

# МЕТОД СЕМАНТИЧЕСКОГО ПРЕОБРАЗОВАНИЯ ОБОБЩЕННЫХ ЗАПРОСОВ НА ОСНОВЕ БАЗЫ ЦЕЛЕЙ

*А.В. Кириллов,*

*аспирант кафедры инноваций и бизнеса в сфере информационных технологий  
Национального исследовательского университета «Высшая школа экономики»*

*Адрес: г. Москва, ул. Кирпичная, д. 33/5*

*E-mail: antonv.kirillov@gmail.com*

*Статья описывает новый метод построения семантических расширений поисковых запросов обобщенного характера для улучшения результатов веб-поиска. Метод базируется на теории K-представлений – теории проектирования семантико-синтаксических анализаторов естественно-языковых текстов с широким использованием формальных средств для представления входных, промежуточных и выходных данных. Изложенный метод реализован на языке программирования Java в экспериментальной поисковой системе AOS (Aspect Oriented Search), эта система в настоящее время тестируется.*

**Ключевые слова:** семантический поиск, семантическое преобразование обобщенного запроса, концептуальный базис, СК-язык, концептуальная база целей, шаблон семантической трансформации.

## Введение

Сегодня параллельно с ростом объемов информации в Интернете происходит разработка новых и совершенствование существующих подходов к ее поиску [1]. Все большую актуальность приобретают средства семантического поиска, под которыми понимаются системы, принимающие на вход некоторый запрос, обрабатывающие его с использованием рассуждений над специфичной базой знаний и возвращающие совместимые результаты [2-8]. Входным запросом может являться, например, вопрос на естественном языке (ЕЯ), представление вопроса при помощи триплетов, графическое представление, набор

ключевых слов, отдельные фразы и т.д. В роли базы знаний могут выступать онтологии, аннотированные массивы текста, текстовые документы, веб, XML документы, RDF документы, HTML документы и т.д.

В данной работе предлагается решение для семантически-ориентированного поиска, использующее возможности традиционных поисковых систем, дополняемые семантическим анализом и расширением подаваемых на вход естественно-языковых запросов. С учетом вычислительных мощностей наиболее крупных существующих систем поиска по ключевым словам, предлагается сместить фокус с детального семантического анализа и ин-

дексации содержимого электронных документов на анализ подаваемых на вход системе запросов и последующую генерацию множества семантически расширенных (адаптированных) запросов, которые затем будут передаваться синтаксической поисковой системе. Результаты поиска по каждому из запросов будут анализироваться и сравниваться между собой с целью их фильтрации и ранжирования по степени семантической релевантности.

В работах автора и В.А. Фомичева [8-9] проводился анализ форм ЕЯ-запросов, и были выделены три наиболее интересных, с научной точки зрения, типа вопросов: вопросы достижения целей, аспектно-ориентированные вопросы и вопросы изменения множеств. Рассмотрим данные вопросы на примере информации об успехах компаний, продуктах, выпускаемых ими, и изменениях состава тех или иных организационных единиц компаний, представленных множествами (например, профсоюз или совет директоров). Последовательность поисковых запросов изображена на *рисунке 1*.

**Пример.** С учетом развития голосовых интерфейсов и средств синтеза и анализа речи пример поиска информации об успешности деятельности некоторой компании X и достигнутых результатах может быть представлен в виде следующего диалога:

*Пользователь:* Какие успехи были у компании X в этом году?

*Система:* Компания X выпустила продукт Y, увеличила объем выручки на 7%, открыла новое представительство в Москве.

*Пользователь:* Каковы особенности продукта Y?

*Система:* Высокая стабильность, отказоустойчивость, низкая цена и т. д.

*Пользователь:* Чем продукт Y отличается от продукта Z?

*Система:* Y более производителен чем Z, Y превосходит Z по следующим показателям ...

*Пользователь:* Какие изменения были в составе совета директоров в прошедшем году?

*Система:* В совет директоров компании вошел Петр Иванов.

Таким образом, если пользователь хочет найти информацию о компании, изменениях в составе ее руководства, успехах или неудачах, различных характеристиках продуктов, то полный цикл поиска покрывается предлагаемыми типами вопросов и соответствующими методами преобразования, тем самым увеличивая скорость, удобство и релевантность поиска.

