

**НАУЧНО-  
ТЕХНИЧЕСКАЯ  
ИНФОРМАЦИЯ**

ISSN 0548-0027



СЕРИЯ 2 · ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И СИСТЕМЫ

**2** 1986

# АВТОМАТИЗАЦИЯ ОБРАБОТКИ ТЕКСТА

УДК 801

## К ОПИСАНИЮ ВОПРОСО-ОТВЕТНОГО ОТНОШЕНИЯ

Е. В. Рахилина

Создание модели человеко-машинного диалога на естественном языке предполагает, как известно, прежде всего описание вопросо-ответного отношения, т. е. отношения, которое, связывая вопрос и ответ на него, создает минимальную единицу диалога, своеобразный микродиалог — вопросо-ответную пару. Каково это отношение, т. е. как в действительности человек строит ответ на заданный ему вопрос? В данной работе указываются некоторые трудности, связанные с описанием вопросо-ответного отношения\*.

В свое время М. Н. Петерсон, рассматривая различные употребления вопросительных предложений, отмечал существование особого типа вопросов, названные им «школьными» [1]. Это вопросы, используемые в школьной практике для распознавания второстепенных членов предложения и их грамматических форм: *платок какой? — теплый* (определение); *купил что? — книгу* (прямое дополнение, винительный падеж)\*\*. Основной недостаток этой практики заключается, по лаконичному выражению М. Н. Петерсона, в том, что подобные вопросы «не встречаются ни в жизни, ни в литературе\*\*\*».

Независимо от того, насколько изменилось положение в педагогической практике к настоящему времени, критические высказывания А. М. Пешковского и М. Н. Петерсона, по нашему мнению, до сих пор не утратили актуальности. Дело в том, что школьные вопросы нельзя считать простой случайностью неудачной методики. Подобное метаязыковое употребление вопросов опирается на некоторую простейшую модель вопросо-ответной пары. Согласно этой модели, ответ на вопрос строится с помощью операции субституции: для получения ответа вопросительное слово должно быть заменено языковым выражением из определенного синтаксического класса, таким, что при этой замене получилось бы грамматически правильное предложе-

ние\*. Приведем типичный пример вопросо-ответной пары, построенной по этой модели, из [3]:

(1) *Какие цветы растут в саду? — В саду растут красивые цветы.*

Такая модель могла бы претендовать на адекватное описание языковой реальности, если бы цель отвечающего заключалась в воспроизведении в своей реплике синтаксической структуры, заданной вопросом. Как представляется, нет необходимости специально объяснять, что в действительности цель отвечающего совсем иная: он должен прежде всего сообщить спрашивающему некоторую информацию, а сделать это можно, как правило, различными способами, и, в частности, совершенно игнорируя синтаксическую структуру вопроса, ср.:

(2) — *Скажите, маркиз, как я должен поступить по карточным правилам в таком странном случае?*

— *Сир, вам надлежит ударить его по физиономии подсвечником.* (Булгаков)

(3) — *Да где же денег взять прикажете? [Ведь восемь тысяч на улице не найдешь?]*

— *Пускай твой господин продаст хоть тебя самого, а мне он заплатит в назначенное время.* (Лермонтов)

Следовательно, из всего множества способов передачи требуемой информации простейшая модель, которой, как мы видим, руководствуется не только школьная, но и лингвистическая традиция, признает лишь синтаксическую субституцию.

Действительно, пользуясь этой моделью, можно получить некоторое количество правильных ответов на данный вопрос; в этом смысле метод субституции оправдан, если принять во внимание его простоту. Однако полученное таким образом (полное) множество ответов не будет ни достаточным, ни необходимым. Многие ответы, бесспорно, правильные (и, возможно даже, наиболее «естественные», если здесь уместно это нетерминологическое употребление), заведомо останутся за пределами полученного синтаксического класса — ср. примеры (2) и (3). С другой стороны, среди элементов этого класса обязательно окажутся и такие выражения, которые никак нельзя будет признать правильными ответами. К ним относится, например, ответ в (1): если мы имеем дело с реальной ситуацией, в которой спрашивающему действительно необходимо узнать, какие цветы растут в саду, такой ответ может быть воспринят только как нарочито неправильный,

\* Критерий выбора такого синтаксического класса определяется, например, следующим образом: «Удобным способом описания класса выражений, соответствующих вопросительным элементам, ... является сопоставление этому классу неопределенного местоимения: считается что в некоторый класс входят те выражения, которые могут быть замещены сопоставленным ему неопределенным местоимением» [3].

