

УДК 004.51

ПРОТОТИП СЕМАНТИЧЕСКОЙ ПОИСКОВОЙ СИСТЕМЫ НА ОСНОВЕ ЛОГИКО-СЕМАНТИЧЕСКОЙ СЕТИ «ВОПРОС-ОТВЕТ-РЕАКЦИЯ»

Добрынин Владимир Николаевич¹, Лобачева Марина Владимировна²

¹ *Кандидат технических наук, профессор Института системного анализа и управления;*

*ГОУ ВПО «Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;*

141980, Московская обл., г. Дубна, ул. Университетская, 19;

e-mail: arbatsolo@yandex.ru.

² *Магистр техники и технологии по направлению «Системный анализ и управление»;*

*ГОУ ВПО «Международный Университет природы, общества и человека «Дубна»,
Институт системного анализа и управления;*

141980, Московская обл., г. Дубна, ул. Университетская, 19;

e-mail: Marina@Style.ru.

В работе предлагается описание прототипа семантической поисковой системы на основе логико-семантической сети (ЛСС) «Вопрос-Ответ-Реакция» адекватной смыслу текста. Описан процесс постановки вопроса, поиска ответа и формирования реакций на него. Рассматриваются структурно-функциональные модели создания ЛСС и разработки семантической поисковой системы, которая позволяет пользователю найти ответ, на сформулированный на естественном языке вопрос, а также при желании прочесть дополнительную информацию либо осуществить продвижение вверх (обобщение) или вниз (уточнение) по ЛСС.

Ключевые слова: вопрос, логико-семантическая сеть, ответ, прототип, реакция, семантика, семантическая поисковая система.

THE PROTOTYPE OF THE SEMANTIC SEARCH ENGINE BASED ON THE LOGICAL- SEMANTIC NETWORK «QUESTION-ANSWER-REACTION»

Dobrynin Vladimir Nikolaevich¹, Lobacheva Marina Vladimirovna²

1 PhD, professor of Institute of system analysis and management;

International university of the nature, society and man «Dubna», Institute of system analysis and management;

141980, Dubna, Moscow reg., Universitetskaya str., 19;

e-mail: arbatsolo@yandex.ru.

² *Master of engineering and technology «Systems analysis and management»;*

International university of the nature, society and man «Dubna», Institute of system analysis and management;

141980, Dubna, Moscow reg., Universitetskaya str., 19;

e-mail: Marina@Style.ru.

In this work proposes a description of the prototype of the semantic search engine based on the logical-semantic network (LSN) «Question-Answer-Reaction» adequate sense of the text. It describes the process of raising the question, answering, and the formation of reactions to it. It consider the structural and functional model of the creation of the LSN and the development of semantic search engine, which allows the user to find the answer, formulated in native language question, and if you want to read more information or to carry out promotion up or down of the LSN.

Keywords: question; logical-semantic network; answer; prototype; reaction; semantics; semantic search engine.

Что такое семантическая поисковая система? В общем смысле под семантической поисковой системой понимается система, которая способна в той или иной степени понимать и запрос пользователя, и текст, в котором будет производиться поиск ответа. При этом используются алгоритмы, схожие с человеческими алгоритмами восприятия и понимания. Такая поисковая система выдает точные результаты обработки запроса с первого раза и полностью. Иначе говоря, нет необходимости открывать страницу, так как поисковик сразу выдает ответ на поставленный пользователем вопрос. Но есть и одно серьезное условие: семантическая поисковая система не является семантической, если не знает языка. Следование этому условию влечет за собой исключение тех алгоритмов поиска, которые основываются на алгоритмах сбора статистики по ссылкам, символам, словам, кликам и т.п. Ведь статистика является скорее инструментом, а отнюдь не решением проблемы. Для семантической поисковой системы нужны алгоритмы, которые сопоставляют значение концепций (скорее концепций, нежели слов) и имитируют понимание.

Создание и наполнение семантических поисковых систем - это трудная, затратная, монотонная работа. Но когда все будет сделано, пользователи получат в свое распоряжение совершенно новый механизм поиска ответов на заданный вопрос.