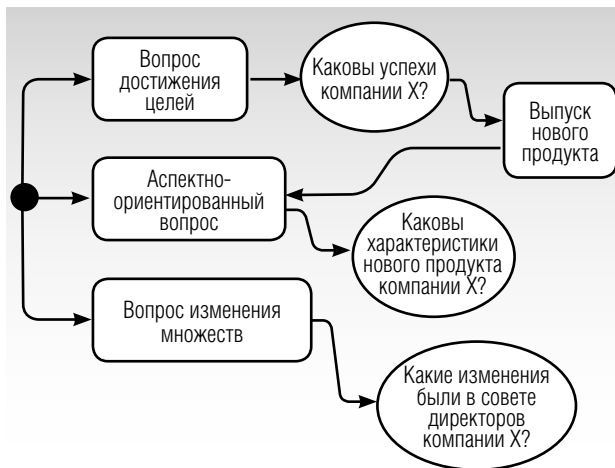


Рис. 1. Схема сценария бизнес-поиска.

Далее в статье будут подробно рассмотрены и структурированы вопросы достижения целей.

### Вопросы достижения целей

Успешность функционирования (либо существования) объекта или системы определяется достижением рассматриваемой сущностью целей, которые перед ней поставлены.

Вопросами достижения целей будем называть предложения, в которых запрашивается информация, касающаяся результатов функционирования того или иного объекта или системы. Другими словами, это вопросы, касающиеся успехов либо неудач рассматриваемых объектов или систем. В [9] такие вопросы названы вопросами о достижениях и неудачах.

Примеры вопросов достижения целей: «Какие успехи были у сборной России по футболу в 2009 году?», «Какие неудачи испытала компания Sun в 2010 году?», «Каковы успехи компании Intel за 2010 год?». Из представленных примеров видно, что если подать их на вход поисковой системе в таком виде, то результаты поиска будут низкого качества и не будут содержать в себе ответов на поставленный вопрос.

Для дальнейшего анализа была выбрана область предпринимательской деятельности: анализироваться будут успехи и неудачи организаций (компаний). Под целью компании будем понимать конечный желаемый результат, который определяется в процессе планирования и регулируется функциями управления. Рассмотрим фрагмент дерева целей отдельно взятой компании (*рисунок 2*):