\* Первые варианты настоящей работы были внимательно прочитаны Г. Е. Крейдлиным, Е. В. Падучевой, В. А. Плунгяном и А. К. Поливановой, советы и замечания которых в значительной степени помогли автору при переработке статьи.

\*\* Разнообразные критические замечания в адрес этой методики высказывались и до М. Н. Петерсона (см., например, статью А. М. Пешковского [2]).

\*\*\* По отношению к методическому средству обучения это весьма серьезный упрек — ведь сказанное фактически означает, что такие вопросы не облегчают школьнику усвоения грамматических понятий, апеллируя к его языковой компетенции, а ставят его перед необходимостью запоминать и сам механизм запоминания.

«издевательский»\*. Правильными были бы ответы типа *Розы, тюльпаны и маки*. Это замечание становится весьма существенным, если мы решаем задачи, связанные с построением программы автоматического синтеза правильных ответов на заданный вопрос: даже если мы не имеем в виду получить все такие ответы, мы, бесспорно, должны воспрепятствовать появлению ответов типа (1).

«Спасти» идею субституции, вообще говоря, можно было бы, например, с помощью расширения самого понятия субституции. Такая попытка сделана в [3] для *какой-вопросов*. Однако, описав таким образом один тип вопросо-ответных пар со словом *какой*, авторы оказываются перед проблемой описания других типов употребления этого вопросительного слова, требующих, по-видимому, принципиально иных решений (подробнее см. [4]).

Очевидно, что причины, в силу которых некоторый правильный ответ не попадает в синтаксический класс, задаваемый с помощью вопросительного слова, различны; соответственно, различаются и типы таких «несинтаксических ответов». Нам представляется, что на данном этапе исследования полезно обратить внимание на следующее различие: (а) с одной стороны, существуют «несинтаксические ответы», для которых можно указать синонимичные им выражения, порождаемые с помощью синтаксического метода; (б) с другой стороны, существуют такие «несинтаксические ответы», для которых указанное соответствие построить принципиально невозможно; ср., в частности, пример (3).

Естественно, что для задач автоматического синтеза наиболее важным является класс (б), так как «потеря» ответов типа (а) при моделировании, по-видимому, всегда можно пренебречь. Ниже мы рассмотрим примеры, которые, с нашей точки зрения, могут дать представление о действительной сложности задач, связанных с описанием вопросо-ответного отношения.

Рассмотрим сначала одну разновидность вопросов со словом *где*, а именно: такие вопросы, которые содержат связку *быть* в настоящем актуальном времени, поверхностно выраженную нулевой формой\*\*:

- (4) — *Где Арбенин?*  
— *Ушел... не слышит и не видит; как бешеный бросился в дверь...* (Лермонтов)
- (5) — *Да где же тот проклятый мешок?*  
— *Я кинула его в угол, там больше ничего нет.* (Гоголь.)

При анализе таких вопросо-ответных пар обнаруживается, что в большинстве случаев в ответе на *где*-вопрос место нахождения объекта вообще не указывается, ср. (4), а также:

- (6) — *Друзья мои... Где же они все?*  
— *Едва я вошел, как они вышли на цыпочках.* (Шварц).
- (7) — *«Скажите, где же ваш приятель?»*  
*Ему вопрос хозяйки был: —*

\* В работе [3], из которой заимствован пример (1), метод субституции провозглашен «общим подходом к определению отношения «текст Т может служить ответом на вопрос Q». Показательны и другие примеры, приводимые в [3] в качестве иллюстраций правильно построенных вопросо-ответных пар:

(а) — *Сколько зерна привезли на склад?*  
— *На склад привезли уйму зерна.*

(б) — *Как проявляют фотопленку?*  
— *Фотопленку проявляют аккуратно.*

(в) — *Как называется соседняя деревня?*  
— *Неудачно (так, что и не выговоришь).*

(г) — *Как передать электрический заряд от одного тела к другому?*  
— *Передать заряд надо быстро и умело.*

Все эти примеры в действительности представляют собой типичные «школьные» вопросо-ответные пары, которые «не встречаются ни в жизни, ни в литературе».

\*\* Это наиболее распространенная разновидность вопросов со словом *где*; не будет преувеличением сказать, что в восьмидесяти случаях из ста вопросительное слово *где* встречается именно в таком контексте.

