

А. Ф. Верлань, И. А. Чмырь, Д. А. Верлань

ДИАЛОГОВЫЕ ПРОЦЕССЫ И КОГНИТИВНЫЕ ТЬЮТОРЫ

УДК 004.8:519.87

Рецензенты:

В. Е. Быков, академик Национальной академии педагогических наук Украины, д. т. н., проф., директор Института цифровизации образования Национальной академии педагогических наук Украины;

А. Е. Коваленко, д. т. н., ведущий научный сотрудник Института проблем математических машин и систем НАН Украины

Научно-издательский отдел физико-математической
и технической литературы
Редактор **А. Я. Бельдий**

Книга посвящена формальному моделированию вопросно-ответных диалоговых процессов и использованию этих моделей для синтеза искусственных диалоговых агентов. В первой части книги предлагается онтологическая модель естественных диалоговых процессов. На основе понятий, составляющих онтологическую модель естественного диалога, строятся несколько формальных моделей вопросно-ответного диалога, различающихся сложностью и сферой применимости. Сетевая диалогическая модель является теоретической основой для синтеза искусственного диалогового агента. Во второй части книги рассматривается применимость искусственного диалогового агента для решения плохо формализуемых проблем. Внимание фокусируется на решении проблем лингводидактики. При решении этих проблем искусственный диалоговый агент рассматривается как когнитивный тьютор, реализующий лингводидактическую методику опытного преподавателя.

Для специалистов и аспирантов в области искусственного интеллекта, когнитивных наук и компьютерного обучения.

ISBN 978-966-00-1848-8

© А. Ф. Верлань, И. А. Чмырь,
Д. А. Верлань, 2022

СОДЕРЖАНИЕ

ПРЕДИСЛОВИЕ	5
1. ЕСТЕСТВЕННЫЙ ДИАЛОГ	13
1.1. Онтологический базис естественного диалога	17
1.1.1. Диалоговый агент и его роли. Внешний и внутренний диалоги	18
1.1.2. Диалоговая транзакция	20
1.1.3. Типы посылок активного агента	24
1.1.4. Типы посылок реактивного агента	31
1.1.5. Истинность информационной посылки реактивного агента. . .	32
1.1.6. Диалоговое поведение	34
1.1.7. Знания, переносимые вопросно-ответной транзакцией	35
1.2. Логическая структура вопросно-ответной транзакции.	37
1.2.1. Онтологический базис эротетической логики	38
1.2.2. Интеррогативная формула ли- и какой-вопросов.	46
1.2.3. Спецификация размера выборки	53
1.2.4. Спецификация требования полноты	57
1.2.5. Спецификация требования различения.	62
1.2.6. Дескрипторные вопросы	65
2. ВОПРОСНО-ОТВЕТНЫЕ ТРАНЗАКЦИИ	69
2.1. Вопросно-ответные транзакции в контексте представления знаний. . .	71
2.1.1. Язык Тернарного Описания Уёмова	74
2.1.2. Представление декларативных знаний в вопросно-ответной транзакции	78
2.2. Эпистемологический взгляд на знания в вопросно-ответной транзакции	86
2.2.1. Конвергентные знания	91
2.2.2. Представление знаний-wh, включающее вопрос	96
3. ПРОБЛЕМНО-НЕЗАВИСИМЫЕ МОДЕЛИ ДИАЛОГА	99
3.1. Когнитивные основы диалогового процесса	100
3.2. Конечно-автоматная модель вопросно-ответного диалога.	108
3.2.1. Проблема моделирования неискренности реактивного агента	115

3.3. Модель вопросно-ответного диалога в виде сети Петри	120
3.3.1. Применимость модели вопросно-ответного диалога в виде сети Петри	129
3.4. Сетевая даталогическая модель вопросно-ответного диалога	135
3.4.1. Отображение сетевой модели в реляционную базу данных.	143
3.4.2. Даталогическое представление вопроса	152
3.4.3. Искусственный активный агент диалога.	167
4. КОГНИТИВНЫЕ ТЬЮТОРЫ	171
4.1. Системы коммуникации знаний	172
4.1.1. Интеллектуальные тьюторы.	173
4.1.2. Обучающие среды с интеллектуальной поддержкой.	182
4.1.3. Интеллектуальные репетиторы	186
4.1.4. Обобщённая структура системы коммуникации знаний.	191
4.2. Теория обучения	194
4.2.1. Философия обучения	195
4.2.2. Сущность обучения.	198
4.2.3. Отношение между учеником и обучающим материалом	199
4.2.4. Обучение и деятельность	200
4.3. Классические методы лингводидактики	202
4.3.1. Ситуативный метод обучения	202
4.3.2. Аудио-лингвистический метод обучения	205
4.3.3. Метод общего физического отклика	209
4.4. Сцены лингводидактических тьюторов	212
4.4.1. Диагностические сцены	216
4.4.2. Обучающие стимулы и сцены	225
4.5. Диагностика и логический вывод	240
4.5.1. Диалог и логический вывод в экспертных системах.	241
4.5.2. Дедуктивные связи в сети диалогового метода	251
4.5.3. Логический вывод в сети диалогового метода	259
4.5.4. Вопросно-ответная транзакция в контексте логического вывода.	265
СПИСОК ЛИТЕРАТУРЫ	273
ТЕЗАУРУС	283
1. Естественный диалог.	283
2. Вопросно-ответные транзакции.	290
3. Проблемно-независимые модели диалога.	291
4. Когнитивные тьюторы.	294

ПРЕДИСЛОВИЕ

В современной классификации диалоговых систем и искусственных «разговаривающих» агентов превалирует дихотомия. Системы делятся на два класса: (1) чатботы и (2) диалоговые системы, ориентированные на решение проблем.

Чатботы в состоянии поддерживать неструктурированный вербальный вопросно-ответный диалог с фиксированным распределением ролей. Как правило, чатбот выступает в роли реактивного (отвечающего на вопросы) партнёра/агента разговора. Протокол диалога чатбота с партнёром нельзя назвать целенаправленным, а основным принципом формирования ответа чатбота на вопрос партнёра является трансформация полученного вопроса. Публикации, посвященные чатботам, имеют выраженное прагматическое измерение и часто описывают полезность и целесообразность использования чатботов в системах электронной коммерции и социальных сетях.

Диалоговые системы, ориентированные на решение проблем, позволяют решать некоторые проблемы, относящиеся к классу плохо формализуемых проблем. В таких системах, как правило, человек выступает в роли активного (задающего вопросы) агента, однако, допускается смена ролей в ходе диалогового процесса. Протокол диалога такой системы с партнёром является целенаправленным, поскольку сам диалог направлен

на решения проблемы и, следовательно, на достижение некоторой цели. Для формирования ответов на вопросы используются различные сложные методы и сценарии. Примером плохо формализуемой проблемы, которая может решаться в ходе диалогового процесса, является проблема обучения или проблема передачи знаний.

Книга посвящена, главным образом, вопросам моделирования и синтеза искусственных диалоговых агентов, ориентированных на решение плохо формализуемых проблем вообще и в области обучения в частности. В качестве предметной области обучения рассматривается компьютерная лингводидактика, которая занимается моделированием и разработкой компьютерных программ обучения иностранным языкам.

Первый раздел посвящён онтологическому базису естественного диалога и логической структуре вопросно-ответной транзакции в естественном диалоге.

Под естественным диалогом понимается свободный диалог или разговор, в котором участвуют люди, а не искусственные системы. Предлагаемые термины и понятия, характеризующие естественный диалог, сформулированы и определены на основе анализа протоколов диалогов, описанных в диалогической литературе. В частности, для иллюстраций терминов и понятий, характеризующих естественный диалог, в первом разделе часто используется фрагмент одного из диалогов Платона, названного «Протагор». Понятия, введенные в первом разделе, такие, например, как «активный диалоговый агент», «реактивный диалоговый агент», «диалоговая транзакция», «информационная посылка активного агента», «информационная посылка реактивного», используются во всех последующих разделах книги. Смысловое наполнение этих понятий последовательно уточняется и обобщается. Так, в первом разделе такое понятие как «информационная посылка активного агента диалога» понимается как вопрос в лингвистическом смысле этого слова. В дальнейшем под информационной посылкой активного агента понимается любой стимул, имеющий интеррогативный статус,

который может быть представлен как вербально, так и не вербально.

Информационная посылка активного агента диалога и связанная с ней информационная посылка реактивного агента составляют диалоговую транзакцию. Диалоговую транзакцию можно рассматривать как «молекулу» диалогового процесса, которая имеет внутреннюю логическую структуру. Логическую структуру диалоговой транзакции для одного из типов естественного диалога — вопросно-ответного диалога изучает эротетическая логика. В первом разделе книги приводится критический обзор основных положений классической эротетической логики. Один из важных выводов, который следует из этой теории, заключается в том, что в вопросно-ответном диалоге информационная посылка активного агента явно или неявно содержит информационную посылку реактивного агента. Иными словами, вопрос явно или неявно содержит ответ. Логическая структура вопроса в эротетической логике называется интеррогативной формулой и конструируется исходя именно из этого положения. Ключевая часть интеррогативной формулы, называемая субъектом вопроса, для целого ряда вопросов содержит ответ на вопрос, и поэтому может рассматриваться как ответ, представленный со значительной степенью неопределённости. Интеррогативная формула, рассмотренная в первом разделе, используется в последующих разделах книги.

Во втором разделе развивается теория логической структуры вопросно-ответной транзакции. Вопросно-ответная транзакция рассматривается как транспортное средство для передачи декларативных знаний от реактивного агента к активному агенту. Предлагается ряд моделей для представления декларативных знаний, ассоциированных с вопросом и ответом. В этих моделях декларативные знания, ассоциированные с вопросно-ответной транзакцией, представляются не в виде пропозиций, а с помощью лингво-независимых категорий «объект-вещь», «объект-свойство», «объект-отношение», а также

категорий «определённый объект», «неопределённый объект», «произвольный объект».

Во второй части второго раздела знания, которыми обладает реактивный агент, рассматриваются с эпистемологической точки зрения как знания-that и знания-wh. Показано, что для решения проблемы конвергентных вопросов (вопросов, имеющих различные субъекты, но предполагающие один и тот же истинный ответ) формула, кодирующая знания реактивного агента, выражаемые пропозицией p , должна включать вопрос, для которого p является истинным ответом.

Третий раздел посвящён синтезу проблемно-независимых моделей вопросно-ответного диалога. В разделе описаны три формальные модели: (1) конечно-автоматная модель; (2) модель на основе сети Петри; (3) сетевая даталогическая модель. Основа всех трёх моделей — когнитивная модель диалогового цикла. Когнитивная модель диалогового цикла является развитием перцептивного цикла Найссера и описывает диалоговый цикл с точки зрения когнитивных структур и процессов, участвующих в вопросно-ответной транзакции. Компоненты и онтологический базис когнитивной модели диалогового цикла используются во всех трёх формальных моделях и определяют независимость этих моделей от предметной области диалога. Из когнитивной модели диалогового цикла следуют два важных заключения: (1) на каждом диалоговом цикле активный агент оперирует относительно небольшим множеством ответов, которые он ожидает получить от реактивного агента; (2) очередной вопрос активного агента формируется путём поиска в памяти, названной памятью вопросов, а метод доступа к этой памяти является диалоговым методом решения проблемы.

Конечно-автоматная модель ассоциирует диалоговый метод решения проблемы с функционированием автомата Мили. В основу формализмов этой модели положены следующие основные предположения. К моменту начала диалогового процесса определены конечные множества вопросов активного агента и ответов реактивного агента, а на каждом диалоговом цикле множество ответов, ожидаемых активным агентом, под-

вергается модификации. Модель может описывать диалоговый процесс аналитически, в табличной форме или графически.

Модель на основе сети Петри является дальнейшим развитием конечно-автоматной модели. Главное преимущество модели на основе сети Петри — легкость её реализации на основе даталогического подхода к проектированию программ диалоговых агентов. Даталогический подход означает, что большая часть системы представляется не программным кодом, а отображается в базу данных. Решающим преимуществом такой системы является возможность конструирования и редактирования диалогового метода его автором с помощью полноэкранный редактора базы данных. Модель на основе сети Петри более универсальна и функциональна, чем конечно-автоматная модель. Она включает память ответов реактивного агента, а также компоненты, с помощью которых осуществляются стандартное и нестандартное управления вопросом-ответным циклом.

Сетевая даталогическая модель построена на предположении, что диалоговый метод имеет нерегулярную сетевую структуру. Сеть состоит из разнотипных узловых элементов, а отдельный узел сети моделирует один потенциальный вопросно-ответный цикл. Сеть отображается в базу данных. Приведен пример отображения сети диалогового метода в реляционную базу данных. В завершающей части описана архитектура искусственного активного агента вопросно-ответного диалога. Архитектура объединяет концепции всех трёх формальных моделей.

Четвёртый раздел посвящён когнитивным тьюторам вообще и лингводидактическим тьюторам в частности. Идеальный когнитивный тьютор определяется с помощью следующих характеристик: (1) непрерывно адаптируется к текущим знаниям ученика и степени понимания им учебного материала; (2) выбирает учебный материал, для которого имеет место несоответствие между текущими и требуемыми знаниями и/или степени понимания ученика; (3) автоматически или по запросу ученика дополняет учебный материал тематически связанной

информацией (например, информацией справочного характера); (4) тестирует ученика и определяет его когнитивные характеристики и когнитивный профиль; (5) проводит процесс обучения методом, соответствующим текущим когнитивным способностям и когнитивному профилю ученика.

Системы коммуникации знаний имеют отношение к когнитивным тьюторам и некоторые характеристики систем коммуникации знаний приближаются к характеристикам идеального когнитивного тьютора. В начальной части раздела приведен обзор ранних, ныне классических, систем коммуникации знаний. Системы коммуникации знаний распределены между тремя классами: (1) интеллектуальные тьюторы; (2) обучающие среды с интеллектуальной поддержкой; (3) интеллектуальные репетиторы.

Несмотря на то, что метод обучения является одним из базовых концептов систем коммуникации знаний, описанные в литературе системы коммуникации знаний, как правило, не опираются на результаты, полученные в теории обучающих процессов. С точки зрения авторов книги, теория обучения является методологической основой проектирования методов обучения и должна учитываться при их проектировании. Кратко рассмотрены основные компоненты теории обучения применительно к лингводидактике: (1) философия обучения; (2) сущность обучения; (3) отношение между учеником и обучающим материалом; (4) обучение и деятельность.

Компьютерная лингводидактика занимается моделированием и разработкой компьютерных программ обучения иностранным языкам и представляет собой предметную область, удобную для развития идей диалоговых когнитивных тьюторов и их проверки путём проектирования реальных систем. Обучение иностранным языкам является относительно стабильной предметной областью и не подвержено резким методологическим изменениям. Здесь применимы практически все идеи как систем коммуникации знаний, так и теории обучения. Кроме того, привлекательной является попытка встроить известные, но «некомпьютерные» методы обучения иностранным языкам,

в архитектуру системы коммуникации знаний. В четвёртом разделе приведен обзор классических методов обучения иностранным языкам: (1) ситуативный метод; (2) аудио-лингвистический метод; (3) метод общего физического отклика.

Идеальный когнитивный тьютор в процессе диалогового взаимодействия с учеником адаптируется к текущему уровню знаний ученика и степени понимания учебного материала, а также осуществляет процесс обучения методом, соответствующим текущим когнитивным способностям и стилю ученика. Поэтому диалоговая сессия когнитивного тьютора и ученика должна включать два типа эпизодов: (1) эпизоды, осуществляющие тестирование ученика с целью диагностики текущего уровня знаний и определения его когнитивных способностей и когнитивного стиля; (2) эпизоды, осуществляющие непосредственно обучение. Сцены, используемые для проектирования обучающих эпизодов, названы обучающими сценами, а сцены, используемые для проектирования эпизодов, осуществляющих тестирование — диагностическими сценами. Приведена классификация и описание ряда диагностических и обучающих сцен.

В завершающей части развивается идея совмещения процесса обучения и процесса диагностики когнитивного статуса ученика в едином диалоговом процессе. Как правило, анализ наборов данных, полученных в результате применения психодиагностических методик, построенных с использованием сложных диагностических сцен, осуществляется непосредственно экспертом психодиагностом. Есть все основания предполагать, что такой анализ представляет собой логический вывод, использующий процедурные знания эксперта и что психодиагност приходит к заключению о когнитивных характеристиках ученика в результате дедуктивных умозаключений. Сетевые модели диалогового метода доступа, описанные в третьем разделе, в контексте логического вывода могут рассматриваться как сети вывода, а вопросно-ответный диалоговый процесс может использоваться для моделирования дедуктивных умозаключений. Таким образом, для реализации психоди-

агностических методик, построенных на основе сложных диагностических сцен, может использоваться вопросно-ответный диалоговый процесс. Это открывает перспективы для конструирования лингводидактических сценариев, совмещающих диагностику и обучение. Основной гипотезой, лежащей в основе этих исследований, является предположение о том, что вопросно-ответный диалоговый процесс обладает свойствами процесса логического вывода на основе дедуктивных умозаключений. В результате исследований получена формальная модель дедуктивных связей в сети диалогового метода доступа.