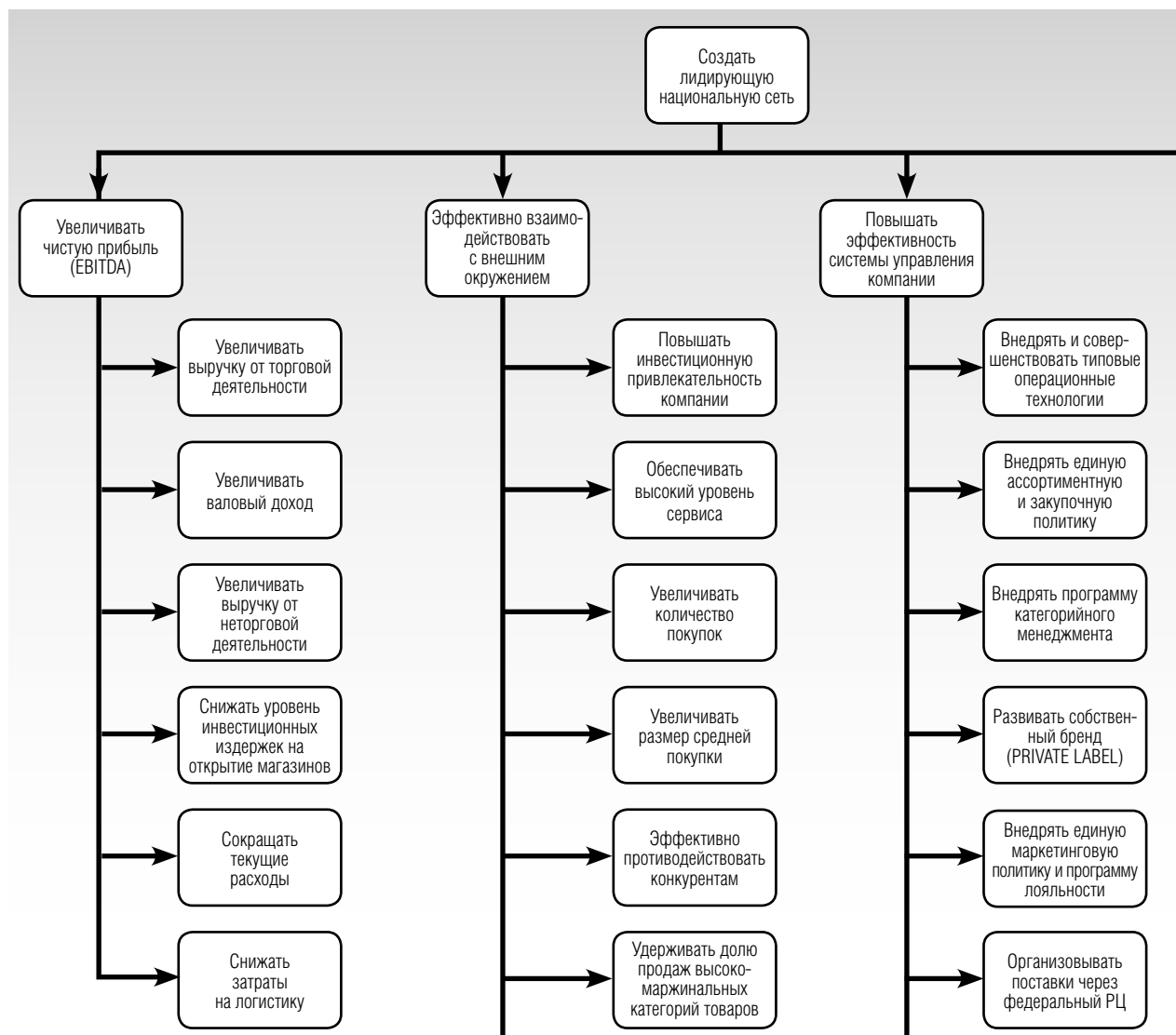


Рис. 2. Фрагмент дерева целей реально существующей компании.

Анализ данного дерева целей, а также деревьев целей других компаний показал, что успешность компании определяется фактами достижения компанией поставленных целей. Дальнейшая детализация целей и формирование фраз-индикаторов, определяющих нахождение в тексте указания на достижение цели, продемонстрировали, что, действительно, указания на детализированные цели встречаются в электронных документах гораздо чаще, чем описание непосредственно успехов или неудач каждой конкретной компании. Примеры текстов, извлеченных поисковой системой и указывающих на достижение некоторых из рассматриваемых показателей, представлены в табл. 1.

Таким образом, чтобы обнаружить документы, содержащие в себе информацию об успехах либо

неудачах компании, необходимо декомпозировать первоначальный ЕЯ-запрос на множество словосочетаний – индикаторов: ЕЯ-выражений, наличие которых в тексте документов позволяет определить соответствие этого документа первоначальному запросу. Анализ целей нескольких компаний позволил обобщить эти цели и сформировать факты, указывающие на достижение компаниями этих целей (табл. 2). В данной таблице обозначение *pos* соответствует тому факту, что данное событие является индикатором успешности деятельности компании, *neg* соответствует неудаче. Пара (+, *pos*) обозначает, что рост данного показателя является позитивным фактором, (-, *pos*) обозначает, что снижение данного показателя является позитивным фактором.

Таблица 1.