*Он что-то нас совсем забыл.*

«Сегодня быть он обещал, —  
Старушке Лечский отвечал, —  
Да, видно, почта задержала». (Пушкин)

Эти и многие другие примеры иллюстрируют такое употребление *где*-вопроса, при котором правильным, удовлетворительным оказывается «неместный» ответ. Лицу, задающему *где*-вопрос, всякий раз достаточно той информации, которая сообщается в подобного рода ответах, спрашивающий ведет себя так, как если бы он не был заинтересован в том, чтобы узнать место нахождения объекта, — другими словами, как если бы и не спрашивал, *где*? Только так можно объяснить и то обстоятельство, что во многих случаях замена реального «неместного» ответа на «местный» дает неудачную вопросо-ответную пару с *где*-вопросом, ср.:

- (8) — *Где же гости?* — *спросила Маргарита у Коровьева.*  
— *Будут, королева, будут, сейчас будут.* (Булгаков)

По всей видимости, Маргариту, ожидающую гостей, не удовлетворили бы сведения о действительном их местонахождении.

- (9) — *Где же учитель?* — *спросил он у вошедшего слуги.*  
— *Нигде не найдут-с, — отвечал слуга.* (Пушкин)

Слуга не знает, где учитель, которого ищут по всему дому. Но если бы знал и на вопрос Троекурова ответил бы, например: *В саду*, — вряд ли Троекуров был бы доволен таким поведением своего слуги: совершенно очевидно, что он не о том спрашивал. Еще более показателен пример:

- (10) — *Где Алеша с Николкой, в самом деле?*  
— *Я сам беспокоюсь... Пять минут жду, а после этого иду навстречу.* (Булгаков)

Отвечающий знает почти наверняка, где должны быть Алеша с Николкой, — иначе он не знал бы, куда ему идти. Между тем, он реагирует на заданный вопрос так, как если бы не был в состоянии на него ответить, поскольку никаких сведений о том, где находятся Алеша с Николкой, он не сообщает. Более того, спрашивающий и не ожидает, что ему будут сообщены эти сведения, так как он оказывается совершенно удовлетворен полученным ответом.

В примере:

- (11) — *Ступай на улицу... или нет, стой!*... *Да другие-то где? Неужели ты только один? Ведь я приказывал, чтобы и Прохоров был здесь.* *Где Прохоров?*

— *Прохоров в частном доме, да только к делу не может быть употреблен...: привезли его поутру мертвецки.* (Гоголь)

ответ содержит прямое указание на местонахождение объекта: *Прохоров в частном доме*. Однако, если бы отвечающий ограничился только этим высказыванием, спрашивающий явно был бы неудовлетворен. Напротив, как это ни парадоксально, если бы он случайно «забыл» сказать, где Прохоров, провала коммуникативного акта, по-видимому, не было бы, ср.:

- (11а) — *Ступай на улицу... или нет, стой? Да другие-то где? Неужели ты только один? Ведь я приказывал, чтобы и Прохоров был здесь. Где Прохоров?*

— *Прохоров к делу не может быть употреблен... привезли его поутру мертвецки.*

Все это свидетельствует о том, что подобные вопросо-ответные пары нельзя квалифицировать как содержащие вопрос (и, соответственно, требующие ответ) о местонахождении.

Целесообразно считать, что в таких случаях в ответе указывается причина отсутствия объекта в некотором пространстве, находящемся в поле зрения спрашивающего. Иначе говоря, отвечающий на вопрос *Где X?* часто строит свою реплику так, как если бы он отве-

чал на вопрос *Почему X-а здесь нет?* Ср. в этой связи замечание Е. В. Падучевой, высказанное в связи с проблемой классификации ответов: «На интуитивном уровне различие между прямым и непрямым ответом обычно очевидно. (Есть, впрочем, и пограничные случаи. Так, на вопрос *Где Ваня?* ответная реплика *Он ушел* в определенной ситуации воспринимается как прямой ответ. Вероятно, дело в том, что этот вопрос может приобретать нестандартную семантику и означать нечто вроде *Почему я не вижу Вани?*») [5].