В книге отношения между сущностями часто иллюстрируются графически простыми диаграммами классов унифицированного языка моделирования UML. Эти диаграммы интуитивно понятны, однако для чтения книги желательно ознакомиться с основами моделирования пространственных структур систем с помощью диаграмм классов унифицированного языка моделирования UML.

Наиболее важные слова, словосочетания и фразы выделены курсивом. На них рекомендуется фокусировать внимание при чтении книги. Большая часть этих слов и словосочетаний включены в тезаурус, размещенный в конце книги.

1. ЕСТЕСТВЕННЫЙ ДИАЛОГ

Естественный диалог — это не регламентированный, свободный диалог или разговор, в котором участвуют люди, а не искусственные системы. Естественный диалоговый процесс противопоставляем искусственному диалогу, который, может использоваться в качестве интерфейсного средства для взаимодействия конечного пользователя с компьютерной программой, а также как средство для решения некоторых плохо формализуемых проблем. Искусственный диалог проще естественного и часто регламентирован набором правил и ограничений, которые должен соблюдать конечный пользователь. Примером искусственного диалога может служить протокол взаимодействия человека с компьютерной системой резервирования авиабилетов.

Построение онтологического базиса естественного диалога и его категоризация предполагают анализ некоторого количества протоколов естественных диалоговых процессов, отражающих их логическое и операционное многообразие. Существует множество источников таких протоколов. Например, протоколы допросов подозреваемых, когда следователь пытается определить вину подозреваемого, а последний — убедить следователя в том, что он невиновен. Однако, наиболее доступным и богатым источником протоколов естествен-

ных диалогов является диалогическая литература, особенно литература философского либо теологического направления. Одними из наиболее известных и значимых источников протоколов естественных диалогов являются диалогические произведения Платона [1]. Философская и теологическая литература изобилует диалогическими произведениями. В качестве примера можно также отметить менее известные, чем Платоновские диалоги, диалогические произведения Григория Сковороды [2].

В этом разделе понятия, формирующие онтологический базис естественного диалога и его категоризацию, будем иллюстрировать на примере небольшого фрагмента одного из диалогов Платона, названного «Протагор». Фрагмент диалога «Протагор», приведенный ниже, представлен в форме, немного отличающейся от оригинала. Незначительные изменения и добавления сделаны с целью его структуризации, а также для облегчения ссылок на элементы протокола. Например, отмечены начало и конец каждой диалоговой транзакции, а также явно отмечены субъекты диалогового общения.

Диалог происходит между Сократом (С) и Гиппократом (Г) перед их встречей с Протагором [3]. Термины «транзакция», «активный агент» и «реактивный агент» определены в дальнейшем изложении. Читателю настоятельно рекомендуется ознакомиться с этим фрагментом, прежде чем далее читать текст.

Начало транзакции 1. С — активный агент.

С: Так мы идём платить деньги Протагору, потому что он софист?

Г: Конечно.

Конец транзакции 1. Начало транзакции 2.

С: А если бы спросили тебя ещё и вот о чём:
«Сам-то ты кем намерен стать, раз идёшь к Протагору?»

Г: Если сообразоваться с прежде сказанным, то ясно, что я собираюсь стать софистом.

Конец транзакции 2. Начало транзакции 3.

С: А тебе не стыдно было бы, клянусь богами, появиться среди эллинов, в виде софиста?

Г: Клянусь Зевом, стыдно, Сократ, если говорить то, что я думаю.

Конец транзакции 3. Начало транзакции 4.

С: Но пожалуй, Гиппократ, ты полагаешь, что у Протагора тебе придётся учиться иначе, подобно тому, как учился ты у учителя грамоты, игры на кифаре или гимнастики? Ведь каждому из этих предметов ты учился не как будущему своему мастерству, а лишь ради своего образования, как это подобает частному лицу и свободному человеку.

Г: Конечно, мне кажется, что Протагорово обучение скорее такого рода.

Конец транзакции 4. Начало транзакции 5.

С: Так сам-то ты знаешь, что собираешься делать, или тебе это неясно?

Г: О чём это ты?

Конец транзакция 5. Начало транзакции 6. Обмен ролями. Далее С — реактивный агент.

С: Ты намерен предоставить попечение о твоей душе софисту, как ты говоришь; но, право, я бы очень удивился, если бы ты знал, что такое софист. А раз тебе это неизвестно, то ты не знаешь и того, кому ты вверяешь свою душу и для чего — для хорошего или дурного.

Г: Я думаю, что знаю.

Конец транзакции 6. Начало транзакции 7. Обмен ролями. Далее С — активный агент.

С: Так скажи, что такое софист по-твоему?

Г: Я полагаю, что по смыслу этого слова он — знаток в мудрых вещах.

Конец транзакции 7. Начало транзакции 8.

С: Да ведь это можно сказать и про живописцев, и про строителей: они тоже знатоки в мудрых вещах; но если бы кто-нибудь спросил у нас, в каких именно мудрых вещах знатоки живописцы, мы бы сказали, что в создании изображений; и в других случаях ответили бы так же. А вот если бы кто спросил, чем мудр софист, что бы мы ответили? В каком деле он наставник?

Г: А что если бы мы так определили его, Сократ: это тот, кто наставляет других в искусстве красноречия?

Конец транзакции 8. Начало транзакции 9. Обмен ролями. Далее С — реактивный агент.

С: Может быть мы и верно бы сказали, однако недостаточно, потому что этот наш ответ требует дальнейшего вопроса: если софист делает людей искусными в речах, то о чём эти речи? Кифарист, например, делает человека искусным в суждениях о том, чему он его научил, — то есть об игре на кифаре. Не так ли?

Г: Да.

Конец транзакции 9. Начало транзакции 10. Обмен ролями. Далее С — активный агент.

С: Допустим. Ну а софист, в каких речах он делает искусным? Не ясно ли, что в речах о том, в чём он и сам сведущ?

Г: Похоже на то.

Конец транзакции 10. Начало транзакции 11.

С: А в чём же софист и сам сведущ и ученика делает сведущим?

Г: Клянусь Зевсом, не знаю, что тебе ответить.

Конец транзакции 11. Начало транзакции 12.

С: Как же так? Знаешь, какой опасности ты собираешься подвергнуть свою душу? Ведь когда тебе бывало нужно верить кому-нибудь своё тело и было неизвестно пойдёт ли это на пользу или

во вред, ты и сам немало раздумывал, верить его или не верить, и друзей, и домашних призывал на совет и обсуждал это дело целыми днями. А когда речь зашла о душе, которую ты ведь ставишь выше, чем тело, потому что от того, будет она лучше или хуже зависит хорошо или дурно пойдут все твои дела, ты ни с отцом, ни с братом и ни с кем из нас, твоих друзей, не советовался верить ли тебе или не верить свою душу этому пришлому чужеземцу. Лишь вчера ввечеру, по твоим словам, услышав о нём, ты уже сегодня идёшь спозаранку, не поразмыслив и не посоветовавшись о том, нужно ли верить ему себя или нет, и сразу готов потратить и собственные деньги и деньги друзей, как будто ты уже дознался, что тебе нужно непременно сойтись с Протагором, которого, как ты говоришь, ты не знаешь и не разговаривал с ним никогда. Ты называешь его софистом, а что такое софист, оказывается, совсем не ведаешь, хоть и собираешься верить себя ему.

Г: Так оно и выходит, Сократ, как ты говоришь.
Конец транзакции 12. Начало транзакции 13.

С: А что, Гиппократ, не будет ли наш софист чем-то вроде торговца или разносчика тех припасов, которыми питается душа? По-моему, во всяком случае, он таков.

Г: Но чем же питается душа, Сократ?
Конец транзакции 13. Обмен ролями. Далее С — реактивный агент.

1.1. Онтологический базис естественного диалога

Здесь термин «онтологический базис» используется для обозначения терминологической базы предметной области, названной «естественный диалог», а в последующем подразделе — для обозначения терминологической базы одной из те-

орий, посвященных исследованию логической структуры диалоговой транзакции. Отметим, что онтологический базис для любой предметной области не может быть абсолютным и должен рассматриваться по отношению к цели, которая преследуется при его формировании. Приведенный онтологический базис является общим для всех разделов книги, а его цель — использование единого тезауруса в изложении разных разделов. Предлагаемый онтологический базис не претендует на полноту, является базовым и в дальнейшем будет пополняться новыми терминами. Понятия, формирующие онтологический базис, как правило, выделены курсивом.

1.1.1. Диалоговый агент и его роли. Внешний и внутренний диалоги

Начнём с понятия *диалоговый агент*. Диалоговые агенты формируют *диалоговую транзакцию*, которая представляет собой элементарный завершённый цикл обмена знаниями между диалоговыми агентами. Хотя участников диалога может быть сколь угодно много, диалоговую транзакцию всегда формируют только два диалоговых агента. Например, в приведенном выше диалоге «Протагор» в различные моменты времени транзакции формируют следующие пары агентов: Сократ-Гиппократ, Сократ-Протагор, Сократ-Каллий и др.

Диалоговый агент в процессе диалогового взаимодействия со своим партнёром может исполнять одну из двух альтернативных ролей: роль *активного диалогового агента* либо роль *реактивного диалогового агента*. Активный агент — это спрашивающий агент. Та часть диалоговой транзакции, которую формирует активный агент, по отношению к реактивному агенту имеет статус вопроса. Главная побудительная причина диалогового агента исполнять роль активного агента заключается в нехватке знаний, необходимых для продолжения диалога. Агент исполняет роль активного диалогового

агента в том случае, когда ему необходимы дополнительные знания, которые он предполагает получить от своего партнёра по диалогу.

Реактивный агент — это отвечающий агент и та часть диалоговой транзакции, которую он формирует, имеет статус ответа по отношению к активному агенту. Агент, исполняя роль реактивного агента, предоставляет своему партнёру некоторые знания, релевантные запрашиваемым.

В приведенном фрагменте диалога «Протагор», после того, как Гиппократ заявил, что он знает, что есть софист (транзакция номер шесть), Сократу для продолжения диалога необходимы дополнительные знания, которыми обладает Гиппократ. Сократ хочет выяснить, что именно понимает Гиппократ под словом «софист». Поэтому в следующей транзакции (транзакция номер семь), исполняя роль активного агента, Сократ запрашивает у Гиппократа требуемые знания.

С: Так скажи, что такое софист, по-твоему?

Г: Я полагаю, что по смыслу этого слова он — знаток в мудрых вещах.

В приведенном фрагменте диалога «Протагор» Сократ и Гиппократ обмениваются ролями. Вначале, в транзакциях с первой по пятую, Сократ исполняет роль активного агента, а Гиппократ — роль реактивного. Затем агенты несколько раз обмениваются ролями.

Инициатором смены ролей, как правило, является реактивный агент. Активный агент обычно удовлетворён своей ролью и после смены ролей старается восстановить свою активность при первой возможности.

В том случае, когда диалоговую транзакцию формируют две личности, между которыми осуществляется диалоговое взаимодействие, будем говорить, что имеет место *внешний диалог*. Однако диалоговая транзакция может быть сформирована и одной личностью, попеременно исполняющей роль активного или реактивного агентов. Такой диалог назовём *внутренним*

диалогом. Вербальный внутренний диалог часто называют монологом, а ментальный — размышлением.

В случае внешнего диалога знания, которые запрашивает активный агент, хранятся в долговременной памяти реактивного агента, а знания, на которые ссылается реактивный агент в своём ответе — в долговременной памяти активного агента. Внешний диалог возможен только в том случае, когда активный и реактивный агенты обладают одними и теми же, или примерно одними и теми же знаниями. В случае внутреннего диалога и запрашиваемые, и получаемые знания хранятся в долговременной памяти одного и того же агента.

Большая часть приведенного фрагмента диалога «Протагор» иллюстрирует внешний диалог, однако в нём присутствует и внутренний диалог. В тринадцатой транзакции Сократ ведёт диалог/разговор «сам с собой» и то, что он говорит, можно представить в виде транзакции внутреннего диалога следующим образом:

С (активный) : А что, Гиппократ, не будет ли наш софист чем-то вроде торговца или разносчика тех припасов, которыми питается душа?

С (реактивный) : По-моему, во всяком случае, он таков.

1.1.2. Диалоговая транзакция

Диалоговый процесс может быть записан в виде *протокола диалога*, представляющего собой последовательную цепь *диалоговых транзакций*. Характерной особенностью естественного диалога, которая отличает его от других форм интерактивного взаимодействия, например, от работы человека с прикладной системой базы данных, является наличие сильной логической зависимости не только между компонентами отдельной транзакции, что очевидно, но и между смежными звеньями цепи транзакций. Именно наличие такой логической

зависимости превращает набор диалоговых транзакций в целенаправленный интеллектуальный процесс.

Каждая диалоговая транзакция состоит из двух компонентов, которые будем называть *информационными посылками*. *Информационная посылка активного агента* имеет статус вопроса. Слово «статус» здесь означает, что информационная посылка активного агента не обязательно является вопросом в грамматическом смысле. Это запрос на получение знаний, который может принимать разные формы. Форма информационной посылки активного агента не регламентирована. Она не влияет ни на цель, которую пытается достичь активный агент с помощью отдельной транзакции, ни на цель, которую он пытается достичь с помощью некоторой последовательности диалоговых транзакций. Для активного агента важным является получение доступа к запрашиваемым знаниям, а не форма запроса, с помощью которого осуществляется этот доступ. Очевидно, что в естественном диалоге активный агент обладает свободой выбора формы информационной посылки. Иными словами, доступ к одним тем же знаниям активный агент может получить с помощью различных информационных посылок. В некоторых эпистемологических публикациях эта естественная свобода выбора активным агентом формы доступа к знаниям реактивного агента считается проблемой и называется «проблема конвергентных знаний». Так, Шаффер называет конвергентными различные вопросы, предполагающие одинаковый ответ, а знания, ассоциированные с конвергентными вопросами — конвергентными знаниями [4]. В приведенном фрагменте диалога «Протагор» Сократ, исполняя роль активного агента, может по разному формировать свои информационные посылки для доступа к одним и тем же знаниям Гиппократу. Например, транзакция тринадцатая имеет вид:

С: А что, Гиппократ, не будет ли наш софист чем-то вроде торговца или разносчика тех припасов, которыми питается душа? По-моему, во всяком случае, он таков.

Г: Но чем же питается душа, Сократ?

Понятно, что эта же транзакция могла бы выглядеть следующим образом:

С: Скажи мне Гиппократ, а не похож ли софист на продавца того, чем питается душа?

Г: Но чем же питается душа, Сократ?

Информационные посылки активного агента могут варьироваться от простого «да/нет» вопроса, до продолжительного внутреннего диалога. Однако, даже в том случае, когда информационная посылка активного агента представляет собой некоторые его продолжительные рассуждения, она всё равно имеет статус вопроса по отношению к реактивному агенту. Поэтому уместно говорить о *степени интеррогативности информационной посылки активного агента*.

Часто информационную посылку активного агента можно понимать как *поисковое предписание*, необходимое для получения доступа к знаниям реактивного агента, или как некоторую *ссылку*, обеспечивающую доступ к области долговременной памяти реактивного агента, хранящей запрашиваемые знания.

Информационная посылка реактивного агента имеет статус ответа. Эту информационную посылку можно также понимать как поисковое предписание, необходимое для получения доступа к знаниям активного агента, или как некоторую ссылку на участок долговременной памяти активного агента, где хранятся запрашиваемые знания. В дальнейшем, с целью сокращения длины предложений вместо словосочетания «информационная посылка» будем использовать слово «посылка».

В приведенном фрагменте диалога «Протагор» одной из наиболее простых транзакций, с точки зрения структуры её посылки, является первая транзакция.

С: Так мы идём платить деньги Протагору, потому что он софист?

Г: Конечно.

В этой транзакции активный агент Сократ передаёт реактивному агенту Гиппократу посылку, обеспечивающую Сократу доступ к знаниям, которыми обладает Гиппократ и которые отсутствуют у Сократа. Сократ не знает, намерен ли Гиппократ платить деньги Протагору только на основании того, что Протагор софист. Эта посылка имеет высокую степень интеррогативности и представлена одним вопросительным предложением.

В двенадцатой транзакции диалога «Протагор» посылка Сократа, выступающего в роли активного агента, имеет более сложную структуру. Она представляет собой его рассуждения, которым, однако, присуща некоторая степень интеррогативности.