Примеры текстов,  
извлеченных поисковой системой

Цели	Примеры текстов
Увеличивать выручку	<b>Выручка</b> индийской <b>компании</b> в отчетный период <b>увеличилась</b> в 3,2 раза - до 31,5 млн долл. <b>Выручка</b> РБК <b>увеличилась</b> на 46% благодаря уверенному росту основного бизнеса <b>компании</b> .
Увеличивать доход	<b>Компания</b> «Визард» после внедрения МАНХ <b>увеличила доход</b> более чем в два раза по сравнению с докризисными показателями. <b>Доход компании</b> Nусomed в I квартале <b>увеличился</b> на 30 % .
Увеличивать прибыль	Чистая <b>прибыль</b> АФК «Система» по US GAAP за 9 месяцев 2010г. составила 471 млн 452 тыс. долл. против 2 млрд 108 млн 46 тыс. долл. годом ранее. Чистая <b>прибыль компании</b> ЛУКОЙЛ за первое полугодие 2006 года <b>увеличилась</b> на 55 процентов.
Снижать затраты	Совет директоров <b>ЛУКОЙЛа</b> принял решение, что капитальные <b>затраты компании</b> в 2009 году будут <b>снижены</b> почти вдвое.
Сокращать расходы	<b>Компания</b> Intel планирует в текущем году сократить свои <b>расходы на</b> несколько сотен миллионов долларов.
Применять новейшие технологии и ресурсы	<b>Компания</b> Softline <b>внедрила технологии управления</b> лицензиями на продукты Adobe Systems в ЗАО «Международный центр финансово-экономического развития». Компания <b>Lenovo внедрила</b> в линейку бизнес ноутбуков ThinkPad <b>технологии</b> NVIDIA Optimus.
Эффективно противодействовать конкурентам	<b>Компания Google</b> <b>поглотила</b> фирму VmрrТор, занимающуюся разработкой инновационного трехмерного интерфейса для управления рабочим столом с активным использованием мультитач навигации.
Контролировать соблюдение стандартов	Производство <b>компании</b> «Союз-Виктан» <b>соответствует стандартам</b> ISO (ИСО) 14000.
Выходить на новые рынки	<b>Компания WELHOME вышла на рынок</b> инвестиционных услуг с новыми предложением. <b>Компания ZTE вышла на рынок</b> планшетных компьютеров.

Таблица 2.

Сгруппированные факторы достижения целей компаний

КОМПАНИЯ				
Внешние факторы	Внутренние факторы		Изменения	
поглощение компании (pos) поглощение компанией (neg) лидер рынка\сферы\отрасли\ сегмента\направления (pos) выпуск продукта (pos) открытие нового офиса\представительства (pos) заключение выгодной\успешной сделки\контракта (pos) создание нового предложения (pos) выход на новый рынок (pos) завоевание рынка (pos) медаль\приз (pos) победа\выигрыш в конкурсе (pos) работодатель года (pos) компания года (pos) заключить контракт (pos) расторгнуть контракт (neg)	автоматизация оптимизация улучшение модернизация	производство процесс оборудование технология	увеличение уменьшение повышение снижение расширение	
	соответствие	требования стандарты		объем продаж (+, pos) рейтинг (+, pos) привлекательность (+, pos) выручка (+, pos) доход (+, pos) прибыль (+, pos) уровень сервиса (+, pos) стоимость бренда (+, pos) реализация продкции (+, pos) число офисов (+, pos) расходы (-, pos) затраты (-, pos) себестоимость продукции (-, pos) количество партнеров (+, pos) число поставщиков (+, pos) ассортимент (+, pos)
	внедрение	система методика методология технология		