Действительно, если понимать *где*-вопросы в приведенных выше примерах именно так, окажется, что реакция отвечающего вполне естественна, ср.:

- (6') — *Друзья мои... Почему же их нет?*  
— *Едва я вошел, как они вышли на цыпочках.*  
(7') — *Почему с вами нет вашего приятеля?*  
— *Сегодня быть он обещал... да, видно, почта задержала.*  
(8') — *Почему же нет гостей?*  
— *Будут, королева, будут, сейчас будут.*  
(9') — *Почему здесь нет учителя? (≈ 'почему до сих пор учителя не привели сюда')*  
— *Нигде не найдут-с.*  
(10') — *Почему нет Алеши с Николкой, в самом деле?*  
— *Я сам беспокоюсь... Пять минут жду, а после этого иду навстречу.*

Такое причинное понимание *где*-вопросов возможно не всегда, а только в том случае, когда говорящий обнаруживает несоответствие между имевшейся у него презумпцией о том, что объект *X* должен находиться в поле его зрения, и реальным положением вещей.

Так, герой пьесы Е. Шварца (6') только что был окружен своими друзьями, он обращается к ним: «Друзья мои!» — и вдруг обнаруживает, что никого нет. Ларина ожидала увидеть с Ленским Онегина, а Ленский приехал один (4'); Елена считает, что братья должны бы уже вернуться, а их все нет (10') и т. д.

Итак, говорящий считает, что в момент речи объект *X* должен находиться в поле его зрения *Q'*, — вот условие, необходимое для того, чтобы вопрос *Где X?* понимался как *Почему X-а нет в Q'?*

Разумеется, есть ситуации, в которых у говорящего (т. е. у задающего вопрос *Где X?*) такой презумпции нет. В этом случае *где* понимается как «местное» *где* и ответ на вопрос есть информация о местонахождении *X-а*:

- (12) ... *Спросил, здесь ли живет чиновник Красинский.*  
— *Пожалуйте в сорок девятый номер, — был ответ.*  
— *А где вход?*  
— *Со двора.* (Лермонтов).

В этом отношении интересны вопросы с местоимениями первого лица — *Где я?* и *Где мы?* Для того чтобы эти вопросы осмыслились как «неместные», необходима презумпция «говорящий ожидает увидеть себя», но она возможна лишь в ситуации с изображением говорящего, которое он может ожидать увидеть (на картине, фотографии и т. д.). Во всех других случаях *где*-вопрос к местоимению первого лица — это вопрос о месте нахождения, ср.:

- (13) — *Но все-таки где мы?*  
— *Мы с вами перенесли в волшебную страну.* (Шварц).  
(14) — *Где я? Какой это город?*  
— *Ну, Ялта.* (Булгаков)

С другой стороны, можно найти и такие контексты употребления *где*, в которых преимуществом пользуется «неместный» анализ, — это *где*-вопросы к отглагольным именам, ср.:

- (15) — *А где же поминки?*  
— *Видишь ли, я не успел как следует подготовиться.* (Вампилов).

Более того, в случае, если имя образовано от глаго-

ла состояния, и, следовательно, обозначает ситуацию, которая, как принято говорить, не способна «развертываться» во времени и пространстве, «местная» интерпретация, предполагающая возможность выбора из множества точек пространства, вообще исключена\*, ср.: *Где Ваша доблесть?; Где Ваши знания, молодой человек?* Так, распространенный ответ на последний вопрос — *Я учил* — есть отрицание презумпции спрашивающего «У Вас нет знаний»: «Я учил, следовательно делал все от меня зависящее, чтобы у меня были знания; у меня должны быть знания». Но указанная презумпция может быть только у задающего «причинный», а не «местный» *где*-вопрос, так как «местный» *где*-вопрос содержит как раз обратную презумпцию, т. е. презумпцию наличия объекта в некоторой точке пространства (говорящему неизвестной).

Итак, «местные» и «причинные» употребления *где*-вопросов противопоставлены: имеются вопросы с *где*, на которые возможны только (или практически только) ответы, содержащие информацию о местоположении объекта (таковы, например, *где*-вопросы к местоимениям первого лица), и вопросы с *где*, которые могут пониматься только как вопросы о причине отсутствия объекта в данном месте (таковы вопросы к отглагольным именам состояния).