С: Как же так? Знаешь, какой опасности ты собираешься подвергнуть свою душу? Ведь когда тебе бывало нужно верить кому-нибудь своё тело и было неизвестно, пойдёт ли это на пользу или во вред, ты и сам немало раздумывал, верить его или не верить, и друзей, и домашних призывал на совет и обсуждал это дело целыми днями. А когда речь зашла о душе, которую ты ведь ставишь выше, чем тело, потому что от того, будет она лучше или хуже, зависит, хорошо или дурно пойдут все твои дела, ты ни с отцом, ни с братом и ни с кем из нас, твоих друзей, не советовался, верить ли тебе или не верить свою душу этому пришлому чужеземцу. Лишь вчера вечером, по твоим словам, услышав о нём, ты уже сегодня идёшь спозаранку, не поразмыслив и не посоветовавшись о том, нужно ли верить ему себя или нет, и сразу готов потратить и собственные деньги, и деньги друзей, как будто ты уже дознался, что тебе нужно непременно сойтись с Протагором, которого, как ты говоришь, ты не знаешь и не разговаривал с ним никогда. Ты называешь его софистом, а что такое софист, оказывается, совсем не ведаешь, хоть и собираешься верить себя ему.

Г: Так оно и выходит, Сократ, как ты говоришь.

Сократ хочет узнать, согласен ли Гиппократ с его выводом, заключающимся в том, что нельзя верить душе человеку, с которым ты не знаком и которого называют софистом, не зная, что означает это слово, поскольку душе может быть нанесён вред так же, как он может быть нанесён телу. Посылка Сократа представляет собой внутренний диалог, начинающийся с вопроса. Сократ задаёт вопрос: «Знаешь, какой опасности ты собираешься повергнуть свою душу?», затем эксплицирует этот вопрос и сам отвечает на него. Можно предположить, что цель внутреннего диалога Сократа — уточнить характер запрашиваемых знаний и что без этих уточнений его посылка не обеспечит ему доступ именно к тем знаниям, которые необходимы.

Из изложенного следует, что диалоговая транзакция является: (1) структурным элементом диалогового процесса, (2) имеет вопросно-ответную (интеррогативную) природу.

1.1.3. Типы посылок активного агента

Посылки типа «поисковое предписание» и посылки типа «задача»

Ранее мы интерпретировали посылку активного агента как поисковое предписание, детерминирующее область памяти реактивного агента, в которой находятся запрашиваемые знания. Такая интерпретация предполагает, что реактивный агент уже обладает запрашиваемыми знаниями. Однако это только один из возможных случаев. Можно привести примеры посылок активного агента, запрашивающие знания, отсутствующие в памяти реактивного агента и требующие от него выполнения некоторой ментальной работы или ментального усилия с участием ресурса внимания с целью формирования запрашиваемых знаний. Поэтому по отношению к участию ресурса внимания реактивного агента в формировании ответа разделим посылки активного агента на два класса: *посылки типа «поисковое предписание»* и *посылки типа «задача»*.

Для формирования знаний, запрашиваемых с помощью посылки типа «задача», реактивный агент должен решить задачу, ассоциированную с этой посылкой. Ответом является вариант решения, полученный реактивным агентом. Рассмотрим следующий пример посылки активного агента.

Сократ родился в 469 году до нашей эры. Сколько лет было бы Сократу сегодня?

Ясно, что, скорее всего, знания, запрашиваемые приведенной посылкой, не хранятся в памяти реактивного агента в готовом виде, и требуется использование ментального ресурса для их формирования.

Посылка типа «поисковое предписание» предполагает, что в момент получения посылки в памяти реактивного агента уже имеются запрашиваемые знания, а структурные элементы посылки позиционируют память на требуемый участок. Для формирования ответа на посылку типа «поисковое предписание» не требуется использование ресурса внимание. Примером может служить следующая посылка активного агента.

Как Вас зовут?

Класс посылок типа «поисковое предписание» может быть подвергнут декомпозиции, и в нём могут быть выделены подклассы. Например, может случиться так, что посылка активного агента типа «поисковое предписание» будет запрашивать знания, которые, во-первых, отсутствуют у реактивного агента, а во-вторых, которые нельзя сформировать решением задачи. Такие посылки могут поставить реактивного агента в тупиковую ситуацию, для выхода из которой понадобится смена ролей. Например, посылка

Почему Вы скрываете, что Вы мужчина?

может поставить в тупик женщину, которая не является мужчиной.

Классификация посылок активного агента на посылки типа «поисковое предписание» и посылки типа «задача» не является абсолютной и должна рассматриваться по отношению к знаниям реактивного агента. Если задача, порождаемая посылкой типа «задача», решается впервые или очень редко (настолько редко, что процесс забывания превалирует), то для её решения, согласно гипотезе Канемана, должен быть выделен ресурс внимания [5]. В случае многократного и частого ответа на одни и те же или близкие посылки типа «задача» происходит следующее: (1) решение запоминается, (2) для формирования ответа более не требуется использование механизма внимания, (3) посылка активного агента превращается в посылку типа «поисковое предписание». Если, например, реактивный агент впервые получает упомянутую выше посылку:

Сократ родился в 469 году до нашей эры. Сколько лет было бы Сократу сегодня?

то она по отношению к реактивному агенту должна рассматриваться как посылка типа «задача», поскольку реактивный агент не держит в памяти знания о том, сколько лет было бы Сократу сегодня, если бы он не был казнён в Афинах много лет тому назад. Для получения этих знаний необходимо решить арифметическую задачу. Однако, если реактивный агент получает эту посылку часто и многократно, то он, в конце концов, запоминает результат решения (гипотетический возраст Сократа), и отмеченная посылка переходит в класс посылок типа «поисковое предписание».

Неопределённость посылок активного агента

Анализ протоколов естественных диалогов позволяет обнаружить, что почти всегда посылки типа «поисковое предписание» обладают некоторой степенью неопределённости по отношению к классу запрашиваемых знаний. В них отсутствуют сведения о том, к какому точно классу относятся запраши-

ваемые знания. Чем более неопределённой является посылка активного агента, тем большей свободой обладает реактивный агент при выборе класса знаний для формирования ответа. Поэтому по отношению к точности указания класса запрашиваемых знаний, все послышки типа «поисковое предписание» можно разделить, по крайней мере, на два подкласса: неопределённые и определённые. Это грубая классификация, поскольку она не учитывает степень неопределённости посылки. Более точная классификация должна учитывать *степень неопределённости*, которая характеризуется количеством классов, к которым могут относиться запрашиваемые знания. Очевидно, что чем больше это количество, тем выше степень неопределённости посылки типа «поисковое предписание».

Неопределённое поисковое предписание означает, что ответная посылка реактивного агента может принадлежать одному из нескольких классов ответов. Отвечая на неопределённую посылку типа «поисковое предписание», реактивный агент может использовать одну из двух стратегий: (1) по своему усмотрению выбрать класс, к которому принадлежит ответ; (2) с целью уменьшения или снятия неопределённости вместо ответа сформулировать уточняющую посылку. Последний случай означает обмен ролями между диалоговыми агентами, при котором реактивный агент берёт на себя роль активного агента, а уточняющая посылка имеет статус вопроса. Таким образом, получение реактивным агентом неопределённой посылки типа «поисковое предписание» может рассматриваться как одна из причин смены ролей в диалоговом процессе. Очевидно, что процесс уточнения неопределённой посылки типа «поисковое предписание» может быть итерационным, если агент воспринимает новую посылку тоже как неопределённую. Например, посылка активного агента:

Кто живёт за этой дверью?

является неопределённой, поскольку запрашиваемые знания принадлежат одному из нескольких классов: (1) знания о воз-

расте жильца; (2) знания о внешности жильца и т.д. Реактивный агент, получив такую посылку и желая уточнить, какие знания запрашивает активный агент, может, изменив роль, вместо ответа передать своему партнёру следующую уточняющую посылку:

Вы хотите знать имя жильца, который живёт за этой дверью, либо что-то другое?

Если последующая посылка сформулирована в виде

Не могли бы Вы его описать?

то это опять неопределённая посылка типа «поисковое предписание», поскольку она предполагает, что запрашиваемые знания принадлежат одному из несколько классов.

Определённое поисковое предписание предполагает, что запрашиваются знания, принадлежащие только одному классу, например

Назовите фамилию человека, который живёт за этой дверью?

Посылки типа «ли» и посылки типа «какой»

Восприняв неопределённую посылку, реактивный агент в итоге отвечает на определённую посылку. Он приходит к этому либо путём смены роли и уточнения класса знаний, к которому принадлежит ответ, либо произвольным выбором этого класса. Каждая определённая посылка содержит *структурный компонент*, который задает класс запрашиваемых знаний. В приведенном, выше, примере определённой посылки

Назовите фамилию человека, который живёт за этой дверью?

этот класс знаний представляет собой некоторое множество фамилий.

Таким образом, один из структурных компонентов посылки типа «поисковое предписание» предназначен для того, чтобы детерминировать класс запрашиваемых знаний. Белнап и Стил в своей версии эротетической логики называют этот компонент применительно к вопросительному предложению «субъектом вопроса» [6]. Следуя Белнапу и Стилу, компонент определённой посылки типа «поисковое предписание», детерминирующий класс запрашиваемых знаний, будем называть *субъектом поискового предписания*.

Субъект поискового предписания может детерминировать класс, в котором есть конечное количество экземпляров, либо класс, который содержит бесконечное количество экземпляров. Поисковые предписания, субъекты которых детерминируют конечный класс, назовём *ли(whether)-посылками*, а поисковые предписания, субъекты которых детерминируют классы с бесконечным количеством экземпляров — *какой(which)-посылками*.

В первой транзакции приведенного фрагмента диалога «Протагор» Сократ, являясь активным агентом, передает Гиппократу следующую посылку.

Так мы идём платить деньги Протагору, потому что он софист?

Если предположить, что Гиппократ имеет заранее сформированное мнение относительно того, по какой причине он хочет платить деньги Протагору, то приведенная посылка является посылкой типа «поисковое предписание». Далее, поскольку ответ Гиппократа принадлежит только одному классу ответов, то эта посылка является определённой. Этот единственный класс ответов есть конечным и состоит только из двух экземпляров: (1) утвердительный ответ и (2) отрицательный ответ. Поэтому в соответствии с только что введенной классификацией можем классифицировать рассматриваемую посылку как *ли-посылку*.

Множество примеров *какой-посылок* приведены в книге Белнапа и Стила [6]. Например,

Какова температура замерзания воды по Фаренгейту при нормальных условиях?

Если предположить, что реактивный агент знает ответ, то это — определённая посылка, поскольку запрашиваемые знания принадлежат одному классу, экземпляры которого представляют собой температуры замерзания жидкостей при нормальных условиях. Количество экземпляров этого класса настолько велико, что можно считать эту посылку какой-посылкой.

Отметим, что, в ряде случаев субъекты ли-посылок могут быть заданы перечислением экземпляров соответствующих классов. Примером может быть субъект следующей ли-посылки, заимствованной из книги Белнапа и Стила.

Курение табака — это порок, добродетель, причуда, сумасбродство или панацея от всех невзгод?

На рис. 1.1 введенные выше типы посылок активного агента представлены в виде UML диаграммы классов.

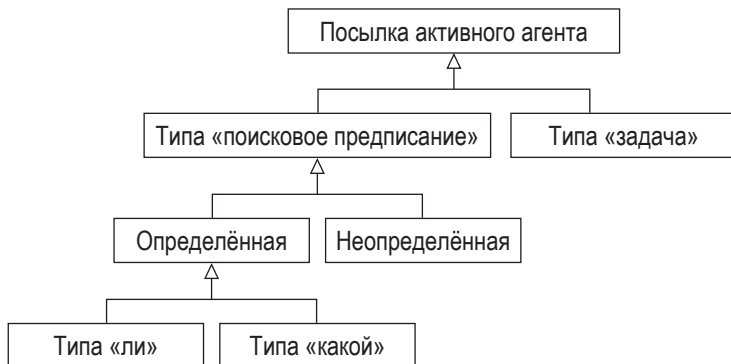


Рис. 1.1. Типы посылок активного агента

1.1.4. Типы посылок реактивного агента

Прямые и не прямые послылки

Посылку реактивного агента мы понимаем как поисковое предписание, необходимое для получения доступа к запрашиваемым знаниям, или как некоторую ссылку на участок долговременной памяти активного агента, где хранятся запрашиваемые знания. Реактивный агент, формируя свою посылку, может непосредственно сослаться на запрашиваемые знания. Такую посылку будем называть *прямой посылкой реактивного агента*. Примером транзакции, в которой посылка реактивного агента непосредственно ссылается на запрашиваемые знания, может служить следующая.

- Активный:** Какова температура замерзания воды по Фаренгейту при нормальных условиях?
Реактивный: Температура замерзания воды при нормальных условиях равна 32 градусам по шкале Фаренгейта.

Однако реактивный агент, формируя свою посылку, может сослаться на запрашиваемые знания опосредованно. Такую посылку будем называть *непрямой посылкой реактивного агента*. Приведенную выше транзакцию несложно представить в форме, когда посылка реактивного агента ссылается на запрашиваемые знания опосредовано.

- Активный:** Какова температура замерзания воды по Фаренгейту при нормальных условиях?
Реактивный: Температура замерзания воды при нормальных условиях по Фаренгейту указана в приложении 1.

Кодированные и не кодированные послылки

В протоколах естественных диалогов часто встречаются случаи, когда реактивный агент сокращает размер прямой по-

сылки с помощью кодирования или замены её на компактный код, смысл которого определяется контекстом посылки активного агента. Так в первом из приведенных выше примеров реактивный агент вместо прямой посылки

Температура замерзания воды при нормальных условиях равна 32 градусам по шкале Фаренгейта

мог бы передать её код, например, в виде числа, соответствующего запрашиваемой температуре.

Активный: Какова температура замерзания воды по Фаренгейту при нормальных условиях?

Реактивный: 32.

В первой транзакции приведенного выше фрагмента диалога «Протагор»

С: Так мы идём платить деньги Протагору, потому что он софист?

Г: Конечно.

Слово «конечно», возвращаемое Гиппократом в качестве прямой посылки, является кодом, заменяющим посылку «Да, мы идём платить деньги Протагору, потому что он софист».

Приведенные примеры транзакций позволяют разделить прямые послылки реактивного агента на два класса: *кодированные послылки* и *не кодированные послылки*.

1.1.5. Истинность информационной посылки реактивного агента

Темой, требующей отдельного изучения, является понятие истинности информационной посылки реактивного агента. Если речь идёт об отдельной и изолированной диалоговой транзакции, то можно согласиться с абсолютной трактовкой истинности ответа, принятой в эротетической логике [6]. Од-

нако если речь идёт о диалоговой транзакции, встроенной в естественный диалоговый процесс, то понятие *истинности информационной посылки реактивного агента становится релятивным* и может определяться по отношению к ряду факторов: (1) цель реактивного агента; (2) знания реактивного агента; (3) намерения реактивного агента (намерен ли он возвращать активному агенту то, что он действительно знает или намерен в ответе исказить свои знания).

Таким образом, проблема истинности информационной посылки реактивного агента при моделировании естественного диалога не может решаться простой дихотомией, как это принято в эротетической логике и эпистемологии, где утверждается, что посылка-ответ может быть либо истинной, либо ложной [6,7]. Такое абсолютное понимание истинности ответа, по всей видимости, связано с тем, что как в эпистемологии, так и в эротетической логике анализу подвергаются изолированные вопросно-ответные пары (или только отдельные информационные посылки) и связанные с ними знания, а не протоколы естественных диалогов. Целенаправленное диалоговое взаимодействие в виде логически связанных транзакций *предопределяет относительность понятия истинности информационной посылки реактивного агента*.

Реактивный агент в ходе вопросно-ответного диалога решает свою проблему и формирует ответы, исходя не только из содержания вопросов, но и из той цели, которую он старается достичь в ходе диалога. По отношению к цели реактивного агента можно сформулировать примерный список правил, которыми он может руководствоваться при конструировании ответа на определенный вопрос. Например, (1) если вопрос затрагивает нежелательные для меня предметы, то я искренне отвечаю на него; (2) если вопрос затрагивает нежелательные для меня предметы, то в ответе я скрываю, либо искажаю часть того, что знаю; (3) если вопрос затрагивает нежелательные для меня предметы, то в ответе я скрываю или искажаю всё, что знаю.

Для того чтобы при формировании посылки реактивный агент мог выбрать одно из приведенных выше правил, он дол-

жен «продумать ответ» или провести некоторые рассуждения, предваряющие ответ. В процессе этих рассуждений реактивный агент оценивает целесообразную степень полноты/искренности ответа по отношению к цели реактивного агента. Отметим, что такие рассуждения являются опцией, которой может воспользоваться опытный и осторожный агент или не воспользоваться неопытный и темпераментный реактивный агент.