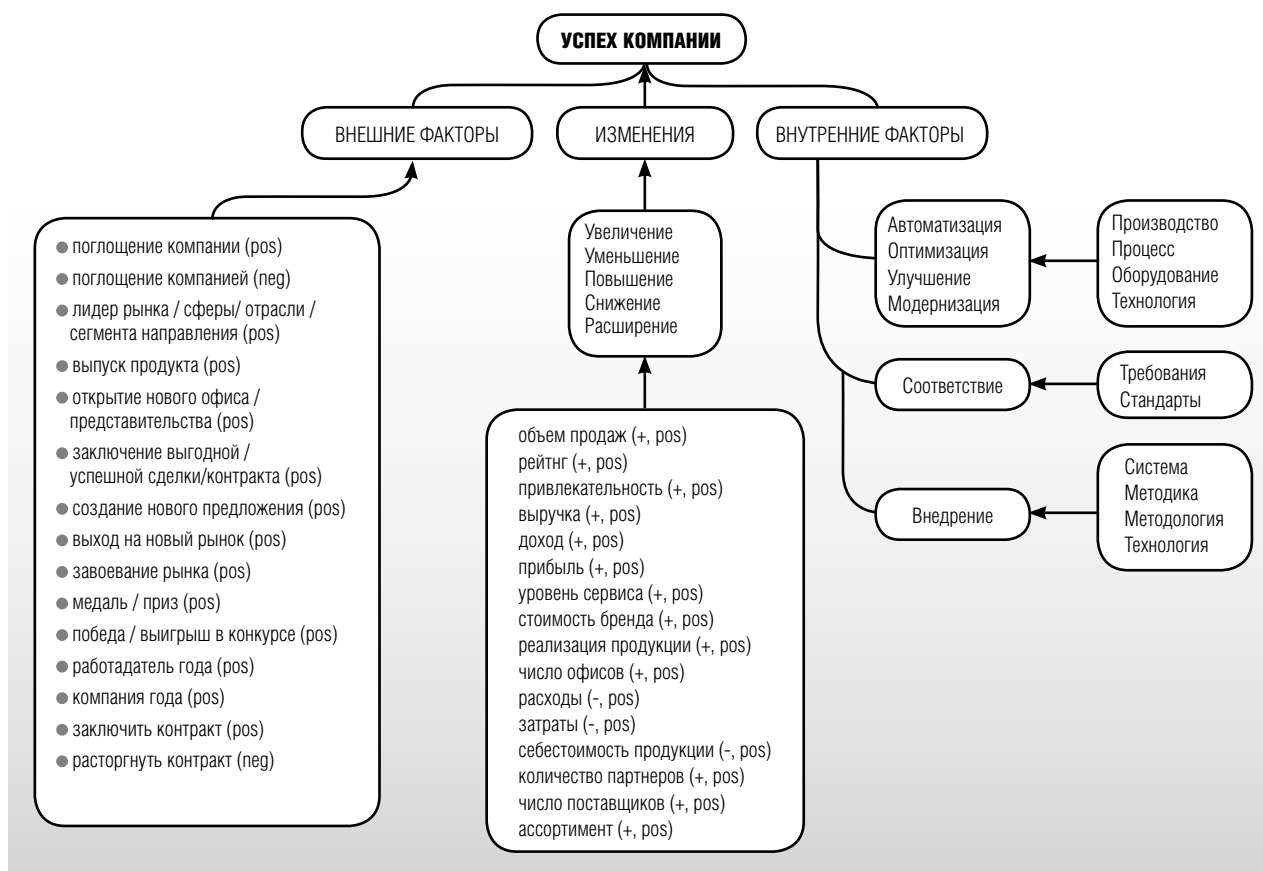


Рис. 3. Графическое представление дерева факторов достижения целей.

**Формальная модель базы знаний для представления целей**

Представим таблицу 2 в виде графической схемы (рис. 3) для большей наглядности:

На представленной схеме видно, что факторы достижения успеха имеют общую форму.

В монографиях В.А. Фомичева [10-12] определен класс формальных объектов, названных концептуальными базисами (к.б.) и задающих базовые сведения о системе первичных концептуальных единиц, используемой прикладной интеллектуальной системой.

**Определение 1.**

Упорядоченная тройка вида  $(B, цел, рац)$  называется концептуальным базисом с числовой разметкой  $\Leftrightarrow$  когда  $B$  – произвольный концептуальный базис,  $цел$  и  $рац$  – два различных сорта из  $Sf(B)$ , и выполняются следующие условия:

1) первичный информационный универсум  $X(B)$  включает подмножества  $Natural$ ,  $Pos-rational$ ,  $Z1$ ,  $Z2$ ,  $Numbers$ , где  $Natural$  – множество всех цепочек

вида  $d_1, \dots, d_n$ , где  $n \geq 1$ , и для  $k=1, \dots, n$   $d_k$  – цифра из множества  $\{‘0’, ‘1’, \dots, ‘9’\}$ , причем из  $d_1 = 0$  следует, что  $n = 1$ ;  $Pos-rational$  – множество всех цепочек вида  $b, c$ , где  $b, c \in Natural$ ;  $Z1$  – объединение множества  $Natural$  и множества всех цепочек вида  $-nb$ , где  $nb$  – цепочка из множества  $Natural$ ;  $Z2$  – объединение множества  $Pos-rational$  и множества всех цепочек вида  $-numb$ , где  $numb$  – цепочка из множества  $Pos-rational$ ;  $Numbers$  – объединение множеств  $Z1$  и  $Z2$ ;

2) для каждого элемента  $d$  множества  $Z1$  типом  $tp(d)$  является сорт  $цел$ , и для каждого элемента  $h$  множества  $Z2$  типом  $tp(h)$  является сорт  $рац$ ;

3) сорт  $цел$  является конкретизацией сорта  $рац$  для отношения общности  $Gen$ .

**Пример.**

Множество  $Natural$  включает цепочки 123 и 4125; множество  $Numbers$  включает, в частности, цепочки 12,78, -0,315 и -542.