Естественно, первое, на что окажет влияние использование семантической поисковой системы, это низкочастотные запросы, так называемый Длинный Хвост, то есть запросы, состоящие из большого количества связных по смыслу слов. «Длиннохвостные» запросы могут обрабатываться семантическими алгоритмами сразу же, на лету, не используя статистические данные, но выдавая при этом правильные результаты, с адекватным контекстом. С такими возможностями можно найти ответы на непопулярные, длинные и, чаще всего, сложные запросы.

В настоящее время ведутся активные разработки в области создания и усовершенствования семантических поисковых систем. В качестве примеров таких поисковиков можно назвать Nakiа, Powerset, CognitionSearch, Lexxe, которые в настоящее время активно развиваются в заданном направлении. Но, тем не менее, все эти семантические поисковые системы в качестве результата на запрос пользователя приводят список ссылок на сайты, где можно найти ответ. Этот результат, конечно, будет лучше, чем при поиске в обычных поисковых машинах. По крайней мере, тут не будет спама, рекламы, и другой не относящейся к делу информации. Но все-таки это будет не ответ на вопрос пользователя, а список мест, где он может найти ответ. Таким образом, способность к выделению непосредственного ответа на вопрос пользователя можно назвать следующим шагом на пути развития семантических поисковых систем. В то же время, еще одной немаловажной проблемой использования поисковых систем в данный момент является тот факт, что пользователь, задавая вопрос, очень часто не является экспертом в данной области, что влечет за собой возможность некорректной постановки вопроса, и как следствие — получение другого ответа. Зачастую бывает очень трудно сориентироваться и понять, является ли полученный ответ ответом на наш вопрос или в силу неправильной или неточной формулировки вопроса был выдан близкий, но неточный, неполный ответ. Ввод некоторых комментариев (или реакций) может помочь решить эту проблему. В качестве реакций может выступать дополнительная информация по теме вопроса и ответа, ссылки на сайты, словари и т. д. Такими реакциями может сопровождаться как вопрос, так и полученный ответ, что позволит пользователю лучше и быстрее сориентироваться в предметной области.

Прежде чем говорить о семантической поисковой системе, основанной на логико-семантической сети (ЛСС) «Вопрос-Ответ-Реакция», следует разобраться в том, что из себя представляет ЛСС подобного вида.

1. Реакции вопроса — это описание преобладающей области вопроса (для того, чтобы адресат мог осознать, откуда и по какой причине возникает вопрос и далее связать в смысловое соответствие с областью ответа).
2. Реакции ответа — это описание области ответа (для того, чтобы адресат мог осознать смысл вопроса и смысловую связь с ответом).

2. Основные этапы процесса постановки вопроса, поиска ответа и формирования реакций

Процесс постановки вопроса, поиска ответа на него и формирования реакций вопроса и ответа – это сложный многоэтапный процесс. На рис. 2 представлена схема (в методологии IDEF) этого процесса.