1.1.6. Диалоговое поведение

Диалог — это целенаправленный интеллектуальный процесс, в котором каждый из агентов диалога стремится к достижению собственной цели. Агенты формируют потоки информационных посылок, которые согласуются с их целью. В этом смысле диалоговые процессы являются процессами решения проблем. Легко заметить целенаправленный характер информационных посылок Сократа в приведенном фрагменте диалога «Протагор». Последовательность его вопросов направлена на достижение цели, которую можно сформулировать, например, следующим образом: «убедить Гиппократ в том, что его первоначальное желание стать учеником Протагора, заплатив ему за это любые деньги, является ошибочным».

В связи с очевидной связью проблематики естественного диалога с проблематикой решения проблем введём понятие *диалогового поведения*, под которым будем понимать реализацию метода решения проблемы с помощью диалогового процесса. В приведенном фрагменте диалога Сократ демонстрирует некоторое диалоговое поведение, которое является реализацией метода, известного как *метод вопросов Сократа* или *диалог Сократа* [8].

Рассмотрим, кратко, сущность метода вопросов Сократа. Сократ мог бы достичь свою цель несколькими способами. Например, он мог бы, апеллируя к своему авторитету, сразу предложить окончательный вывод, сказав: «Поверь мне Гиппократ,

не следует стремиться стать учеником Протагора любой ценой, не выяснив, чему он будет тебя учить и будут ли полезны приобретённые у него знания». В этом случае Сократ передал бы Гиппократу скорее не знания, а веру, опирающуюся на свой авторитет, а не на логический вывод.

Однако Сократ предпочитает вступить в диалог с Гиппократом, предложив ему серию интеррогативных информационных посылок с небольшим количеством предопределённых вариантов ответов. Информационные послышки Сократа сформулированы таким образом, чтобы ответы Гиппократа составили цепь рассуждений, которая завершается сформулированным выше выводом. Это более действенный способ, поскольку Гиппократ, ведомый Сократом, сам выводит сформулированное выше заключение. В этом суть метода Сократа. Находясь в роли активного агента, Сократ, таким образом, управляет диалоговым процессом, что побуждает оппонента отвечать так, как это требуется для формирования цепи умозаключений, завершающейся целевым заключением. Секрет успеха метода Сократа в том, что с каждым вопросом ассоциировано некоторое множество возможных ответов, которое не зависит от знаний и ответа реактивного агента, а детерминируется исключительно самим вопросом. Зная это, Сократ таким образом строит свои информационные послышки, чтобы максимально сузить количество вариантов возможных ответов и по возможности свести их к единице.

1.1.7. Знания, переносимые вопросно-ответной транзакцией

В вопросно-ответном диалоге информационная посылка активного агента имеет стопроцентную степень интеррогативности и форму вопроса в грамматическом смысле этого слова. В вопросно-ответном диалоге знания, которые запрашивает активный агент при помощи вопроса, и знания, которые возвращает реактивный агент при помощи ответа, являются *декларативными*. Поэтому модель структуры диалоговой транзакции в

вопросно-ответном диалоге в контексте представления знаний должна основываться на некоторой приемлемой концепции представления декларативных знаний. Хотя система декларативных знаний человека, по всей видимости, когерентна и подобна ткани, в которой близлежащие части сцеплены между собой, вопросно-ответная транзакция имеет дело с отдельной небольшой частью или порцией системы декларативных знаний человека. Поэтому нам в дальнейшем понадобятся средства представления только порции декларативных знаний, ассоциированных с вопросно-ответной транзакцией.

В эпистемологических публикациях, как правило, не используется терминология декларативно-процедурной дихотомии знаний, которая принята в когнитивной психологии. В эпистемологических публикациях встречается категоризация, согласно которой знания делятся на *знания-that* и *знания-wh*. Однако из контекста этих публикаций можно заключить, что *знания-that* и *знания-wh* — это не что иное, как эпистемологическое наименование порции декларативных знаний реактивного агента, участвующего в вопросно-ответной транзакции [4]. Эти знания рассматриваются как естественно-языковые пропозиции, являющиеся ответом на *that*- или *wh*-вопрос, соответственно.

Wh-вопрос — это наименование типа вопросов, обобщающего шесть подтипов специальных вопросов английского языка: (1) *who/кто-вопрос*; (2) *what/что-вопрос*; (3) *when/когда-вопрос*; (4) *where/где-вопрос*; (5) *how/как-вопрос*; (6) *why/почему-вопрос*.

Знания, выражаемые ответами на вопросы из перечисленных подтипов, формируют шесть подтипов декларативных знаний. Таким образом, *знания-wh* — это наименование типа знаний, обобщающего шесть подтипов: (1) *who/кто-знания*; (2) *what/что-знания*; (3) *when/когда-знания*; (4) *where/где-знания*; (5) *how/как-знания*; (6) *why/почему-знания*.

Эпистемологическое понимание *знаний-wh* можно проиллюстрировать примерами следующих шести утверждений о *знаниях-wh*: (1) «Я знаю кто есть Протагор»; (2) «Я знаю, что

он будет говорить»; (3) «Я знаю, когда он начнёт разговор»; (4) «Я знаю, где будет происходить разговор»; (5) «Я знаю, как Протагор начнёт говорить» и (6) «Я знаю, почему его аргументы некорректны». В контексте диалоговой транзакции приведенные утверждения являются примерами декларативных знаний реактивного агента.

1.2. Логическая структура вопросно-ответной транзакции

Логическая структура отдельной диалоговой транзакции в том случае, когда информационная посылка активного агента является вопросом в грамматическом смысле этого слова, а информационная посылка реактивного агента — ответом, рассматривается в публикациях, посвященных *эротетической логике*. Версии эротетической логики, ставшие классическими, изложены в работах: Н. Белнапа и Т. Стила (Nuel Belnap & Thomas Steel); [6] Т Кубински (Tadeusz Kubinski); [9] Д. Хинтикка (Jaakko Hintikka) [10] и А. Вишиевского (Anrzej Wisniewski) [11]. Этот подраздел посвящен критическому обзору теории, предложенной Белнапом и Стилом. Обзор проведен в контексте вопросно-ответного диалогового процесса с акцентом на проблематику представления знаний и положения когнитивной психологии. Например, формализмы вопросно-ответных пар рассматриваются как внеязыковые, семантические структуры, а проблемы интеррогативного вывода — в контексте теории интеллектуальных агентов, решающих проблемы (problem-solving agents).

Прежде, чем углубляться в основы эротетической логики, отметим один важный когнитивный аспект, связанный с вопросно-ответным диалоговым процессом. Что побуждает активного агента диалогового процесса задавать вопросы и поддерживать вопросно-ответный диалог? По нашему убеждению, *стимулом формирования любого вопроса является потребность в знани-*

ях, которые активный агент диалога надеется получить от реактивного агента. Находясь в состоянии перманентного решения проблемы, активный агент детерминирует знания, необходимые для продолжения решения проблемы. Этот процесс можно мыслить как внутренний диалог активного агента с самой собой. Определив, какие знания ему необходимы, активный агент транслирует релевантную семантическую структуру вопроса в вербальную либо вербально-образную форму и передаёт вопрос реактивному агенту. Тот очевидный факт, что стимулом для формирования вопроса является потребность в знаниях со стороны активного диалогового агента, как правило, явно не отмечается в публикациях по эротетической логике, однако с точки зрения проблематики вопросно-ответного диалогового процесса является одним из ключевых.

1.2.1. Онтологический базис эротетической логики

Теория Белнапа и Стила ориентирована исключительно на естественно-языковые вопросно-ответные пары и это отразилось как на их классификации, так и на тех формализмах, которые предлагаются для моделирования логической структуры вопросов и релевантных им ответов. Теория не является универсальной по отношению ко всем возможным естественно-языковым вопросам и фактически базируется на примерах естественно-языковых вопросов, приведенных в книге. Большинство примеров вопросов и ответов данного подраздела заимствованы из [6].

Примером вопроса, который не охватывается теорией Белнапа и Стила, является вопрос, приведенный в подразделе 1.1.3 в качестве примера информационной посылки активного агента, которая обладает высокой степенью неопределённости.

Кто живёт за этой дверью?

Реактивному агенту сложно ответить на этот вопрос, поскольку он не знает, что имеет в виду активный агент: фами-

лию жильца, его профессию и занимаемую должность, описание внешности, или что-то другое.

Важным элементом онтологического базиса теории Белнапа и Стила является понятие прямого ответа (*direct answer*). *Прямой ответ* — это те знания, которые нужны активному агенту и ради которых он задаёт вопрос. Эти знания активный агент ожидает получить от реактивного агента. Прямой ответ удовлетворяет потребность в знаниях активного агента на некотором этапе решения им ментальной проблемы. Поскольку есть прямые ответы, то, очевидно, существуют и непрямые ответы. Для иллюстрации прямого и непрямого ответов Белнап и Стил приводят следующий пример:

Вопрос

Какова температура замерзания воды по Фаренгейту при нормальных условиях?

Прямой ответ

Температура замерзания воды при нормальных условиях равна 32 градусам по шкале Фаренгейта.

Непрямой ответ

Температура замерзания воды при нормальных условиях указана в справочнике по физике.

Прямой ответ понимается Белнапом и Стилом как некоторое предложение, которое может быть истинным или ложным. Так, приведенный выше прямой ответ считается истинным. Ответ, содержащий любую другую температуру, например 35°, считается ложным. Теория Белнапа и Стила ориентирована исключительно на прямые ответы, поэтому в дальнейшем во всех случаях, когда будем употреблять термин ответ, всегда будем подразумевать прямой ответ.

Приписывая ответу истинность или ложность, Белнап и Стил абсолютизируют понятие истинности ответа, утверждая, что с любым ответом необходимо ассоциировать одну из альтернативных оценок «истина» или «ложь». Абсолютизация понятия истинности или ложности ответа является важным элементом философии Белнапа и Стила, которая нашла отражение практически во всех предложенных ими формально-логических моделях вопросов.

Однако, как отмечено в подразделе 1.1.5, такая абсолютизация понятия истинности или ложности является упрощением реальной ситуации. В общем случае понятие истинности или ложности ответа относительно и определяется по отношению к диалоговому агенту. Это становится очевидным, если рассматривать вопросно-ответную пару не саму по себе, а в контексте вопросно-ответного диалогового процесса. Вопросы и ответы формируются различными агентами, обладающими разными системами знаний и преследующими, возможно, разные цели в диалоговом процессе. Ответ формируется реактивным агентом и его истинность или ложность должны рассматриваться по отношению к системе знаний реактивного агента и цели, которую он преследует в ходе диалогового процесса. Поскольку система знаний и цели агентов диалога могут не совпадать, то ответ, который является ложным для активного агента, может быть вполне истинным для реактивного. Например, в вопросе о курении табака, который приведен в подразделе 1.1.3, отношение к курению табака активного агента может не совпадать с отношением к курению реактивного агента. Активный агент может считать, что курение — это порок, а реактивный агент — что это добродетель. Если на вопрос о курении табака (этот вопрос рассматривается в книге Белнапа и Стила как пример ли-вопроса) реактивный агент отвечает, что — это «добродетель», то по отношению к реактивному агенту мы должны считать этот ответ истинным. Однако, если активный агент убеждён, что курение табака — «это порок», то по отношению к активному агенту этот ответ — ложный.

В некоторых случаях, например во время устного экзамена, экзаменатор, являясь активным агентом, решает задачу оценки знаний реактивного агента-студента. Экзаменатор формирует вопрос, на который он а priori знает правильный ответ (истинный с его точки зрения). Однако знания студента и экзаменатора могут не совпадать, и студент вполне может сформировать истинный с его точки зрения ответ, не совпадающий с эталонным, ожидаемым экзаменатором.

В других случаях активный агент а priori не знает ответа на сформированный им вопрос и должен доверять (или не доверять) ответам реактивного агента. В тех случаях, когда активный агент полностью доверяет ответам реактивного агента, то любой полученный ответ он должен считать истинным.

Белнап и Стил постулируют, что логическая структура *любого* вопроса представлена двумя компонентами:

- *субъектом* (subject) и
- *требованием* (request).

В русскоязычном издании книги Белнапа и Стила слово request переводится как *предпосылка* [12].

Субъект вопроса явно или неявно определяет множество утверждений. Утверждения понимаются в смысле ассерторических суждений, в которых констатируется наличие или отсутствие у предмета некоторого признака. В субъекте вопроса *могут быть как истинные, так и ложные утверждения*. Таким образом, Белнап и Стил абсолютизируют понятие истинности или ложности не только по отношению к прямому ответу, но и по отношению к утверждениям субъекта вопроса. Абсолютизация понятия истинности или ложности по отношению к утверждениям субъекта вопроса предполагает, что имеется некоторая эталонная система знаний вне сознания активного или реактивного агента, по отношению к которой осуществляется оценка истинности или ложности утверждения субъекта вопроса. Истинность или ложность утверждений субъекта, в общем случае, должны рассматриваться как относительные, а не как абсолютные, а оценка утверждения осуществляется по отношению к системе знаний активного или реактивного агента.

Формируя ответ, реактивный агент осуществляет выбор некоторого количества утверждений субъекта, которые переносятся в ответ. Белнап и Стил называют утверждения субъекта *альтернативами* (alternatives), что не совсем точно, поскольку в ответ может переходить несколько утверждений субъекта, а альтернативность этого не предполагает. Поэтому в дальнейшем будем называть эти утверждения *элементами субъекта*. Например, вопрос

Является ли стекло жидкостью при температуре 70° F?
предполагает, что субъект состоит из двух элементов:

Стекло является жидкостью при температуре 70° F.
Стекло не является жидкостью при температуре 70° F.

В приведенном примере множество элементов субъекта конечно. Однако субъект вопроса может состоять из бесконечно-го множества элементов. Так, рассмотренный ранее вопрос

Какова температура замерзания воды по Фаренгейту при нормальных условиях?

предполагает бесконечно много элементов в субъекте. Например,

Температура замерзания воды при нормальных условиях равна 0° F.
Температура замерзания воды при нормальных условиях равна 1° F
и так далее.

Требование вопроса является, по сути, набором ограничений, накладываемых на возможные ответы. Требование необходимо для специфицирования того, в каком виде должен быть представлен ответ, формируемый реактивным агентом из субъекта. Для иллюстрации того, что в логической структуре

вопроса кроме субъекта должен присутствовать компонент, названный требованием, Белнап и Стил приводят примеры нескольких вопросов, имеющих одинаковые субъекты, но предполагающие разные ответы.

Вопрос 1.

Какие простые числа находятся между 10 и 20?

Вопрос 2.

Что является примером простого числа, находящегося между 10 и 20?

Оба вопроса обладают одним и тем же субъектом — это, например, множество целых чисел, находящихся между числами 10 и 20. Ответ на первый вопрос представляет собой некоторый список простых чисел, находящихся между числами 10 и 20, а ответ на второй вопрос — одно из простых чисел, выбранных реактивным агентом в качестве примера.

Как видно из приведенных примеров и как будет показано в дальнейшем, одному вопросу в общем случае соответствует некоторое количество ответов, а *отдельный ответ может рассматриваться как конъюнкция элементов субъекта*.

Требование вопроса описывается тремя компонентами:

- *спецификация размера выборки* (selection-size-specification),
- *спецификация требования полноты* (completeness-claim-specification),
- *спецификация требования различения* (distinctness-claim-specification).

Под выборкой понимается подмножество элементов субъекта, переходящих в ответ. Спецификация размера выборки детерминирует количество элементов субъекта, которые должны войти в выборку и перейти в ответ. Размер выборки специфицируется указанием нижнего и верхнего пределов для количества элементов.

Спецификация требования полноты дополняет спецификацию размера выборки и специфицирует долю истинных элементов субъекта, переходящих в ответ по отношению ко всему множеству истинных элементов субъекта. Иными словами, это требование того, какая часть всех истинных элементов субъекта должна перейти в ответ.

Спецификация требования различения — это компонент требования, который индицирует, необходимо ли в ответе указывать, являются ли элементы субъекта реально или номинально различными. Реальные элементы субъекта — это некоторые сущности мира, а номинальные — их имена. Необходимость этого компонента объясняется необходимостью различать в качестве ответа различные имена для одной и той же сущности. Например, число «семь» может иметь два разных имени: арабская цифра «7» или римская цифра «VII».

Требование конкретного вопроса включает некоторое количество спецификаций. *Обязательной является спецификация размера выборки.* Требование любого вопроса должно обязательно включать эту спецификацию. Спецификации выбора полноты и различения являются необязательными и включаются в требование только некоторых вопросов. Таким образом, требование вопроса может быть представлено одной, двумя или тремя спецификациями. Формально требование вопроса может быть представлено при помощи одного из четырёх выражений:

$$s \wedge c \wedge d, s \wedge c, s \wedge d, s, \quad (1.1)$$

где:

s — спецификация размера выборки (selection);

c — спецификация требования полноты (completeness);

d — спецификация требования различения (distinctness).