Однако в большинстве своем *где*-вопросы, изъятые из контекста, допускают и ту, и другую интерпретацию. Неоднозначность иногда сохраняется даже в том случае, если мы имеем не один вопрос, а целую вопросо-ответную пару. Так, ответ в примере:

- (16) — *А где, если не секрет, жена ваша?*  
— *Она сейчас, видите ли, в Сухуми. Отдыхает.* (Вампилов)

может быть проинтерпретирован и как ответ на местный вопрос (говорящего интересует, где находится жена его собеседника), и как ответ на причинный вопрос (говорящего интересует, почему жена не пришла вместе с мужем). Здесь сообщение о действительном месте пребывания есть уже достаточное объяснение причины отсутствия: «е нет потому, что она на курорте».

Однако встречаются примеры, показывающие, что носитель языка может осознавать эти интерпретации как два разных вопроса, ср., например

- (17) — *Где Ирина?*  
— *Ирина на кухне, кофе готовит.*  
— *А Вероника где?*  
— *Сейчас придет.*  
— *Где она?*  
— *Дома.* (Розов)

Спрашивающий вынужден был переспросить именно потому, что адресат ответил не на тот вопрос, который имелся в виду. Но при этом второй вопрос — тот, который должен объяснить слушающему, что его реакция была неверной, что от него требуется другая информация, — практически совпадает с первым! Поэтому для того чтобы отвечающий ответил на второй вопрос не просто иначе, а так, чтобы ответ удовлетворил спрашивающего, он должен осознать, что отвечает на два разных вопроса. Первый — о том, почему в комнате нет Вероники. «Вопрос не по существу, — отвечает адре-

\* Ср. [6]: «Высказывания, содержащие предикаты рассматриваемого типа, могут быть истинными (или ложными) в течение определенного периода, но они не описывают процессы, протекающие во времени. Лаура могла сказать дону Карлосу: *Мне двух любить нельзя. Теперь люблю тебя* (Пушкин). — но это не значит, что она говорила об актуальном «процессе любви», который развертывается во времени, включающем в себя момент речи».

То же можно сказать и о возможности «развертывания» состояний в пространстве, — об этом свидетельствует, в частности, запрет на конструкции типа:

\* *Он ненавидит в саду; знает в транспорте* и т. д.

сат: — Вероника сейчас придет, я сию же минуту презумпция, дающая право задавать такой вопрос, перестанет иметь место». Второй вопрос — омонимичный первому — это вопрос о месте, в котором в данный момент находится Вероника. Именно так его и понимает отвечающий:

— Где она?

— Дома.

Замечание. Назвав эти два вопроса омонимичными, мы, конечно, несколько огрубели положение. В действительности русский язык располагает средством, позволяющим противопоставить разные интерпретации одного где-вопроса — это смена фразового ударения в (17):

(17а) — А Вероника где?

— Сейчас придет.

— Где она?

— Дома.

Значительно хуже:

(17б) — А Вероника где?

— Сейчас придет.

— Где она?

— Дома

И совсем, на наш взгляд, неприемлемо:

(17в) — А Вероника где?

— Где она?

Фразовое ударение на имени, входящем в сферу действия вопросительного слова, характеризует причинные где-вопросы, фразовое ударение на самом вопросительном слове — местные. Где я? — вопрос человека, потерявшего ориентацию; вопросы Где я? в таком случае, видимо, невозможен. Это, скорее, вопрос человека, который тщетно пытается отыскать на фотографии свое изображение среди других.

Заметим, что обратный порядок вопросов в примере (17) невозможен в силу того общего правила, что где-вопрос о месте нахождения не может предшествовать где-вопросу о причине отсутствия, ср.:

\* (17г) — А Вероника где?

— Дома.

— Где она?

— Сейчас придет.

Для того чтобы иметь возможность в данном контексте задать причинный вопрос, говорящему необходимо иначе его сформулировать; ср. вполне допустимый диалог (17д):

(17д) — <Так вы нашли Веронику?> Где она?

— Дома.

— А почему не здесь?

— Сейчас придет.

Запрет на (17г) объясняется тем, что «неместный» где-вопрос не сводится только к требованию указать причину отсутствия объекта в наблюдаемом говорящим пространстве, — этот вопрос имеет свою специфику употребления, которая обусловлена презумптивной частью его семантического представления. В тот момент, когда говорящий в (17) задает второй где-вопрос, уже невозможен эффект «обманутого ожидания»: говорящий уже знает (ему только что сообщили), что Вероники нет. Между тем именно ситуация «обманутого ожидания» создает условия, необходимые для употребления причинного где-вопроса: говорящий должен быть уверен, что Вероника в комнате, и вдруг обнаружить, что ее нет.