**Определение 2.**

Концептуальной базой целей (к.б.ц.) называется набор  $Gbase$  вида

$(B, \text{цел, рац, событие, Goals}), \quad (1)$

где тройка  $(B, \text{цел, рац})$  является концептуальным базисом с числовой разметкой, *событие* является выделенным сортом из множества  $St(B)$ , и выполняются следующие условия:

1) множество  $Acts(B) = \{y \text{ из } X(B) \mid tp(y) = \uparrow \text{событие}\}$  не пусто и конечно;

2) первичный информационный универсум  $X(B)$  включает такой элемент *#Объект-интереса*, что тип  $tp(\text{\#Объект-интереса})$  является конкретизацией базового типа [*объект*];

3) универсум  $X(B)$  включает элементы (обозначения тематических ролей) *Агент* и *Роль2*, причем  $tp(\text{Агент}) = \{(\text{событие}, s)\}$ ,  $tp(\text{Роль2}) = \{(\text{событие}, t)\}$ , где  $s$  и  $t$  – элементы множества сортов  $St(B)$ , причем ни один из этих сортов не является конкретизацией сорта  $P$  (сорт «смысл сообщения») или сорта *событие*;

4) *Goals* является некоторым конечным множеством выражений СК-языка  $Ls(B)$  вида

$\langle \text{event-concept} * (\text{Агент}, d1)(\text{Роль2}, d2), \text{estimation} \rangle,$

где *event-concept* принадлежит множеству  $Acts(B)$ ,  $d1$  и  $d2$  – различные элементы универсума  $X(B)$ , причем какой-либо из этих элементов является символом *#Объект-интереса*, *estimation* – элемент множества *Numbers*, обозначающий рациональное число от -1 до 1, отличное от 0.

**Пример.**

Множество *Goals* может включать цепочку  $\langle \text{поглощение-орг} * (\text{Агент}, \text{\#Объект-интереса})(\text{Роль2}, \text{нек компания}), 1 \rangle.$

Данная модель позволит хранить записи о фактах успехов либо неудач какой-либо компании. Хранение информации в базе знаний именно в такой форме позволяет с легкостью генерировать выражения, которые послужат в качестве новых поисковых запросов, а также расширять спектр данных, хранящихся в базе. База целей может постоянно расширяться, позволяя дополнять ее новыми фактами, относящимися к успехам либо неудачам компаний.

Однако для семантического расширения поисковых запросов этих данных недостаточно. Необходимо иметь возможность преобразовывать данные факты в словосочетания, которые наиболее вероятно встречаются в электронных документах. Следует учитывать, что преобразование должно учитывать морфологические свойства слов и синтаксис фраз,

чтобы наиболее гибко трансформировать факты в ЕЯ-выражения. Для этих целей введем понятие шаблона семантической трансформации.

**Определение 3.**

Пусть *Gbase* – концептуальная база целей (к.б.ц.) вида (1), *Morph-values* – конечное множество символов, интерпретируемых как значения различных морфологических признаков (*существит, глагол, прош-время, наст-время, пассив-зalog* и т.д.). Тогда шаблоном семантической трансформации, порожденным к.б.ц. *Gbase* и множеством *Morph-values*, называется произвольный упорядоченный набор вида

$(\text{sem-pattern}, X, Y, Z, \text{prop-chain}), \quad (2)$

где *sem-pattern* – элемент множества *Goals*,  $(X, Y, Z)$  – произвольная перестановка без повторений из символов *#A#*, *#Pred#*, *#B#*, и *prop-chain* – цепочка вида  $v[1] * v[2] * \dots * v[k]$ , где  $1 \leq k$ ,  $v[1], \dots, v[k]$  – элементы множества *Morph-values*.

**Пример.**

Пусть *sem-pattern* – цепочка  $\langle \text{поглощение-орг} * (\text{Агент}, \text{\#Объект-интереса})(\text{Роль2}, \text{нек компания}), 1 \rangle.$  Тогда набор  $(\text{sem-pattern}, \text{\#A#}, \text{\#Pred#}, \text{\#B#}, \text{глаг} * \text{прош-время})$  является одним из возможных шаблонов семантической трансформации.