Задачей обязательного компонента требования вопроса (спецификация размера выборки) является специфицирование мощности подмножества элементов субъекта, переходящих в ответ.

Спецификация требования полноты имеет смысл тогда, когда субъект может включать как истинные, так и ложные

утверждения. Если все элементы субъекта истинны, то спецификация требования полноты автоматически задаётся спецификацией размера выборки.

Спецификация требования различения имеет смысл тогда, когда элементы субъекта представлены именами (например, закрытыми терминами в смысле логики первого порядка), среди которых допускаются различные имена, именующие одну и ту же сущность. Если при конструировании субъекта вопроса запретить использование различных имён для наименования одной и той же сущности, то отпадает необходимость в спецификации требования различения.

Вопросы классифицируются по отношению к мощности субъекта. Субъект может детерминировать либо (1) конечное или небольшое и обозримое количество элементов, либо (2) бесконечное или очень большое количество элементов. В первом случае соответствующие вопросы называются *ли-вопросами* (*whether-questions*), а во втором — *какой-вопросами* (*which-questions*). В подразделе 1.1.3 мы использовали «ли» и «какой» терминологию для категоризации посылок активного агента с любой степенью неопределённости. Здесь речь идёт исключительно об определённых посылках активного агента, которые в случае вопросно-ответного диалога называем *вопросами*.

Примером *ли-вопроса* является вопрос о курении табака, который якобы сформулировал английский король Джеймс первый:

Курение табака — это порок, добродетель, причуда, сумасбродство или панацея от всех невзгод?

Субъект этого вопроса состоит точно из пяти элементов:

- (1) курение табака — это порок,
- (2) курение табака — это добродетель,
- (3) курение табака — это причуда,
- (4) курение табака — это сумасбродство,
- (5) курение табака — это панацея от всех невзгод.

Примером какой-вопроса является следующий вопрос:

Какое натуральное число является наименьшим простым числом, большим, чем 45?

Субъект этого вопроса состоит из бесконечного количества элементов:

1 является наименьшим простым числом, большим, чем 45,
2 является наименьшим простым числом, большим, чем 45,
3 является наименьшим простым числом, большим, чем 45,
и так далее.

1.2.2. Интеррогативная формула ли- и какой-вопросов

Формально-логическая модель ли- или какой-вопроса называется *интеррогативной формулой* или интеррогативным предложением (interrogative formula or sentence) или просто интеррогативом (interrogative). Интеррогативная формула в общем случае записывается в виде

$$\text{Intr} = ? \rho \sigma, \quad (1.2)$$

где:

ρ — требование вопроса;

σ — субъект вопроса;

? — признак интеррогатива.

С учётом трёхкомпонентной структуры требования интеррогативная формула может быть записана в виде

$$\text{Intr} = ?(s c d) \sigma. \quad (1.3)$$

Если интеррогативная формула моделирует ли-вопрос, то она называется *ли-интеррогативом* (whether-interrogative) и записывается в виде

$$\text{Whether-Intr} = ? \rho(A_1, \dots, A_n), \quad (1.4)$$

где:

(A_1, \dots, A_n) — субъект ли-вопроса, или *ли-субъект*.

Ли-субъект представляет собой список утверждений, относительно которого вводится несколько ограничений. Во-первых, предполагается, что среди элементов списка нет повторяющихся. Во-вторых, предполагается, что список может включать как элементы, представляющие собой простые предложения, так и элементы, представляющие собой конъюнкции простых предложений. При этом хотя бы один из конъюнктов не должен являться элементом списка. Например, допустим список вида $(A, A \wedge B)$, но не допустим список $(C, A, B, A \wedge B)$.

Отмеченные ограничения вводятся умозрительно, и их необходимость никак не объясняется. Более того, сомнительным является необходимость включения конъюнкций в субъект вопроса. Спецификация размера выборки на деле означает, что в ответе могут фигурировать конъюнкции некоторого количества простых утверждений субъекта. Поэтому включение в субъект ли-вопроса элементов, представляющих собой конъюнкцию простых утверждений, является избыточным требованием.

Частным случаем ли-вопроса является вопрос типа да/нет. Субъект вопроса типа да/нет состоит из двух простых взаимно противоположных утверждений. Для рассмотренного ранее примера ли-вопроса

Является ли стекло жидкостью при температуре 70^0 F?

субъект представляется двумя взаимно противоположными утверждениями:

Стекло является жидкостью при температуре 70^0 F.
Стекло не является жидкостью при температуре 70^0 F,

а интеррогатив записывается в виде

$$Yes / No-Intr = ? \rho(A, \bar{A}), \quad (1.5)$$

где:

A — утверждение Стекло является жидкостью при температуре 700 F.

\bar{A} — утверждение Стекло не является жидкостью при температуре 700 F.

Рассмотрим пример, при помощи которого Белнап и Стил иллюстрируют необходимость использования конъюнкции в субъекте ли-вопроса. Ли-субъект вопроса

Перестал ли Джон бить свою жену?

может быть представлен двумя утверждениями:

Джон бил свою жену (утверждение A),

Джон сейчас не бьёт свою жену (утверждение B).

Тогда ли-интеррогатив этого вопроса имеет вид

$$\textit{Whether-Intr} = ? \rho((A \wedge B), (A \wedge \bar{B})). \quad (1.6)$$

Отметим, что приведенный пример вопроса можно рассматривать и как вопрос типа да/нет. Тогда его субъект представляет собой два простых взаимно противоположных утверждения.

Ли-интеррогатив вопроса короля Джеймса о курении табака записывается в виде

$$\textit{Whether-Intr} = ? \rho(A_1, A_2, A_3, A_4, A_5). \quad (1.7)$$

Рассмотрим, каким образом Белнап и Стил предлагают формально представлять субъект и интеррогативную формулу какой-вопроса. Вначале приведём несколько примеров.

Для какой-вопроса

Какое натуральное число является наименьшим простым числом, большим, чем 45?

субъект может быть задан при помощи следующих двух предложений с переменной, обозначенной буквой x .

x является наименьшим простым числом, большим, 45.
 x является натуральным числом.

Первое из приведенных предложений названо *матрицей* (matrix) *субъекта какой-вопроса*, а второе — *категорным условием* (category condition) *субъекта какой-вопроса*. Матрица представлена при помощи естественно-языкового предложения, в которое включён символ переменной x . Категорное условие ограничивает значения переменной x в матрице. Заменяя в матрице переменную x её значениями, удовлетворяющими категорному условию, можно получить сколь угодно много элементов субъекта.

В матрице может быть несколько переменных. В этом случае должно быть задано несколько категорных условий, ограничивающих каждую из переменных. Например, субъект вопроса

Какие мальчики являются братьями каких девочек?

может быть задан следующим набором предложений:

x является братом y (матрица).
 x является мальчиком (первое категорное условие).
 y является девочкой (второе категорное условие).

В приведенных примерах какой-вопросов значения переменных в матрице субъекта и в категорных условиях представляют собой имена сущностей, имеющих категориальный статус вещи. Это подсказывает иной взгляд на формально-логическое моделирование вопроса и, в частности, на способ представления декларативных знаний, переносимых субъектом. Нотация, предлагаемая Белнапом и Стилом для формально-логического моделирования вопросов, сильно привязана к естественному языку. Она предполагает, что элементы субъекта, а также ма-

трица и категорные условия являются естественно-языковыми предложениями. Однако элементами субъекта могут быть также образы или изображения. Такого типа субъекты часто используются в вопросах, используемых в психодиагностических тестах. Для того чтобы уйти от привязки к естественному языку и сделать формально-логические модели вопросов более универсальными, целесообразно воспользоваться онтологическим базисом и, возможно, нотацией, которые предложены и развиты Авениром Уёмовым и его учениками в языке *тернарного описания* [13–16]. Полученные в рамках отмеченного онтологического базиса формально-логические модели вопросов и ответов приобретают статус моделей представления знаний в виде семантических структур вопросно-ответных пар и могут транслироваться как в символьное естественно-языковое представление, так и в образное представление.

Таким образом, согласно Белману и Стилу субъект какой-вопроса, или *какой-субъект* (which-subject) представлен тройкой

$$(X, g, Ax_1 \dots x_n), \quad (1.8)$$

где:

X — множество переменных $\{x_1 \dots x_n\}$, в матрице субъекта;

g — категорное отображение;

$Ax_1 \dots x_n$ — матрица субъекта.

Категорное отображение (category mapping) g ставит в соответствие элементам множества X категорные условия. Отношение между множеством X и набором категорных условий не взаимно однозначное. Если некоторой переменной из множества X не поставлено в соответствие категорное условие, то это означает, что переменная никак не ограничена и может принимать произвольные значения. Такая переменная называется *категорно свободной* (category free).

Ясно, что ли-субъект является частным случаем какой-субъекта. Какой-субъект превращается в ли-субъект в случае, когда количество значений переменных, подставляемых в матрицу, конечно.

Вместо символа категорного отображения g в какой-субъекте можно явно указать список категорных условий, а какой-субъект записать в виде

$$(C_1x_1, \dots, C_r x_r, x_{r+1}, \dots, x_n // Ax_1 \dots x_n), \quad (1.9)$$

где:

r — количество переменных, ограниченных категорными условиями;

$C_1x_1, \dots, C_r x_r$ — список категорных условий для ограниченных переменных;

x_{r+1}, \dots, x_n — список категорно свободных переменных;

$Ax_1 \dots x_n$ — матрица.

Нотация какой-субъекта в виде (1.9) используется для записи *какой-интеррогати́ва* (which-interrogative)

$$Which-Intr = ? \rho(C_1x_1, \dots, C_r x_r, x_{r+1}, \dots, x_n // Ax_1 \dots x_n). \quad (1.10)$$

Запись какой-субъекта в формуле (1.10) отличается от записи какой-субъекта в формуле (1.8). Формула (1.10) явно не содержит множество X , наверное поэтому Белнап и Стил не обсуждают вопрос о том, что является значением переменной для этого множества. Из дальнейшего изложения следует, что это закрытый терм, или имя некоторого объекта. Однако, если развивать идею использования онтологического базиса языка тернарного описания Уёмова для формального представления универсального субъекта и его элементов, то более важным является категориальный статус объекта (вещь, свойство или отношение), а не его имя.

Идея представления какой-субъекта в виде матрицы и списка категорных условий обладает тем недостатком, что допускает *многозначность трансляции формально-логической модели одного и того же вопроса* в естественно-языковые предложения. Отмеченная многозначность обусловлена тем, что матрица субъекта и, следовательно, соответствующий ей список категорных условий могут быть записаны множеством различных

способов. Это легко продемонстрировать на примере вопроса о простых числах, находящихся в диапазоне от 10 до 20.

Какие простые числа находятся между 10 и 20?

Какой-интеррогатив этого вопроса имеет вид

$$\textit{Which-Intr} =? \rho(Cx / Ax). \quad (1.11)$$

В интеррогативе (1.11) категорное условие Cx и матрица Ax могут быть представлены различным образом. Например,

x является целым числом // x является простым числом между 10 и 20,
 x является простым числом // x находится между 10 и 20.

Избавиться от этого недостатка можно, если конструировать субъект, состоящий не из естественно-языковых элементов, а из семантических элементов, инвариантных к естественному языку, и, как отмечено выше, хорошим кандидатом для этого является язык тернарного описания и такие категории как вещь, свойство и отношение.

Ответ на любой ли- или какой-вопрос представляет собой часть субъекта, выделенную из всего субъекта в соответствии с требованием вопроса. Часть субъекта вопроса, переходящая в ответ, может быть представлена *в виде конъюнкции либо в виде списка элементов субъекта*. Это эквивалентные представления, поскольку из списка предложений выводится их конъюнкция (правило введения конъюнкции). Количество элементов списка или количество конъюнктов — это размер выборки, который задаётся спецификацией размера выборки. Таким образом, формально-логическая структура ответа в общем случае может быть записана в виде

$$\textit{Ans} = \left(\bigwedge_{i=1}^{i=p} A_i \right) \wedge c \wedge d, \quad (1.12)$$

где:

A_i — элемент субъекта вопроса;

p — количество элементов субъекта вопроса, переходящее в ответ;

c — необязательная спецификация требования полноты;

d — необязательная спецификация требования различения.

1.2.3. Спецификация размера выборки

Ответ формируется реактивным агентом из субъекта вопроса в соответствии с требованием вопроса, состоящим из некоторого количества спецификаций. Напомним, что полный список спецификаций состоит из:

- спецификации размера выборки (s);
- спецификации требования полноты (c);
- спецификация требования различения (d).

Спецификация размера выборки предполагает, во-первых, что в ответ переходит подмножество элементов субъекта, а во-вторых, что *реактивный агент должен осуществить операцию выбора* этих элементов. Сама спецификация задаёт мощность подмножества элементов субъекта в виде значения переменной p в выражении (1.12).

Как было отмечено в подразделе 1.1.5, Белнап и Стил абсолютизируют понятие истинности ответа и считают, что вне зависимости от системы знаний и убеждений агентов вопросно-ответного диалога, каждому элементу субъекта вопроса можно приписать значение истинности или ложности. В эротетической логике утверждается, что в *ответ могут перейти только истинные элементы субъекта вопроса*.

Мощность подмножества элементов субъекта, переходящих в ответ, задаётся указанием верхнего (u) и нижнего (v) пределов выборки. Предлагается нотация, согласно которой верхний предел записывается на месте верхнего индекса, а нижний — на месте нижнего индекса. Таким образом, с учётом нотации для представления спецификации размера выборки при помо-

щи верхнего и нижнего пределов интеррогативная формула (1.1.3) может быть записана в виде

$$\text{Intr} = ? \binom{u}{v} cd \sigma, \quad (1.13)$$

где:

u — наибольшее количество элементов субъекта, переходящих в ответ (наибольшее количество элементов выборки);

v — наименьшее количество элементов субъекта, переходящих в ответ (наименьшее количество элементов выборки).

Предполагается, что верхний и нижний пределы принимают значения на множестве натурального ряда чисел $\{1, 2, \dots\}$ и, таким образом, исключается пустая выборка. Для указания того, что верхний предел не установлен и выборка никак не ограничена сверху, вместо целого положительного числа записывается символ дефиса.

Ниже приведены несколько примеров записи требования вопроса с указанием спецификации размера выборки в предлагаемой нотации.

$\binom{1}{1} cd$ — требует, чтобы выборка состояла точно из одного элемента субъекта.

$\binom{-}{1} cd$ — требует, чтобы выборка состояла не менее чем из одного элемента субъекта.

$\binom{-}{3} cd$ — требует, чтобы выборка состояла не менее чем из трёх элементов субъекта.

$\binom{-}{v} cd$ — требует, чтобы выборка состояла не менее чем из v элементов субъекта.

$\binom{5}{3} cd$ — требует, чтобы в ответ перешло не менее трёх, но не более пяти элементов субъекта.

$\begin{pmatrix} u \\ v \end{pmatrix} cd$ — требует, чтобы выборка состояла не менее чем из v , но не более чем из u элементов субъекта.

Рассмотрим несколько примеров вопросов, иллюстрирующих нотацию для представления спецификации размера выборки.

Вопрос

Что является примером простого числа, находящегося между 10 и 20?

можно представить при помощи интеррогативной формулы, с учётом только спецификации размера выборки в виде

$$Intr = ? \begin{pmatrix} 1 \\ 1 \end{pmatrix} (Cx // Ax), \quad (1.14)$$

где:

Cx — категорное условие: x является простым числом;

Ax — матрица: x находится между 10 и 20.

Спецификация размера выборки в интеррогативной формуле (1.14) требует, чтобы в ответ перешёл точно один элемент субъекта. Поэтому синтаксически корректные ответы, соответствующие формуле (1.14), имеют вид

11 находится между 10 и 20

13 находится между 10 и 20

и так далее.

Ответ

$(11 \text{ находится между } 10 \text{ и } 20) \wedge (13 \text{ находится между } 10 \text{ и } 20)$

является синтаксически некорректным, поскольку не соответствует спецификации выбора размера выборки, утверждающей, что в ответ должен перейти точно один элемент субъекта.

Ответ

14 находится между 10 и 20

также является синтаксически некорректным, поскольку не соответствует категорному условию. Число 14 не является простым.

Вопрос

Курение табака – это порок, добродетель, причуда, сумасбродство или панацея от всех бед?

можно представить при помощи интеррогативной формулы следующим образом:

$$Intr = ? \left(\begin{array}{c} - \\ 1 \end{array} c d \right) (A_1, A_2, A_3, A_4, A_5), \quad (1.15)$$

где:

A_1 – утверждение: Курение табака – это порок;

A_2 – утверждение: Курение табака – это добродетель;

A_3 – утверждение: Курение табака – это причуда;

A_4 – утверждение: Курение табака – это сумасбродство;

A_5 – утверждение: Курение табака – это панацея от всех бед.