Из сказанного следует, что причинные где-вопросы и вопросы с почему далеко не всегда взаимозаменяемы. Но различие в их значении не затрагивает того аспекта коммуникативного акта, который принято понимать под вопросительностью: этим вопросам сопоставляется одно и то же множество ответов и, следовательно,

одна и та же структура вида 'Сделай так, чтобы я знал P'. Таким образом, для адекватного описания этих вопросов, учитывающего и разницу в их употреблении, требуется некоторая дополнительная (относительно формулы 'Сделай так, чтобы я знал P') информация — она касается условий, побудивших говорящего задать вопрос. Вообще говоря, такие «побудительные» условия являются естественной предпосылкой любого вопроса. Однако в некоторых случаях эта информация составляет настолько существенную характеристику диалога, что даже специальным образом маркируется в языке. Причинное где здесь не исключение. Так, специально маркируются, например, «побудительные» условия, обратные тем, которые характеризуют причинное где: если говорящий не ожидал увидеть объект X (т. е. если у него имелась презумпция, что объект X не должен быть в поле его зрения), в качестве вопроса о причине появления X-а обычно употребляется вопрос со словом откуда. Так, встречая знакомого, которого мы не ожидали увидеть, мы спрашиваем: «Откуда ты?» Семантическая структура этого вопроса примерно следующая: 'Я полагал, что ты не должен находиться здесь; сделай так, чтобы я знал причину того, что ты здесь находишься'. Ср. другие характерные вопросо-ответные пары:

(18) — Откуда гроши, рассказывай.

— Украл. (Вампилов)

(19) — Откуда у вас машина?

— Подарок судьбы... (Арбузов)

Собственно «местные» ответы в данных контекстах неприемлемы (так же как и в случае причинного где). Ср. пример (20), в котором восприятие ответа как иронически-уклончивого возникает, в частности в силу того, что «местные» ответы запрещены языковыми правилами:

(20) — Откуда у тебя такие мысли?

— Из учебника политэкономии. (Вампилов)

Показательны в этом отношении диалоги с вопросом типа Откуда ты знаешь? Это всегда вопрос удивления: спрашивающий не ожидал, что адресат знает нечто, его вопрос — реакция на этот неожиданный для него факт — направлен, так сказать, на выявление «причины знания»:

(21) — Откуда советник знает, что я здесь?

— Он следил за тобой. (Шварц)

(22) — А вы знаете, что такое обворожительная женщина?

— Конечно.

— Откуда?

— Это во мне сидит. (Арбузов)

Многочисленны и примеры использования вопросов со словом куда как высказываний, представляющих собой реакцию на обнаружившееся несоответствие презумпции говорящего и действительного положения вещей. В этом случае у говорящего имеется презумпция, что объект должен в течение некоторого времени находиться в поле его зрения, и когда оказывается, что в действительности это не так, говорящий пытается выяснить причину несоответствия:

(23) — Ты куда?... Вы куда? — крикнул он на трех солдат, которые, без ружей, подобрав полы шинелей, проскользнули мимо него в ряды. — Стой, каналы. (Л. Толстой).

В (23) говорящий спрашивает солдат не о том, куда они бегут (это он, скорее всего, и сам знает), а о том, почему они бегут. Ср. также:

(24) — Куда ты, сынок?

— Нам пора. (Вампилов)

(Ответ вполне удовлетворительный, если считать, что вопрос задан о причине поспешного ухода героя).

Итак, выделяется целая большая область употребления частных вопросов с «нестандартной» семанти-

кой, т. е. такой, которая предписывает говорящему давать ответы, не подчиняющиеся правилам субституции. Подробно рассмотрен только один класс «нестандартных» вопросов. Как нам кажется, уже он дает представление о степени сложности информации, необходимой для построения правильной вопросо-ответной пары, и, следовательно, о неприемлемости операции субституции (ведь она учитывает лишь синтаксическую структуру вопроса) в качестве универсального «механического» средства порождения ответа на вопрос. Уместно еще раз вспомнить А. М. Пешковского: «Что механичность вредна при творческом наблюдении над языком и при истолковании индивидуальной стороны текста — ясно из самих понятий творчества и индивидуальности» [2].