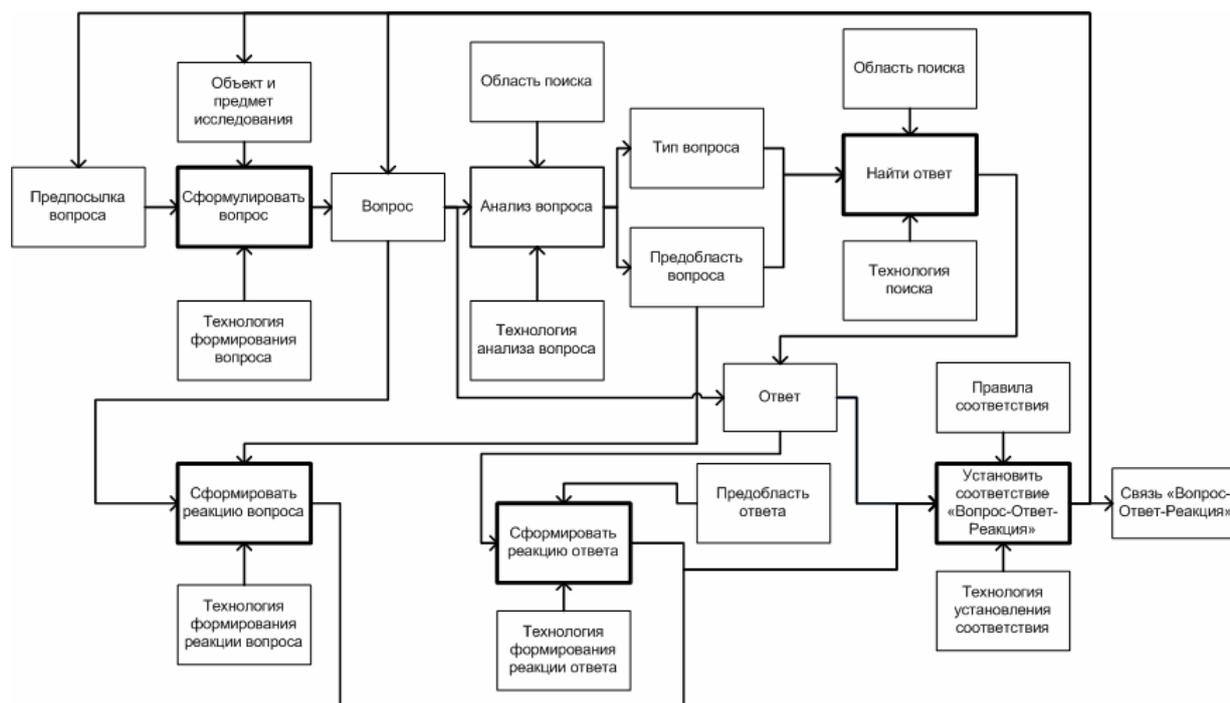


Рис. 2. Процесс постановки вопроса и поиска ответа

3. Прототип семантической поисковой системы на основе логико-семантической сети «Вопрос-Ответ-Реакция»

Под прототипом понимается полное описание информационно-функционального и программно-технологического обеспечения АРМ и его частичная реализация. К основным компонентам реализации прототипа относятся: интерфейс поисковой системы, технология формирования ЛСС, механизм поиска информации на множестве ЛСС.

Для построения множества ЛСС предметной области необходимо выделить общие требования к ЛСС, определиться с теоретическими основами и технологиями создания ЛСС, а также собрать исходные данные для построения ЛСС, в которые входят:

- описание предметной области,
- описание области предпосылок вопросов,
- описание областей поиска ответов,
- описание областей формирования реакций,
- информация, необходимая для решения задачи.

Таким образом, можно построить структурно-функциональную модель создания ЛСС разработчиком (Рис. 3).

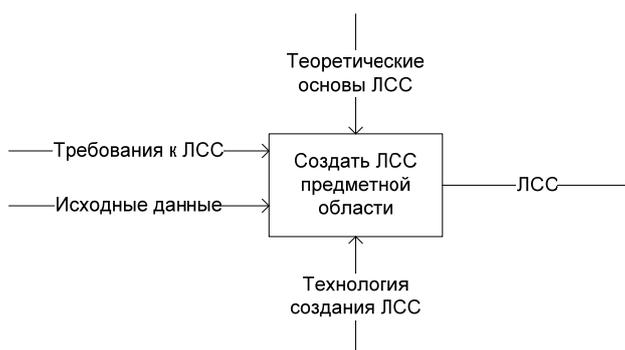


Рис. 3. Структурно-функциональная модель создания ЛСС

Структурно-функциональная модель создания семантической поисковой системы включает в себя следующие компоненты: создание механизма поиска информации как в режиме ответа на запрос пользователя, так и в режиме движения вверх и вниз по ЛСС, для осуществления которого пользователю предоставляется навигатор движения по ЛСС. Таким образом, можно построить структурно-функциональную модель разработки семантической поисковой системы (Рис. 4).