**Методика преобразования вопросов достижения целей к расширенному виду**

Преобразование происходит при помощи правил трансформации, уникальных для каждого из фактов. В правиле трансформации указываются порядок слов в результирующем словосочетании, а также их формы. На основании хранимых в базе фактов и соответствующих правил трансформации генерируются словосочетания, которые позволят традиционной системе поиска по ключевым словам найти документы, содержащие в себе упоминания данных фактов. После того, как множество словосочетаний-индикаторов сформировано, оно передается на вход синтаксической поисковой системе, которая осуществит отдельный поиск по каждому из них. Множество результирующих документов, возвращенных поисковой системой будет проанализировано с точки зрения количества вхождений различных словосочетаний-индикаторов, т.е. указателей на присутствие упоминания того или иного факта в документе. Релевантность документов будет опре-

деляться во-первых, по количеству вхождений различных фактов, а во-вторых, по рейтингу документа, определяемому по алгоритму *PageRank*. Отсортированные по релевантности первоначальному поисковому запросу документы затем будут переданы пользователю.

*Метод порождения результирующих индикаторов.*

1. По входному запросу находится информационная единица *Studied-object*, обозначающая объект интереса запроса. Эта информационная единица интерпретируется как значение символа *#Объект-интереса*.

2. По шаблону вида (2) строится цепочка *X Y Z*.

3. В цепочке *X Y Z* символ *#A#* заменяется на произвольную лексическую единицу, которой соответствует информационная единица *Studied-object*.

4. Пусть *sem-pattern* – цепочка вида *<event-concept\*(Агент, d1)(Роль2, d2), estimation>*. Тогда в цепочке, полученной на Шаге 3, символ *#Pred#* заменяется на произвольную лексическую единицу *pred-word-form*, которой соответствует информационная единица *event-concept*, причем словоформа *pred-word-form* должна обладать значениями морфологических признаков, задаваемыми цепочкой *prop-chain*.

5. В цепочке, полученной на Шаге 4, символ *#B#* заменяется на произвольную лексическую единицу, соответствующую той из семантических единиц *d1* и *d2*, которая отлична от символа *#Объект-интереса* в исходном шаблоне.

**Пример 1.**

Пусть *Studied-object = firm-Oracle*, *sem-pattern* – цепочка *<поглощение-орг \* (Агент, #Объект-интереса)(Роль2, нек фирма1), 1>*. Тогда по шаблону семантической трансформации вида (*sem-pattern, #A#, #Pred#, #B#, глаг \* ПрошВр*) в соответствии с данным алгоритмом может быть построен результирующий индикатор *Oracle поглотил компания*.

**Пример 2.**

Пусть начальный запрос *w = «Каковы успехи компании Oracle?»*. После определения типа запроса и объекта интереса [8] как пары (успех, Oracle) можно переходить к построению расширенного множества запросов. Примеры записанных фактов, шаблонов трансформации и результирующих словосочетаний-индикаторов представлены в таблице 4.

Запись факта	Правило трансформации	Результирующий индикатор
<поглощение-орг*(Агент, #Объект-интереса)(Роль2, компания), 1>	#A+#Pred + #B + глаг * ПрошВр	Oracle поглотил компания
<выпуск*(Агент, #Объект-интереса)(Роль2, нек продукт1), 1>	#A+#Pred + #B + глаг * ПрошВр	Oracle выпустил продукт
	#B+#Pred + #A + глаг * ПрошВр *ЗалогПасс	Продукт выпущен Oracle
	#A+#Pred + #B + глаг * ПрошВр * ЗалогПасс	Oracle выпущен новый продукт
<открытие*(Агент, #Объект-интереса)(Роль2, нек офис), 1>	#A+#Pred + #B + глаг * ПрошВр	Oracle открыл новый офис
<увеличение-знач-парам*(Агент, #Объект-интереса)(Роль2, объем-продаж), 1>	#A+#Pred + #B + глаг * ПрошВр	Oracle увеличил объем продаж
		Oracle повысил объем продаж
	#B+#A + #Pred + глаг * ПрошВр	Объем продаж Oracle увеличился
		Объем продаж Oracle возрос
		Объем продаж Oracle вырос
<заклучение1 * (Агент, #Объект-интереса)(Роль2, нек сделка), 1>	#A+#Pred + #B + глаг * ПрошВр	Oracle заключил сделку
	#B+#Pred + #A + глаг * ПрошВр *ЗалогПасс	Сделка заключена Oracle
	#B + #A	Успешная сделка Oracle

**Заклучение**

Часто для вопросов на естественном языке отсутствуют методы их анализа для повышения качества поиска, особенно это характерно для вопросов определенных типов. Рассмотренные вопросы достижения целей представляются интересными для исследования и анализа. Для адекватного поиска ответов на такие вопросы был предложен метод, основанный на декомпозиции целей организаций и генерации множества детализирующих запросов, содержащих индикаторы, позволяющие определить соответствие текста документов первоначальному запросу. В результате работы семантического анализатора часть документов будет отсеяна, будет учтено

дублирование документов и количество вхождений индикаторов в документ. На выходе пользователь получит набор документов, упорядоченный по количеству фактов успехов либо неудач той или иной компании. Факторам могут быть присвоены веса, которые будут учитываться при ранжировании результирующего множества документов.