Синтаксически корректные ответы, соответствующие интеррогату (1.15), имеют вид

Курение табака – это порок

Курение табака – это добродетель

(Курение табака – это порок) \wedge (Курение табака – это причуда)

(Курение табака – это добродетель) \wedge (Курение табака – это сумасбродство)

(Курение табака – это добродетель) \wedge (Курение табака – это сумасбродство) \wedge (Курение табака – это панацея от всех бед)

и так далее.

Вопрос

Какие есть несколько городов с населением большим, чем население Бостона?

представим при помощи следующей интеррогативной формулы:

$$Intr = ? \binom{5}{3} (Cx // Ax), \quad (1.16)$$

где:

Cx — категорное условие: x является городом;

Ax — матрица: население x больше, чем население Бостона.

Спецификация размера выборки в формуле (1.16) требует, чтобы в ответ перешло не менее трёх, но не более пяти элементов субъекта. В этом примере при помощи спецификации размера выборки в интеррогативной формуле снимается неопределённость естественно-языкового вопроса по отношению к количеству городов, которые должен назвать реактивный агент. В оригинальной естественно-языковой формулировке это количество задано неопределённо в виде словосочетания «а few cities», а в интеррогативе — это более определённое число «не менее трёх, но не более пяти».

Отметим, что все, рассмотренные выше примеры вопросов порождали не один, а *несколько возможных синтаксически корректных ответов*. Если согласиться с тем, что истинность или ложность ответа это относительная категория, которая определяется по отношению к системе знаний и цели реактивного агента, *то все ответы из множества возможных синтаксически корректных ответов следует считать истинными*.

1.2.4. Спецификация требования полноты

Подмножество элементов субъекта, переходящих в ответ, детерминируется при помощи спецификации размера выборки

указанием нижнего и верхнего пределов количества элементов этого подмножества. Часто этого достаточно для конструирования того *ответа, который ожидает активный агент*. Однако в некоторых случаях, такой способ специфицирования подмножества элементов субъекта, переходящих в ответ, не достаточен для конструирования ответа, ожидаемого активным агентом, и требует уточнения. Рассмотрим два примера естественно-языковых вопросов, имеющих тождественные субъекты и спецификации размера выборки, но предполагающие различные ответы.

- (1) Какие простые числа находятся между 10 и 20?
- (2) Каковы некоторые из простых чисел, находящиеся между 10 и 20?

Активный агент, задавая первый вопрос, *ожидает получить полный список* простых чисел, находящихся между числами 10 и 20. И только такой ответ является для него истинным. Если реактивный агент в своём ответе перечислит только часть простых чисел, то такой ответ активный агент воспримет как ложный. Ответ на второй вопрос будет истинным даже, если он содержит *не полный список* простых чисел.

Если ограничиться использованием только спецификации размера выборки в требовании вопроса, то оба вопроса могут быть представлены при помощи одной и той же интеррогативной формулы

$$Intr = ? \begin{pmatrix} - \\ 1 \end{pmatrix} (Cx // Ax), \quad (1.17)$$

где:

Cx — категорное условие: x является натуральным числом;

Ax — матрица: x является простым числом, находящимся между 10 и 20.

Синтаксически корректные ответы, соответствующие интеррогативной формуле (1.17), имеют вид

11 находится между 10 и 20,
 (17 находится между 10 и 20) \wedge (19 находится между
 10 и 20)
 и так далее.

Однако, несмотря на то, что приведенные примеры ответов соответствуют интеррогативной формуле (1.17) и поэтому синтаксически корректны для обоих вопросов, они являются истинными только для второго вопроса и ложными — для первого.

Рассмотренный пример показывает, что для различения первого и второго вопросов, требование вопроса должно включать не только спецификацию размера выборки, но и спецификацию требования полноты. Поскольку интеррогативная формула (1.17) позволяет генерировать и синтаксически корректные, и истинные ответы для второго вопроса, то спецификацией требования полноты должен быть снабжён интеррогатив, моделирующий первый вопрос.

Таким образом, можем сделать предположение, что в тех случаях, когда спецификация размера выборки позволяет порождать одновременно и синтаксически корректные, и истинные ответы, спецификация требования полноты не нужна. В случае, когда спецификация размера выборки позволяет порождать синтаксически корректные, но ложные ответы, она должна быть дополнена спецификацией требования полноты.

Вопрос

Какие простые числа находятся между 10 и 20?

предполагает, что в ответе должен быть весь список простых чисел, выбранных из множества целых чисел, находящихся в диапазоне между 10 и 20. Поэтому требование этого вопроса должно включать спецификацию требования максимальной полноты (*maximum completeness-claim specification*). В некоторых случаях необходимо специфицировать степень полноты. Например, вопрос

Каковы имена 5% секретарей в списке сотрудников?

предполагает, что ответ должен содержать имена точно 5% от всего списка секретарей. Степень полноты в естественно-языковом вопросе специфицируется при помощи словосочетаний-кванторов: все (all); все, кроме одного (all but one); 5%, большинство (most) и т.д.

Белнап и Стил определяют в общем случае спецификацию требования полноты следующим образом. Требование полноты прямого ответа — это требование того, насколько полным должно быть подмножество истинных элементов субъекта, переходящих в ответ по отношению ко всему множеству истинных элементов субъекта. Однако вариант эротетической логики, предлагаемый Белнапом и Стилом, ограничивается только формальным представлением спецификации требования максимальной полноты. При записи интеррогативной формулы для указания того, что требование вопроса включает спецификацию максимальной полноты, используется символ квантора всеобщности, который записывается на месте символа c :

$$\text{Intr} =?(s\forall d)\sigma \quad (1.18)$$

Если требование вопроса не включает спецификацию требования полноты, то вместо символа c записывается символ дефиса и интеррогативная формула принимает вид

$$\text{Intr} =?(s-d)\sigma. \quad (1.19)$$

Основываясь на предположении, что спецификация размера выборки s и спецификация требования максимальной полноты независимы, Белнап и Стил вводят классификацию требований вопроса путём комбинации значений этих спецификаций. Предлагается четыре класса требований вопроса:

- 1) требование одного элемента субъекта (single-example request);

- 2) требование нескольких элементов субъекта (some-examples request);
- 3) требование уникального элемента субъекта (unique alternative request);
- 4) требование полного списка элементов субъекта (complete list request).

Приведенная классификация требований вопроса иллюстрируется вопросами, имеющими один и тот же субъект:

$$\sigma = (Cx // Ax), \quad (1.20)$$

где:

Cx — категорное условие: x является натуральным числом;

Ax — матрица: x является простым числом между 10 и 20.

Ниже приведены примеры, иллюстрирующие предлагаемую классификацию требований вопроса. Каждый пример представлен интеррогативной формулой вопроса и соответствующей естественно-языковой формулировкой.

Вопрос с требованием одного элемента

$$Single - Example - Intr = ? \begin{pmatrix} 1 \\ 1 \end{pmatrix} - d \sigma. \quad (1.21)$$

Что является примером простого числа, находящегося между 10 и 20?

Вопрос с требованием нескольких элементов

$$Some - Examples - Intr = ? \begin{pmatrix} - \\ 1 \end{pmatrix} - d \sigma. \quad (1.22)$$

Каковы некоторые из простых чисел, находящиеся между 10 и 20?

Вопрос с требованием уникального элемента субъекта

$$Unique - Alternative - Intr = ? \left(\begin{array}{c} 1 \\ \forall d \\ 1 \end{array} \right) \sigma. \quad (1.23)$$

Какое простое число находится между 10 и 20?

Вопрос с требованием полного списка элементов субъекта

$$Complete - List - Intr = ? \left(\begin{array}{c} - \\ \forall d \\ 1 \end{array} \right) \sigma. \quad (1.24)$$

Какие простые числа находятся между 10 и 20?

1.2.5. Спецификация требования различения

В ряде случаев при формировании ответов на какой-вопрос, кроме размера выборки и требования полноты, необходимо учитывать и требование различения. Покажем это на примере. Пусть вопрос

Каковы, по крайней мере, пять примеров простых чисел?

моделируется интеррогативной формулой

$$Intr = ? \left(\begin{array}{c} - \\ 5 - d \end{array} \right) (Cx // Ax), \quad (1.25)$$

где:

Cx — категорное условие: x является натуральным числом;

Ax — матрица: x является простым числом.

Если в ответе на этот вопрос перечислим:

2, 3, 5, 7, и VII,

то такой ответ будет синтаксически корректным и вполне соответствовать интеррогативу (1.25). Однако его нельзя считать истинным ответом, который ожидает активный агент, поскольку в нём, по сути, перечислены только четыре простых числа, ибо 7 и VII — это различные имена одной и той же сущности — числа «семь».

Для того, чтобы гарантировать и синтаксически корректный, и истинный (ожидаемый) ответ на приведенный вопрос, необходимо в требовании вопроса конкретизировать спецификацию требования различения d таким образом, чтобы исключить использование в ответе различных имён для одной и той же сущности.

Спецификация требования различения не является обязательным компонентом требования вопроса. Она не нужна в тех случаях, когда сущности субъекта не имеют нескольких имён и синтаксически корректный ответ всегда является истинным. Например, спецификация требования различения точно не нужна для вопроса, спецификация размера выборки которого требует, чтобы выборка состояла точно из одного элемента субъекта. Спецификация требования различения также не нужна для ли-вопросов, поскольку предполагается, что субъект ли-вопроса всегда состоит из разноимённых сущностей.

Ясно, что не существует градаций для значений спецификации требования различения. Требование различения или есть, или его нет. Поэтому предлагается два способа записи интеррогативной формулы для случая, когда спецификация требования различения отсутствует и присутствует

$$Intr = ?(sc -) \sigma, \quad (1.26)$$

$$Intr = ?(sc \neq) \sigma, \quad (1.27)$$

Если требование вопроса не включает спецификацию требования различения, то в интеррогативной формуле вместо символа d записывается дефис, а если требование вопроса включает требование различения, то в интеррогативной формуле вместо символа d записывается знак неравенства.

Для того чтобы интеррогативная формула приведенного выше вопроса

Каковы, по крайней мере, пять примеров простых чисел?

порождала и синтаксически корректные, и истинные ответы, она должна быть записана в виде

$$Intr = ? \left(\begin{array}{c} - \\ 5 \end{array} - \neq \right) (Cx // Ax), \quad (1.28)$$

где:

Cx — категорное условие: x является натуральным числом;

Ax — матрица: x является простым числом.

Рассмотренные случаи целесообразности использования спецификации требования различения в комбинации с другими спецификациями требования вопроса позволяют расширить классификацию вопросов, введенную в подразделе 1.2.4. Ниже приведен расширенный список классов вопросов и соответствующие им интеррогативные формулы.

Вопрос с требованием одного произвольного истинного элемента субъекта

$$Single - Example - Intr = ? \left(\begin{array}{c} 1 \\ 1 \end{array} -- \right) \sigma. \quad (1.29)$$

Вопрос с требованием нескольких произвольных истинных элементов субъекта (разрешается использование различных имён для одной и той же сущности)

$$Some - Examples - Intr = ? \left(\begin{array}{c} - \\ 1 \end{array} -- \right) \sigma. \quad (1.30)$$

Вопрос с требованием нескольких произвольных истинных элементов субъекта (запрещается использование различных имён для одной и той же сущности)

$$\textit{Some - Distinct - Examples - Intr} = ? \left(\begin{array}{c} - \\ 1 \end{array} - \neq \right) \sigma. \quad (1.31)$$

Вопрос с требованием уникального истинного элемента субъекта

$$\textit{Unique - Alternative - Intr} = ? \left(\begin{array}{c} 1 \\ 1 \end{array} \forall - \right) \sigma. \quad (1.32)$$

Вопрос с требованием полного списка истинных элементов субъекта (разрешается использование различных имён для одной и той же сущности)

$$\textit{Complete - List - Intr} = ? \left(\begin{array}{c} - \\ 1 \end{array} \forall - \right) \sigma. \quad (1.33)$$

Вопрос с требованием полного списка истинных уникальных элементов субъекта (запрещается использование различных имён для одной и той же сущности)

$$\textit{Complete - Distinct - List - Intr} = ? \left(\begin{array}{c} - \\ 1 \end{array} \forall \neq \right) \sigma. \quad (1.34)$$

1.2.6. Дескрипторные вопросы

Кроме ли- и какой-вопросов Белнап и Стил рассматривают и другие классы вопросов. В настоящем обзоре теории Белнапа и Стила ограничимся кратким рассмотрением класса вопросов, названных *дескрипторными вопросами* (description-questions). Отметим, что для всех вопросов, входящих в класс вопросов-дескрипций, неизменной остаётся базовая логическая модель в виде ассоциированных субъекта и требования.

Используя терминологию логики предикатов первого порядка, субъект какой-вопроса можно рассматривать как постулируемую *открытую формулу* (матрица, или предложение с переменными) и ассоциированное с ней *множество ожиданий* (*desiderata*) в виде закрытых термов (имён или констант). Элементы субъекта какой-вопроса формируются подстановкой ожиданий в открытую формулу.

С точки зрения Белнапа и Стила субъект дескрипторного вопроса в этом смысле имеет структуру, противоположную какой-вопросу и определяется как некоторый закрытый терм, с которым ассоциировано множество ожиданий в виде открытых формул, называемых *дескрипторами* (*descriptors*). Подстановкой терма в дескрипторы формируются элементы субъекта. Например, вопрос

Какого цвета Том?

постулирует терм Том и ассоциированное с ним множество ожиданий, которые можно представить в виде следующих дескрипторов:

x — белый,
 x — красный
и так далее.

Для формального представления субъекта дескрипторного вопроса предлагается следующая нотация:

$$Des(Hx / /b), \quad (1.35)$$

где:

Hx — открытая формула, при помощи которой можно порождать множество дескрипторов (в англоязычном варианте книги названо прилагательным *determinable*);

b — закрытый терм (имя).

Для приведенного примера вопроса о цвете Тома субъект в нотации (1.35) можно записать в виде

$$Des(x \text{ is colored} // Tom). \quad (1.36)$$

Субъект (1.36) и порождает элементы

Том — белый,
Том — красный
и так далее.

В классе дескрипторных вопросов выделяются подклассы, названные:

- кто-вопросы (who-questions);
- как-вопросы (how-questions).

Примером дескрипторного кто-вопроса является вопрос

Кто такой был Скотт?

Интеррогативная формула этого вопроса может быть записана в виде

$$Who-Int = ? \left(\begin{array}{c} - \\ 1 \end{array} - \neq \right) Des(Hx // b), \quad (1.37)$$

где:

Hx — открытая формула: *x* был индивидом, представляющим интерес для истории;

b — закрытый терм: Скотт.

Открытая формула *Hx* в интеррогативе (1.37) может порождать следующие элементы субъекта:

Скотт был писателем, жившим в XVIII веке,
Скотт был тайным любовником мадам Помпадур
и так далее.

Примером дескрипторного как-вопроса является вопрос

Как растёт ваш сад?

Интеррогативную формулу этого вопроса запишем в виде.

$$\text{How-Int} = ? \left(\begin{array}{c} 1 \\ - \\ 1 \end{array} \right) \text{Des}(Hx // b), \quad (1.38)$$

где:

Hx — открытая формула: *x* растёт;

b — закрытый терм: ваш сад.

Открытая формула *Hx* в интеррогативе (1.38) порождает следующие элементы субъекта:

ваш сад растёт хорошо,
ваш сад растёт быстро,
ваш сад растёт не очень хорошо
и так далее.

2. ВОПРОСНО-ОТВЕТНЫЕ ТРАНЗАКЦИИ

Материал, изложенный в данном разделе, базируется на некоторых упрощающих предположениях относительно естественного диалога. Первое из этих упрощающих предположений заключается в том, что вместо полномасштабного естественного диалога рассматривается его вариант, называемый *вопросно-ответным диалогом*. [17,18].

Вопросно-ответный диалог является одним из типов естественного диалога, которым могут оперировать искусственные диалоговые агенты, демонстрирующие антропоморфное диалоговое поведение. Транзакции вопросно-ответного диалога ограничены вопросно-ответными парами, и поэтому внешне вопросно-ответный диалог не настолько разнообразен, как ничем не ограниченный естественный диалог, однако он может достигать те же цели и решать те же проблемы, что и полномасштабный естественный диалог. В вопросно-ответном диалоге информационные послышки активного агента имеют стопроцентную степень интеррогативности и часто (но не всегда) представляют собой вопросы в лингвистическом смысле слова, а информационные послышки реактивного агента — ответы на эти вопросы. Вопросно-ответный диалог — это шаг в направлении упрощения естественного

диалога, который, однако, представляется целесообразным. Во-первых, полномасштабный естественный диалог, как правило, можно конвертировать в вопросно-ответный диалог, сохранив при этом исходные цель и метод. Во-вторых, исследование вопросно-ответного диалога может базироваться на более или менее формализованных теориях. В-третьих, результаты, полученные при исследовании вопросно-ответного диалога, могут быть хорошей основой для их обобщения на полномасштабный естественный диалог.