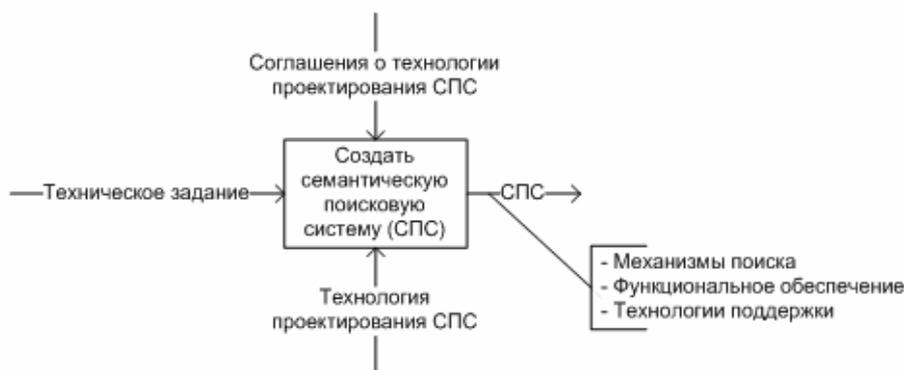


Рис. 4. Структурно-функциональную модель разработки СПС

С точки зрения пользователя данная система позволяет найти ответ на поставленный вопрос. Пользователь задает вопрос и получает на него некоторый ответ с комментариями в виде реакций вопроса и ответа, которые могут подтолкнуть его скорректировать свой вопрос, либо прочитать уточняющие или обобщающие вопросы (т. е. продвинуться вверх или вниз по ЛСС) (Рис. 5).

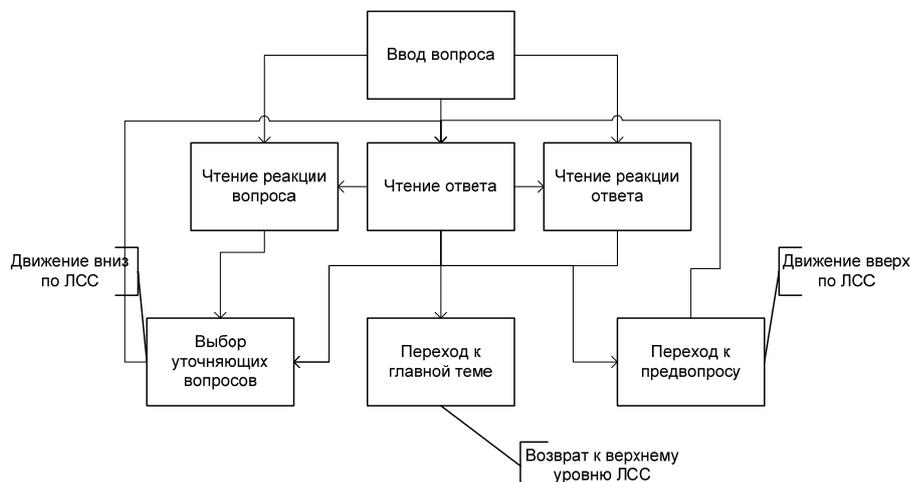


Рис. 5. Варианты действий пользователя

Ниже представлены формы интерфейса, показывающие пример реализации поиска ответа на поставленный вопрос.

Пользователь задает вопрос, после чего происходит поиск по логико-семантической сети и выдается результат (Рис. 6).

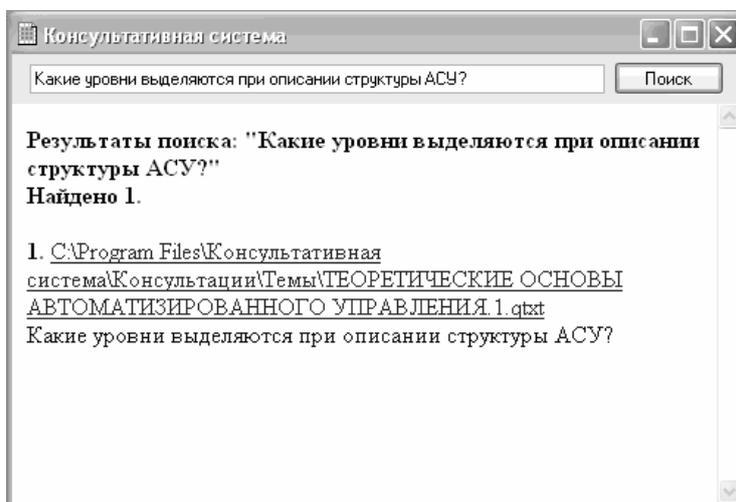


Рис. 6. Результат поиска заданного пользователем вопроса

После выбора результата, пользователь сразу приступает к чтению ответа, сопровождающегося соответствующими реакциями вопроса и ответа (Рис. 7). При желании пользователь может прочитать реакции на вопрос и ответ, а также начать продвижение вверх (Рис. 8) или вниз по ЛСС.

