроса по имени.

Когда все гипотезы сформированы, начинается второй этап вывода – поиск ответов. При этом для каждой гипотезы делается попытка сопоставить каждому из входящих в ее состав фреймов фрейм базы знаний. Критерием сопоставимости является нечеткое совпадение имен либо связь по наследованию (например, фрейму "человек" сопоставляются как само понятие "человек", так и все известные системе конкретные люди). При этом число гипотез может расти или уменьшаться. Очевидно, что для оптимальной работы системы необходимо первыми анализировать фреймы, которые наиболее сильно ограничивают пространство гипотез. Для этого каждый слот любого фрейма базы знаний имеет специальную характеристику, определяющую, в какой мере значение данного слота является определяющим для описания понятия (например, имя человека является более определяющим, чем возраст или цвет волос). Кроме того, если фрейм базы знаний совпадает с фреймом гипотезы не полностью, достоверность этой гипотезы уменьшается.

По окончании этого этапа в пространстве гипотез остаются только фреймы, являющиеся ответами на заданный вопрос. В зависимости от типа вопроса из них выбирается либо самый достоверный, либо множество фактов с достоверностью не ниже заданной.

5. Состояние и перспективы реализации

В качестве базовой для реализации системы "Эфемериды" была выбрана технология "клиент-сервер". База знаний и вся служебная информация хранятся в таблицах базы данных InterBase 5.0. Такое решение позволяет избавить разработчика от проблем, связанных со скоростью и надежностью работы системы, так как их решение возлагается на SQL-сервер. Поскольку язык SQL является промышленным стандартом, серверная часть системы может быть легко перенесена на любую из существующих аппаратных и программных платформ.

В настоящее время реализована серверная часть системы и ведется работа над ее клиентской частью. Клиентская часть разрабатывается на языке Delphi и включает в себя подсистему общения с пользователем на языке, близком к естественному, и подсистему создания и пополнения баз знаний, которая пока реализуется в виде интеллектуального редактора.

В дальнейшем планируется создание подсистемы автоматизированного извлечения знаний из текстовых и гипертекстовых документов, а также из баз данных с использованием технологии искусственных нейронных сетей [3].

Список литературы

[1] Барсуков А.В., Гаврилов А.В., Олейник Е.В. Представление знаний в системе распределенных баз знаний и данных "Сократ". // Труды Международной научно-технической конференции "Научные основы высоких технологий (НОВТ-97)". - Новосибирск, НГТУ, 1997, т. 2. - С. 212-217.

[2] Donna S. Haverkamp, Susan Gauch. Intelligent Information Agents: - Review and Challenges for Distributed Information Sources. - <http://eecs.ukans.edu/iagents>.

[3] Гаврилов А.В., Канглер В.М., Катомин М.Н., Коротенко А.И. Обнаружение ассоциативных взаимосвязей между полями в базах данных с использованием нейронной сети. // Труды Международной научно-технической конференции "Научные основы высоких технологий (НОВТ-97)". - Новосибирск, НГТУ, 1997, т.2. - С. 210-211.

Biography of Gavrilo A.V.

Andrey V.Gavrilo was born per 1952 in Novosibirsk. Has ended in 1974 Novosibirsk Electrotechnical Institute (NETI) (now Novosibirsk State Technical University - NSTU). In 1977 has finished post-graduate school at faculty of computers of NSTU and in 1979 has protected the candidate dissertation on a theme " Control of process of execution of a flow of tasks in the distributed computing systems ". In 1978-1979 years He worked in Institute of nuclear physics of Siberian Department of Russian Academy of Science, and since 1979 works on faculty of computers of NSTU, assistant, and since 1982 - as the senior lecturer. Since 1991 in combination He is chief of venture " INSYCOM Ltd. " (Intelligent Systems Company) at faculty of Computers of NSTU.

He worked in the following directions: the theory of control of the large distributed computing systems, system and applied software of parallel computation in the distributed systems, system software of local networks, software of the flexible automated manufactures, including, machine tools with NC and robots, information systems based on the data and knowledge, dialogue with the computer in natural language.

Nowadays His areas of scientific interests : expert systems and technology of their development, neural networks and their application, methods of representation and processing of knowledge, hybrid and distributed expert systems, dialogue with the computer in natural language, technology Data Mining and Knowledge Discovery in Databases, intellectual processing of the documents.