В вопросно-ответном диалоге знания, запрашиваемые активным агентом и возвращаемые реактивным агентом, являются декларативными. Поэтому модель структуры вопросно-ответной транзакции в контексте представления знаний должна основываться на некоторой приемлемой концепции представления декларативных знаний. Хотя система декларативных знаний человека, по всей видимости, когерентна и подобна ткани, в которой близлежащие части сцеплены между собой, вопросы и ответы оперируют с отдельными небольшими частями или порциями системы декларативных знаний человека. Поэтому при дальнейшем изложении нам понадобятся средства представления только порции декларативных знаний, ассоциированных с вопросно-ответной транзакцией.

В подразделе 1.2 рассматривалась логическая структура отдельной вопросно-ответной транзакции, представленной в виде вопросно-ответной пары. В отдельной вопросно-ответной паре существует очевидная логическая связь между вопросом и ответом. Формально-логическим моделям этой связи, по сути, посвящен весь подраздел 1.2. Протоколы вопросно-ответных диалогов показывают, что кроме логической связи внутри отдельной транзакции существуют *логические связи между самими транзакциями*. Именно это отличает диалог от других видов интерактивного взаимодействия, например от последовательности изолированных друг от друга транзакций с базой данных. В настоящем разделе внимание фокусируется на вопросно-ответном диалоге и на структуре вопросно-ответной транзакции с точки зрения обмена знаниями в пределах

транзакции. Предложен спектр формальных моделей, ориентированных на представление внутренней логической структуры вопросно-ответной транзакции в контексте представления знаний. Отличительной особенностью этих моделей является их ориентация на лингво-независимые сущности представления декларативных знаний, ассоциированных с элементами транзакции. Теоретической основой такого представления знаний является Язык Тернарного Описания Уёмова [13–16].

2.1. Вопросно-ответные транзакции в контексте представления знаний

Диалог является одним из важнейших феноменов ментальной и интеллектуальной деятельности человека и внешне представляется протоколом в виде цепи взаимосвязанных информационных транзакций, которые формируются в процессе обмена знаниями между агентами диалогового взаимодействия. Целенаправленность — одна из фундаментальных характеристик диалога, и из этой целенаправленности необходимо следует логическая зависимость не только между отдельными компонентами диалоговой транзакции, но и между смежными транзакциями. Анализ практически любой когнитивной активности человека обнаруживает, что в её основе лежит диалоговый процесс. Так, рассуждения или умозаключения, как в широком, так и узком понимании, являются, по сути, «разговором с самим собой» или внутренним диалогом, в котором оба диалоговых агента представлены одной и той же личностью.

Диалоговая транзакция служит средством обмена знаниями между диалоговыми агентами, поэтому можем рассматривать диалоговый процесс как процесс, основанный на знаниях, а диалоговую систему как систему, основанную на знаниях. Подавляющее большинство исследователей явно или неявно предполагают, что знания в диалоговой транзакции представлены в виде пропозиций и поэтому (1) в соответствующих

публикациях превалируют вербальные примеры транзакций и их элементов; (2) для формального представления пропозиций чаще всего используется нотация логики первого порядка [11,19]. Такая ориентация на пропозициональное понимание и моделирование диалоговых транзакций определяется традиционным способом представления знаний в этой области и, с нашей точки зрения, не должна рассматриваться как единственно возможная.

Исследования, имеющие отношение к естественно-языковому и вопросно-ответному взаимодействию, можно обнаружить в различных и часто не очень близких научных областях. Перечень, возможно неполный, включает: искусственный интеллект, когнитивную психологию, эротетическую логику и эпистемологию.

Одно из наиболее известных подтверждений важности диалога, которое находим в искусственном интеллекте, является тест, предложенный Тьюрингом для операционной верификации интеллектуальности технической системы, который полностью базируется на диалоговом процессе между человеком и технической системой [20]. Современное изложение искусственного интеллекта, интегрирующее все или, по крайней мере, многие идеи в логически связанный предмет, основано на концепции интеллектуальных агентов [21]. Интеллектуальный агент является также диалоговым агентом. Он не может существовать вне проблемной среды, а транзакции между проблемной средой и интеллектуальным агентом определяют поведение последнего.

Теория когнитивного развития Пиаже, которую обычно относят к области когнитивной психологии и эпистемологии, неявно предполагает, что движущая сила развития схем — это диалоговый процесс [22]. Если согласиться с идеей о том, что человеческое существо появляется на свет с некоторым количеством врожденных схем, необходимых для его выживания в начальный период жизни, а приобретенные знания определяются последующим развитием и ростом этих схем, то методом, осуществляющим эволюцию врожденных схем, явля-

ется диалоговый процесс. В области когнитивной психологии можно обнаружить ещё несколько теорий и моделей, вносящих свой вклад в понимание сущности диалога. Например, циклическая модель восприятия, предложенная Найссером, — это, по сути, модель диалогового процесса, если источник потока воспринимаемых сенсорных событий интерпретировать как один из партнёров диалога [23].

Поскольку диалоговая транзакция, в какой бы форме она не была представлена, служит для обмена знаниями между диалоговыми агентами и транспортным средством этого обмена является вопросно-ответная пара, то результаты, полученные в эротетической логике (изложенные, например, в [6]) применимы для исследования и моделирования логической структуры вопросно-ответной диалоговой транзакции.

Эпистемология имеет непосредственное отношение к пониманию сущности диалогового процесса, поскольку эта наука направлена на изучение и моделирование тех видов знаний, которые циркулируют в пределах диалоговой транзакции и существенны для диалоговых агентов. Особый интерес представляет версия, предложенная Хинтикка и названная им Сократическая эпистемология, в которой ключевую роль играет вопросно-ответный диалог [24].

Большое количество журнальных публикаций, в которых встречаются словосочетания «диалоговый процесс» и «диалоговый агент», отражают сугубо прагматическое отношение к изучению диалога. Как правило, целью этих работ является не понимание феномена диалогового взаимодействия, а построение различных искусственных систем, способных играть роль вербального интерфейса между конечным пользователем и прикладной компьютерной программой. Например, поиск в Web порождает огромное количество ссылок на Web-страницы, включающие словосочетание «диалоговый агент» и описывающие решение тех или иных практических проблем. Диалоговые агенты, описанные в этих публикациях, проектируются исходя из потребности реализации коммуникативных функций и их сложно отнести к системам

«думающим как человек» согласно классификации Рассела и Норвига [21].

Формальное моделирование диалогового процесса должно начинаться с детального изучения протоколов естественных диалогов во всём их многообразии. В тех случаях, когда модели диалоговых агентов основаны на фундаментальных принципах человеческого диалога, можно ожидать, что искусственные диалоговые агенты наследуют богатые когнитивные возможности естественного диалогового процесса. Искусственный диалоговый агент только тогда может считаться таковым, когда он в состоянии поддерживать диалог, уровень сложности которого мало чем отличается от уровня сложности естественного диалога.

Цель подраздела — синтез формальных моделей диалоговых транзакций, направленных, главным образом, на более полное понимание логической сущности диалога, но и достаточно определенных, чтобы служить основой для проектирования искусственных диалоговых агентов. В изложенном материале используется введенный в первом разделе онтологический базис, который используется для синтеза формальных моделей классов декларативных знаний, ассоциированных с вопросно-ответными парами.

2.1.1. Язык Тернарного Описания Уёмова

Особенностью предлагаемого подхода к моделированию диалоговой транзакции является фокусировка внимания на том, что главным предназначением диалоговой транзакции является оперирование декларативными знаниями реактивного агента. Модели, которые будем конструировать, моделируют возможные варианты логической структуры вопросно-ответной транзакции в контексте представления декларативных знаний, ассоциированных с транзакцией. Моделирование будет ограничено вопросами типа «поисковое предписание», что является ещё одним упрощением полномасштабного естественного диалога.

В эпистемологии и эротетической логике декларативные знания, участвующие в вопросно-ответной транзакции, представляются в виде пропозиций, и, следовательно, элементарная порция декларативных знаний рассматривается как предложение. Общепринятый способ формализации такого представления знаний заключается в трансляции естественно-языковых предложений в предложения, записанные в онтологическом базисе и нотации логики первого порядка. Например, класс рассматриваемых в эпистемологии что-знаний [4] в общем случае может быть представлен предложением

$$(\exists x)a \text{ знает, что } (b = x).$$

Моделирование вопросно-ответной транзакции в контексте представления декларативных знаний может осуществляться средствами с более универсальной онтологией, чем формальная система логики первого порядка. Под универсальной онтологией здесь понимается онтология, основанная не на лингво-зависимых переменных (как в логике первого порядка), а на лингво-независимых фундаментальных сущностях декларативных знаний. В этом смысле привлекательным является Язык Тернарного Описания (ЯТО), первоначально предложенный Уёмовым для формального описания параметрического варианта общей теории систем, а затем развитый им и его учениками в самостоятельную неклассическую логику [13–16].

Исходной сущностью в ЯТО является *объект*, в самом общем смысле этого слова. В зависимости от места в структуре знаний объект может существовать в одной из трёх форм: *объект-вещь*, *объект-свойство* и *объект-отношение*.

Категории «вещь» и «свойство» имеют в ЯТО традиционный смысл, принятый в классической логике, а смысл категории «отношение» отличается от традиционного. Общепринято использовать понятие «отношение» для обозначения взаимовлияния нескольких вещей, т.е. говорить об отношении между вещами. Например, между женщиной и ребёнком может существовать отношение «отцовство». В контексте ЯТО под отношением по-

нимается то, что образует вещь, т.е. говорится об отношении на вещи или отношении, имеющем место в вещи. Иными словами, отношение в ЯТО — это в каком-то смысле другое наименование внутренней логической организации вещи.

Бинарная ассоциация объекта-вещи с объектом-свойством порождает два прототипа для представления сущностей в ЯТО:

1. Наименование первого прототипа: *«вещь, обладающая свойством»*, а формальная нотация имеет вид $(*)^*$.
2. Наименование второго прототипа: *«свойство, присущее вещи»*, а формальная нотация имеет вид $(**)^*$.

Бинарная ассоциация объекта-вещи с объектом-отношением порождает ещё два прототипа представления сущностей в ЯТО:

3. Наименование третьего прототипа: *«вещь, в которой имеет место отношение»*, а формальная нотация имеет вид $*(*)$.
4. Наименование четвёртого прототипа: *«отношение, имеющее место в вещи»*, а формальная нотация имеет вид $**(**)$.

Для формального представления сущностей в ЯТО используется специальная скобочная нотация. Символ «звёздочка» указывает на знакоместо для символа объекта. Символ объекта-вещи записывается внутри скобок, символ объекта-свойства — справа от скобок, а символ объекта-отношения — слева от скобок.

Ассоциация объекта-вещи с объектом-свойством или с объектом-отношением имеет направление. Если символ объекта-вещи заключён в обычные (одинарные) круглые скобки, то это означает, что ассоциация направлена от объекта-вещи к объекту-свойству или объекту-отношению. Словами это можно выразить как *«вещь, обладающая свойством»*, или *«вещь, в которой имеет место отношение»*. Ассиметричная (двойная) круглая скобка означает, что ассоциация направлена от объекта-свойства или объекта-отношения к объекту-вещи. Словесная формулировка: *«свойство, присущее вещи»*, или *«отношение, имеющее место в вещи»*.

Отметим, что в публикациях, посвящённых ЯТО, знакоместо внутри скобок используется только для символа объекта-вещи. Это ограничение исключает из набора прототипов следующие два: «отношение, обладающее свойством» и «свойство, присущее отношению».

Объект в зависимости от степени неопределённости знаний о нём существует в одной из трёх форм:

- *определённый объект* (в знакоместо прототипа подставляется символ t),
- *неопределённый объект* (в знакоместо прототипа подставляется символ a),
- *произвольный объект* (в знакоместо прототипа подставляется символ A).

Категории «вещь, свойство и отношение», а также «определённость, неопределённость и произвольность» независимы и формируют девять классов объектов: (1) определённый объект-вещь, (2) неопределённый объект-вещь, (3) произвольный объект-вещь, (4) определённый объект-свойство, (5) неопределённый объект-свойство, (6) произвольный объект-свойство, (7) определённый объект-отношение, (8) неопределённый объект-отношение, (9) произвольный объект-отношение.

Подставляя в прототипы представления сущностей в ЯТО вместо символа звёздочка символы t , a или A , получим множество возможных моделей порций декларативных знаний в онтологическом базисе ЯТО. Ниже приведены несколько моделей.

Модель

$$(t)a \quad (2.1)$$

представляет знания о некоторой определённой вещи, которая обладает каким-то (неопределённым) свойством. Направление ассоциации от вещи к свойству.

Модель

$$(t))a \quad (2.2)$$

представляет знания о каком-то (неопределённом) свойстве, которое присуще некоторой определённой вещи. Направление ассоциации от свойства к вещи.

Модель

$$(a)A \quad (2.3)$$

представляет знания о некотором произвольном свойстве, которое присуще какой-то (неопределённой) вещи. Направление ассоциации от свойства к вещи.

Модель

$$a(t) \quad (2.4)$$

представляет знания о некоторой определённой вещи, в которой имеет место какое-то (неопределённое) отношение. Направление ассоциации от вещи к отношению.

Поскольку есть четыре прототипа представления сущностей в ЯТО, каждый из которых заполняется объектами из девяти классов, то общее количество моделей элементарных порций декларативных знаний в онтологическом базисе ЯТО равно тридцати шести.

2.1.2. Представление декларативных знаний в вопросно-ответной транзакции

Структуру вопроса в вопросно-ответной транзакции в контексте ЯТО-представления порций декларативных знаний будем рассматривать как развитие идеи базовой интеррогативной формулы Белнапа и Стила, постулирующей, что внутренняя логическая структура вопроса детерминируется двумя компонентами: субъектом вопроса и требованием вопроса. Запишем вопрос в виде интеррогативной формулы

$$Que = ? \rho \sigma, \quad (2.5)$$

где:

ρ — требование вопроса;

σ — субъект вопроса;

? — признак интеррогатива.

Важным элементом философии Белнапа и Стила, лежащим в основе их теории, является концепция субъекта вопроса. Из этой концепции следует, что те вопросы, модель которых включает субъект, например вопросы типа «поисковое предписание», не только определяют возможные ответы, но и *содержат их*. Это означает, что субъект вопроса можно мыслить как *неопределённый ответ*.

Субъект вопроса является ключевым компонентом вопросно-ответной транзакции. Реактивный агент при конструировании ответа занимается, по сути, тем, что трансформирует неопределённые знания, на которые указывает субъект вопроса, в определённые знания, ассоциированные с ответом. Иными словами, *ментальные усилия реактивного агента направлены на понижение степени неопределённости знаний субъекта вопроса до уровня приемлемого для ответа*.

Будем моделировать порции декларативных знаний, на которые указывает субъект вопроса, следующими альтернативными моделями:

$$K_{subj} = (t)a, \quad (2.6)$$

$$K_{subj} = (a)t, \quad (2.7)$$

$$K_{subj} = a(t), \quad (2.8)$$

$$K_{subj} = t((a)). \quad (2.9)$$

Модели (2.6)–(2.9) получены из четырёх рассмотренных ранее ЯТО-прототипов порций декларативных знаний для случая ассоциации определённого и неопределённого объектов. Имеется несколько причин выбора моделей (2.6)–

(2.9) для представления порции декларативных знаний. Во-первых, субъекты всех примеров вопросов, описанных в [4,6,11], могут быть представлены одной из моделей (2.6)–(2.9). Во-вторых, это именно те модели (из тридцати шести возможных), которые соответствуют идее о том, что реактивный агент при конструировании ответа трансформирует неопределённые знания субъекта вопроса в определённые знания ответа. Один из объектов в моделях (2.6)–(2.9) является неопределённым, и более конкретные знания именно о нём активный агент ожидает получить от реактивного агента.

Поясним модели (2.6)–(2.9) на содержательном уровне.

- Модель $K_{subj} = (t)a$ представляет знания о конкретной определённой вещи, которой присуще некоторое неопределённое свойство. Вопрос с таким субъектом формируется активным агентом в том случае, когда он хочет узнать, *какими свойствами обладает данная вещь*.
- Модель $K_{subj} = (a)t$ представляет знания об определённом свойстве, которое присуще некоторой неопределённой вещи. Вопрос с таким субъектом формируется активным агентом в том случае, когда он хочет узнать, *каким вещам присуще данное свойство*.
- Модель $K_{subj} = a(t)$ представляет знания о конкретной определённой вещи, в которой имеет место некоторое неопределённое отношение. Вопрос с таким субъектом формируется активным агентом в том случае, когда он хочет узнать, *какое отношение имеет место в данной вещи*.
- Модель $K_{subj} = t(a)$ представляет знания о конкретном определённом отношении, которое имеет место в некоторой неопределённой вещи. Вопрос с таким субъектом формируется активным агентом в том случае, когда он хочет узнать, *в каких вещах имеет место данное отношение*.