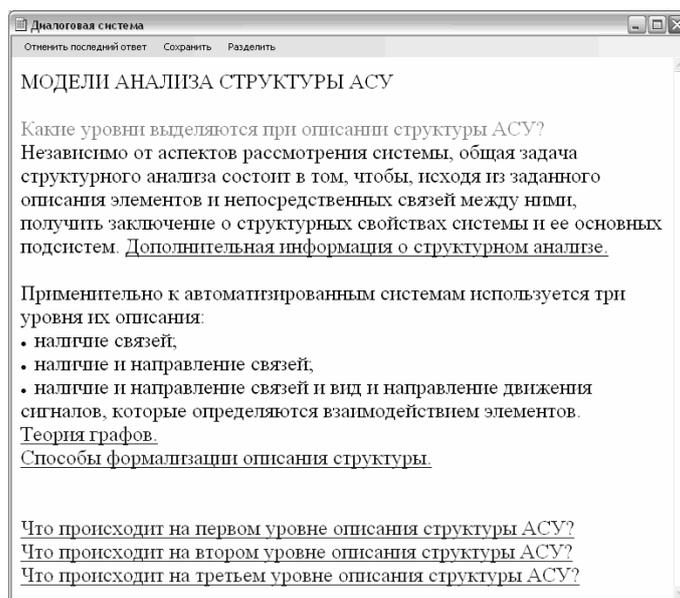


Рис. 7. Ответ на заданный пользователем вопрос

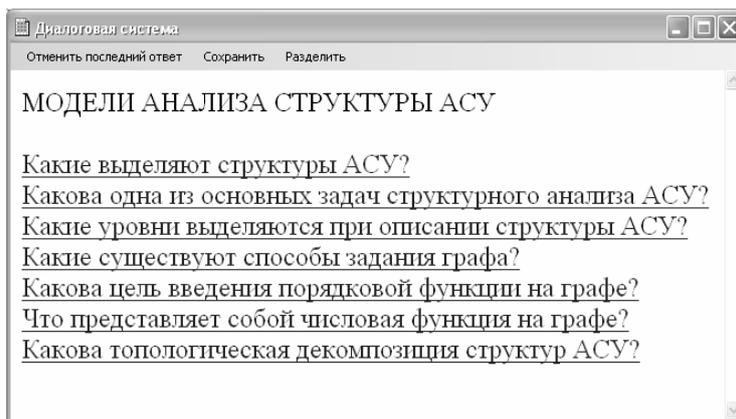


Рис. 8. Просмотр вопросов так же содержащихся в той же теме, что и вопрос, заданный пользователем

Использование логико-семантических сетей типа «Вопрос-Ответ-Реакция» — лишь один из возможных взглядов на способ построения семантических поисковиков. И если «Вопрос-Ответ» отнюдь не новые понятия для такой поисковой системы, то добавления «Реакции» позволяет добиться качественно нового уровня движения по логико-семантической сети, усиливает семантическую составляющую системы построенной по данному принципу. Семантические поисковые системы на основе ЛСС «Вопрос-Ответ-Реакция» имеют широкий спектр применимости, в том числе (и это можно считать одним из основных направлений) — в системе обучения, являясь средством активного усвоения учебных материалов.

В последнее время развитие ИТ было направлено по совершенно другому пути, нежели в самом начале развития эры компьютеризации. Тогда придумывали, как заставить компьютеры работать быстрее, как сделать винчестеры меньше, а мышки удобнее. Сейчас идет активное развитие интерфейсной части. Одна из главных задач, которую сейчас активно пытаются решить — это упрощение общения человека и компьютера. Создатели поисковиков пытаются сделать их все проще и понятнее для пользователей, а постоянный рост и развитие сети Интернет делают задачу поиска все более сложной. В этих условиях решением проблемы представляется создание поисковых машин, способных давать полноценный и семантически полный ответ на сформулированный на естественном языке вопрос. Разработка таких семантических поисковых систем сейчас идет полным ходом, и один из вариантов — это построение семантической поисковой системы на основе множества логико-семантических сетей.

Список литературы

1. Аверьянов, Л. Я. Почему люди задают вопросы? [Текст] / Л. Я. Аверьянов. — М. : «Социолог», 1993.
2. Добрынин, В. Н. Технологии извлечения знаний из текстов — средство активного усвоения учебных материалов [Текст] / В. Н. Добрынин, М. В. Лобачева // Сборник научных трудов. Выпуск 2. Международный университет природы, общества и человека «Дубна», 2008.
3. Кириллов, В.И. Логика [Текст] / В. И. Кириллов, А. А. Старченко. — М., 1987.