Предложенные модели базы целей и шаблонов семантических трансформаций могут легко быть дополнены и расширены, а использование выражений СК-языков для представления знаний позволяет очень гибко манипулировать описаниями информации, хранимой в базе целей.

Использование данного подхода значительно расширяет возможности поиска, потому что дополняет его возможностями рассуждений. Ведь методика, основанная на детализации целей и генерации словосочетаний-индикаторов, выходит за рамки построения обычных отношений вида is-a, применяемых в системах поиска, основанных на семантических сетях, либо использующих как базу знаний какую-либо таксономию. Изложенный метод реализован на языке программирования Java в экспериментальной поисковой системе AOS (Aspect Oriented Search), эта система в настоящее время тестируется. ■

### Литература

1. Кириллов А.В. Поисковые системы: компоненты, логика и методы ранжирования // Бизнес-информатика. 2011. № 4(10). – С. 51-59.
2. Halpin H., Lavrenko V. Relevance Feedback Between Hypertext and Semantic Search // Proc. Conference WWW2009 (April 20-24, 2009, Madrid, Spain).
3. Lei Y., Uren V., Motta E.. Semsearch: A search engine for the semantic web // Proc. 5th International Conference on Knowledge Engineering and Knowledge Management Managing Knowledge in a World of Networks, Lect. Notes in Comp. Sci. – Springer, Podybrady, Czech Republic, 2006. – pp. 238-245.
4. Bhagdev R., Chapman S., Ciravegna F., Lanfranchi V., Petrelli D. Hybrid search: Effectively combining keywords and semantic searches // The Semantic Web: Research and Applications. – Springer, Berlin / Heidelberg, 2008. – pp. 554-568
5. d'Aquin M., Baldassarre C., Gridinoc L., Angeletou S., Sabou M., Motta E. Characterizing knowledge on the semantic web with watson // EON, 2007. – pp. 1-10.
6. Kaufmann E., Bernstein A., Zumstein R. Querix: A Natural Language Interface to Query Ontologies Based on Clarification Dialogs // 5th International Semantic Web Conference (ISWC 2006). – Springer, 2006. – pp. 980-981.
7. Bernstein A., Kaufmann E., Gohring A., Kiefer C. Querying Ontologies: A Controlled English Interface for End-users // 4th International Semantic Web Conference (ISWC 2005), November 2005. – pp. 112-126.
8. Кириллов А.В., Фомичев В.А. О новом подходе к семантическому преобразованию естественно-языковых запросов поисковых систем // Бизнес-информатика. № 1 (15). С. 19-26.
9. Fomichov V.A., Kirillov A.V. Semantic Transformation of Search Requests for Improving the Results of Web Search // Pre-Conference Proceedings of the Focus Symposium on Knowledge Management Systems (August 2, 2011, Focus Symposia Chair: Jens Pohl) in conjunction with InterSymp-2011, 23rd International Conference on Systems Research, Informatics and Cybernetics, August 1 – 5, 2011, Baden-Baden, Germany), Collaborative Agent Design Research Center, California Polytechnic State University. – San Luis Obispo, CA, USA, 2011. – pp. 37-43.
10. Фомичев В.А. Формализация проектирования лингвистических процессоров. – М.: МАКС Пресс, 2005. – 368 с.
11. Фомичев В.А. Математические основы представления содержания посланий компьютерных интеллектуальных агентов. – М.: Изд. дом ГУ-ВШЭ, изд. «ТЕИС», 2007. – 176 с.
12. Fomichov V.A. Semantics-Oriented Natural Language Processing: Mathematical Models and Algorithms / Series: IFSR International Series on Systems Science and Engineering, Vol. 27. – Springer: New York, Dordrecht, Heidelberg, London, 2010. – 354 p.