#### ЛИТЕРАТУРА

1. Петерсон М. Н. О вопросах. — Русский язык в школе, 1940, № 2, с. 38—40.
2. Пешковский А. М. Вопрос о вопросах. — В кн.:

Пешковский А. М. Избранные труды. М.: Учпедгиз, 1959, с. 33—50.

3. Ильин Г. М., Новицкая И. М., Смирнова Л. Н. О вопросительных элементах в системе «запрос — ответ». — В кн.: Лингвистические проблемы функционального моделирования речевой деятельности. Вып. 3. Л., 1976, с. 65—74.
4. Крейдлин Г. Е., Рахилина Е. В. Семантический анализ вопросо-ответных структур со словом *какой*. — Изв. АН СССР. Сер. литературы и языка, 1984, 43, № 5, с. 457—471.
5. Падучева Е. В. Вопросительные местоимения и семантика вопроса. — В кн.: Разработка формальной модели естественного языка. Новосибирск, 1981, с. 80—105.
6. Семантические типы предикатов. — М.: Наука, 1982. — 365 с.

Статья поступила в редакцию 4 мая 1985 г.

УДК 801.56:800.92

## АВТОМАТИЧЕСКАЯ ПРОВЕРКА ПРАВИЛЬНОСТИ ДЕФИСНЫХ ОБРАЗОВАНИЙ

И. А. Большаков

### ВВЕДЕНИЕ

В деловой прозе довольно много двух- и даже трех-четырёхсоставных буквенных цепочек, содержащих в качестве соединителей дефисы. Разнообразные примеры таких цепочек: *научно-технический, англо-русско-немецко-французский, киловатт-час, гамма-функция, мини-ЭВМ, 30-летие, Жатва-84, 5-й, п-ка, S-образный, Сиси-спираль, 8-мм* (сокращение для *восьмимиллиметровый* во всех родах и падежах), *кое-как, какой-нибудь, все-таки, Жан-Поль, св-во* (сокращение для *свойство*), *ин-т* (сокращение для *институт*).

Назовем все эти случаи дефисными образованиями, а их составляющие, не содержащие дефиса, — парциальными цепочками.

При обработке текстов на машинных носителях, в частности, в системах автоматического обнаружения и исправления ошибок [1—3] необходимо решать, как обрабатывать такие объекты. Возможен простейший подход, когда дефисные образования считаются неделимыми. Но такое решение противоречит более экономичному атомарному подходу, ориентирующемуся на максимальное членение языковых сущностей и возобладавшему при описании иных уровней языка — морфологии, синтаксиса и даже семантики.

Следуя атомарному подходу, отнесем дефис к разделителям, но особого рода. Будем анализировать и, где нужно, хранить в словаре каждую парциальную цепочку отдельно, причем при хранении снабдим ее признаками дефисной сочетаемости справа/слева с парциальными цепочками разных классов.

Для таких дефисных образований, как *X-летие, У-го, Жатва-Z*, где *X, Y, Z* — числовые или буквенные переменные, атомарный подход единственно возможен. Для образований типа *киловатт-час* или *гамма-функция* расчленение удобно и экономично\*. В случае образований типа *научно-технический, русско-английский, кое-как, какой-нибудь, по-можковски* расчленение порой порождает цепочки, в качестве независимых словоформ не употребляющиеся, но формально такой прием подходит и здесь.

Хуже всего обстоит дело с сокращениями типа *св-ва, ин-т*. Здесь словоформы разрываются на несуразные, как правило, не существующие отдельно части. Но дефисные сокращения употребимы только в ряде жанров письменной речи (рефераты, энциклопедические издания и т. п.), да и здесь атомарный принцип формально пригоден.

В данной работе анализируются все известные нам классы парциальных цепочек и признаки, однозначно описывающие их взаимную сочетаемость в рамках дефисных образований. Тем самым строятся формальные средства их описания, справедливые для широкого спектра естественных языков, пользующихся кириллическим или латинским алфавитами.

\* К этой же группе можно условно отнести соединенные с помощью тире, но без пробелов по обе стороны от него, пары типа *два-три, культиватор-рыхлитель*. Дело в том, что на конторских машинках и в алфавитах ЭВМ существует лишь единый знак дефиса/тире. Рекомендации ГОСТ 17059-71 изображать тире двумя такими знаками подряд редко соблюдаются [4]. И только тогда, когда этот единый знак окаймлен пробелами, он четко понимается как тире, и к этим случаям наш анализ отношения не имеет.