Недостатком моделей (2.6)–(2.9) с точки зрения инжиниринга программных систем, основанных на знаниях, является их плохая приспособленность для отображения в релевантные структуры данных. Эти модели могли бы быть полезными для целей разработки программных систем в случае, если бы мы нашли способ их трансформации в типы или структуры данных современных систем программирования. Речь идёт, в первую очередь, о даталогической интерпретации неопределённого объекта.

Одной из возможных и, наверное, простейшей *даталогической интерпретацией неопределённости является множественность*. Неопределённый объект можно понимать как множество определённых объектов, а мощность этого множества как меру его неопределённости. Тогда понижение степени неопределённости объекта эквивалентно уменьшению мощности соответствующего множества. Неопределённый объект превращается в абсолютно определённый в том случае, когда мощность множества становится равной единице, или когда множество представлено одним конкретным определённым объектом.

С учётом такой интерпретации неопределённости можем заменить неопределённые объекты списками определённых объектов и представить модели (2.6)–(2.9) следующим образом:

$$K_{subj} = \langle \text{объект-вещь} \rangle \{ \text{список объектов-свойств} \}, \quad (2.10)$$

$$K_{subj} = \langle \text{объект-свойство} \rangle \{ \text{список объектов-вещей} \}, \quad (2.11)$$

$$K_{subj} = \langle \text{объект-вещь} \rangle \{ \text{список объектов-отношений} \}, \quad (2.12)$$

$$K_{subj} = \langle \text{объект-отношение} \rangle \{ \text{список объектов-вещей} \}. \quad (2.13)$$

Модели (2.10)–(2.13) — это *даталогические аналоги* моделей (2.6)–(2.9), которые представляют декларативные знания, передаваемые реактивному агенту посредством субъекта во-

проса. Две дополнительные модели, невыражаемые в нотации ЯТО (в силу того, что внутри скобок можем записывать только символ объекта-вещи), представлены формулами (2.14) и (2.15):

$$K_{subj} = \langle \text{объект-свойство} \rangle \quad (2.14)$$

$$\{ \text{список объектов-отношений} \},$$

$$K_{subj} = \langle \text{объект-отношение} \rangle \quad (2.15)$$

$$\{ \text{список объектов-свойств} \}.$$

Поскольку субъект вопроса является, по сути, ответом с некоторой степенью неопределённости и реактивный агент при конструировании ответа понижает степень неопределённости вопроса до уровня, приемлемого для ответа (который детерминруется требованием вопроса ρ в формуле (2.5)), модели порций декларативных знаний, ассоциированных с ответом, должны быть аналогичны моделям (2.10)–(2.15). Отличие моделей декларативных знаний, ассоциированных с ответом, заключается в мощности множества (списка) объектов. Модели декларативных знаний, ассоциированных с ответом, представим в виде

$$K_{ans} = \langle \text{вещь} \rangle \quad (2.16)$$

$$\text{ОБЛАДАЕТ СВОЙСТВАМИ} \{ \text{список свойств} \},$$

$$K_{ans} = \langle \text{свойство} \rangle \quad (2.17)$$

$$\text{ПРИСУЩЕ ВЕЩАМ} \{ \text{список вещей} \},$$

$$K_{ans} = B \langle \text{вещи} \rangle \quad (2.18)$$

$$\text{ИМЕЮТ МЕСТООТНОШЕНИЯ} \{ \text{список отношений} \},$$

$$K_{ans} = \langle \text{отношение} \rangle \quad (2.19)$$

$$\text{ИМЕЕТ МЕСТО В ВЕЩАХ} \{ \text{список вещей} \},$$

$$K_{ans} = \langle \text{свойство} \rangle \quad (2.20)$$

$$\text{ПРИСУЩЕ ОТНОШЕНИЯМ} \{ \text{список отношений} \},$$

$$K_{ans} = \langle \text{отношение} \rangle$$

$$\text{ИМЕЕТ МЕСТО В СВОЙСТВАХ} \{ \text{список свойств} \}. \quad (2.21)$$

Модели (2.16)–(2.21) представляют декларативные знания, ссылку на которые реактивный агент возвращает активному агенту при помощи ответа на вопрос с релевантным субъектом.

Модели (2.16)–(2.21) проиллюстрируем примерами вопросно-ответных транзакций, построенных на основе вопросов, которые приведены в [6]. Рассмотрим транзакцию

Активный агент: Какова температура замерзания воды по Фаренгейту при нормальных условиях?
Реактивный агент: При нормальных условиях температура замерзания воды равна 32 °F.

Субъект вопроса в приведенной транзакции обеспечивает доступ к декларативным знаниям, представимым моделью (2.10):

$$K_{subj} = \langle \text{вода при нормальных условиях} \rangle$$

$$\{ \text{температуры замерзания жидкостей по Фаренгейту} \}$$

Модель представляет знания относительно данной определённой вещи с именем «вода при нормальных условиях», которая обладает некоторым неопределённым свойством — какой-то температурой замерзания. Активный агент, формулируя вопрос с таким субъектом, хочет перейти от неопределённости к определённости и получить доступ к знаниям реактивного агента относительно конкретного свойства «воды при нормальных условиях». Неопределённость в знаниях, на которые указывает субъект, выражается множеством, состоящим из конечного, но достаточно большого количества альтернатив. Знания, ассоциированные с ответом реактивного агента, вполне определённы и представимы моделью (2.16):

$K_{ans} = \langle \text{вода при нормальных условиях} \rangle$
ОБЛАДАЕТ СВОЙСТВОМ {замерзнуть при 32° F}

Рассмотрим вопрос.

Какие простые числа лежат между 10 и 20?

Субъект этого вопроса обеспечивает доступ к декларативным знаниям, которые представимы моделью (2.11)

$K_{subj} = \langle \text{быть простым числом} \rangle$
{10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20}

Модель представляет знания относительно данного определённого свойства с именем «быть простым числом», которое присуще некоторой неопределённой вещи — какому-то числу в диапазоне от 10 до 20. Активный агент, формулируя вопрос с таким субъектом, хочет перейти от неопределённости к определённости и узнать у реактивного агента, каким конкретным вещам-числам в диапазоне от 10 до 20 присуще свойство «быть простым числом». Неопределённость в знаниях, на которые указывает субъект, выражается множеством, состоящим из одиннадцати элементов. Знания, ассоциированные с ответом реактивного агента представимы моделью (2.17).

$K_{ans} = \langle \text{быть простым числом} \rangle$
ПРИСУЩЕ ЧИСЛАМ {11, 13, 17, 19}

В качестве следующего примера рассмотрим одну из транзакций из диалога «Протагор», приведенного в начальной части первого раздела. Диалог «Протагор» не является исключительно вопросно-ответным, поскольку информационные послышки активного агента не во всех транзакциях представлены вопросами в грамматическом смысле. Однако ряд транзакций этого диалога являются вопросно-ответными. Рассмотрим первую из них

Сократ: Так мы идём платить деньги Протагору, потому что он софист?

Гиппократ: Конечно.

Субъект вопроса Сократа в этой транзакции обеспечивает доступ к декларативным знаниям, представимым моделью (2.10). Модель представляет знания о данной определённой вещи «Сократ-и-Гиппократ», которая обладает некоторым неопределённым свойством — готовностью или неготовностью платить деньги Протагору только на том основании, что он софист. Сократ, формулируя вопрос с таким субъектом, хочет перейти от неопределённости к определённости и узнать у Гиппократа каким конкретным свойством обладает Сократ-и-Гиппократ. Сократ ожидает получить некоторый ответ и, исходя из своих ожиданий, он конструирует субъект вопроса таким образом, чтобы неопределённость в знаниях, на которые он указывает, представляется списком, состоящим только из двух альтернатив:

$$K_{subj} = \langle \text{Сократ-и-Гиппократ} \rangle \\ \{ \langle \text{давать деньги софисту Протагору} \rangle, \\ \langle \text{не давать деньги софисту Протагору} \rangle \}$$

Ответ Гиппократа является указателем на порцию декларативных знаний, выражаемых моделью (2.16):

$$K_{ans} = \langle \text{Сократ-и-Гиппократ} \rangle \\ \text{ОБЛАДАЕТ СВОЙСТВОМ} \{ \langle \text{давать деньги софисту Протагору} \rangle \}$$

В завершении отметим, что модели логической структуры транзакции вопросно-ответного диалога целесообразно строить исходя из структуры знаний, ассоциированных с транзакцией. В этом случае модели не только способствуют лучшему пониманию внутренней сущности диалоговой транзакции, но и имеют прагматическую ценность, поскольку могут служить теоретической основой для проектирования базы знаний диа-

логового агента. Мы проиллюстрировали применимость Языка Тернарного Описания Уёмова для представления порций декларативных знаний, ассоциированных с вопросно-ответной транзакцией. Модели, представленные формулами (2.10)–(2.21), вскрывают внутреннюю логическую сущность вопросно-ответной транзакции в контексте представления знаний.

2.2. Эпистемологический взгляд на знания в вопросно-ответной транзакции

Эпистемология — это теория знаний, а слова «теория знаний» и «эпистемология» являются взаимозаменяемыми. Эпистемология — это раздел философии, поэтому эпистемологическая теория знаний рассматривает природу, типы и источники знаний в рамках философских традиций.

Знания субъекта, как правило, рассматриваются изолированно (вне контекста диалогового процесса) и представляются в виде высказываний (пропозиций), которые могут быть истинными или ложными. Общепринятым наименованием таких знаний является «знание-that». Примером знаний-that некоторого субъекта может быть высказывание «душа питается знаниями».

Существует специальное обозначение для знаний-that. Знания субъектом (человеком, личностью) S пропозиции p обозначаются в виде K_{sp} . Таким образом, приведенный пример знаний о том, что душа питается знаниями, в эпистемологической интерпретации K_{sp} означает, что некоторый субъект S знает что p , а p записывается как «душа питается знаниями». В контексте диалогового процесса K_{sp} — это знания реактивного диалогового агента, доступ к которым хочет получить активный агент. Так, Сократ (субъект S) обладает знаниями K_{sp} о том, чем питается душа. Однако Гиппократ этого не знает. Поэтому в диалоге с Сократом Гиппократ, находясь в роли активного агента, хочет получить доступ к этим знаниям. С этой

целью он конструирует вопрос: «Но чем же питается душа, Сократ?»

Перечислим кратко вопросы, которыми занимается классическая эпистемология.

Одним из центральных вопросов эпистемологии является вопрос: «*При каких условиях субъект знает некоторую пропозицию?*». Классический ответ на этот вопрос формулируется в виде словесной формулы, утверждающей, что *знанием является «обоснованное и истинное убеждение»* (Justified True Belief — JTB). Согласно этой словесной формуле необходимые и достаточные условия K_{sp} могут быть сформулированы следующим образом:

- 1) p является истинной (p is true),
- 2) S убеждён, что p (s believes that p), и
- 3) S обоснованно убеждён, что p (s is justified in believing that p).

Приведенные необходимые и достаточные условия знаний субъекта разделяются многими эпистемологами, однако являются предметом дискуссии. Геттиер показал, что можно найти примеры, которые иллюстрируют, что формула JTB не содержит достаточное условие знаний субъекта о данной пропозиции [24]. Примеры, которые Геттиер привёл в своей статье, иллюстрируют случаи, когда вопросы «Знает ли S , что p ?» и «Обладает ли S обоснованным убеждением, что p ?» имеют различные ответы и, следовательно, формула JTB некорректна и должна быть модифицирована.

Дискуссия о том при каких условиях субъект обладает знаниями о p является, по сути, дискуссией о том, при каких условиях убеждение субъекта в p обосновано. Рассуждения о том, что есть обоснованием убеждения субъекта S в пропозиции p является стандартной темой эпистемологических исследований. Эта тема раскрывается в рамках поиска ответов на следующие вопросы. При каких условиях субъект рационален в своём убеждении в нечто? Когда субъект уверен в чём-то? Когда субъект точно знает, что что-то имеет место? Когда что-то является сомнительным для субъекта? При каких услови-

ях убеждение может быть ложным с точки зрения субъекта? Когда убеждение адекватно подтверждается доказательством? При каких условиях убеждение должно подтверждаться доказательством для того, чтобы быть рациональным?

Другим центральным вопросом эпистемологии является вопрос: «*Что мы знаем?*» или «*Каков объём или масштаб наших знаний?*», который тесно связан с упомянутым выше вопросом об условиях того, что субъект S обладает знаниями, что p . Пессимистический взгляд на масштаб наших знаний, называемый *скептицизмом* (skepticism), заключается в том, что мы знаем меньше, чем мы думаем, что знаем, а радикально пессимистический взгляд — в том, что мы знаем очень мало, или даже не знаем ничего [25]. Таким образом, скептицизм отрицает утверждение о том, что достоверные знания возможны. Несмотря на свою радикальность, такой скептический взгляд на масштаб наших знаний разделяют многие философы и эпистемологи. Странники скептицизма часто строят умозаключения на утверждении о том, что убеждениям, как правило, не хватает определённости.

Эпистемологов интересуют вопросы о том, *когда убеждения становятся обоснованными и когда субъект знает то, в чём он убеждён*. Существует два основных подхода к экспликации ответов на эти вопросы, называемые *интернализм* (internalism) и *экстернализм* (externalism). Согласно интернализму ответы на эти вопросы зависят в первую очередь от факторов, внутренних по отношению к субъекту, убеждённому в некоторой пропозиции. Большинство интерналистов считают, что вопрос о том, истинно ли убеждение, релевантен вопросу о том, составляет ли оно знание. В вопросе знания интернализм обычно придерживается точки зрения, согласно которой только внутренние факторы имеют отношение к тому, являются ли истинные убеждения знаниями. Согласно экстернализму вопросы знания и/или обоснования убеждения зависят исключительно от внешних факторов или от того, насколько надёжен механизм, при помощи которого субъект пришёл к убеждению. Можно представить себе двух субъектов, чья ментальная жизнь

идентична по отношению к восприятию вещей (внутренние факторы), но чьи убеждения различны.

Важным вопросом, рассматриваемым эпистемологами, является *вопрос о структуре знаний* (структуре обоснования убеждений). Основные подходы к экспликации ответов на этот вопрос называются *фундаментализм* (foundationalism) и *когерентизм* (coherentism). Убеждение *A* может составлять часть знаний потому, что оно основано на другом убеждении — *B*, которого придерживается субъект и которое представляет собой свидетельство в пользу *A*. Однако такая схема работает, если убеждение *B* является обоснованным убеждением. Но как убеждение *B* могло быть обосновано? Возможно, оно базировалось на другом обоснованном убеждении субъекта *C*. Где и каким образом может завершиться процесс обоснования убеждений на основе других убеждений? Если потребовать от всех наших обоснованных убеждений, чтобы они основывались на других обоснованных убеждениях, которых мы придерживаемся, и не содержали циклов обоснования (убеждение *A* базируется на убеждении *B*, а убеждение *B* — на убеждении *C*, и т.д. пока не встречается убеждение, которое обосновывается, поскольку базируется на *A*), и допустить, что цепи обоснования убеждений не могут быть бесконечно длинными, то придём к скептическому выводу, что ни одно из наших убеждений не обосновано.

Фундаменталисты отрицают, что все наши убеждения должны основываться на других наших убеждениях, чтобы быть обоснованными. По мнению фундаменталистов, некоторые из наших убеждений являются базовыми (properly basic). Эти базовые убеждения являются знаниями вне зависимости от того, основаны они на других обоснованных убеждениях или нет. Базовые знания образуют фундамент, на котором строятся все остальные наши обоснованные убеждения. Сам по себе фундаментализм — это просто приверженность этой базовой структурной картине.

Когерентисты считают, что только доказуемые отношения между убеждениями субъекта могут обосновывать эти убежде-

ния, поэтому отрицают наличие базовых убеждений, которые не требуют обоснования. Когерентисты допускают наличие циклов в обосновании убеждений (если убеждение A базируется на убеждении B , то убеждение B , опосредованно, через цепь других убеждений, может базироваться на убеждении A). Согласно философии когерентистов, наши убеждения обосновываются не тем, насколько хорошо они согласуются с базовыми убеждениями, а тем, насколько хорошо они согласуются друг с другом.

Кроме поиска ответов на перечисленные классические вопросы эпистемологии усилия ряда эпистемологов направлены на изучение знаний, ассоциированных с вопросно-ответными транзакциями. В ряде публикаций анализу подвергаются знания-*wh*, определение и классификация которых приведена в подразделе 1.1.7, а также знания-*that*, рассматриваемые как утверждения о знаниях-*wh* (субъект S знает, что p , где p — ответ на соответствующий *wh*-вопрос) [4, 26–30]. Поскольку знания-*wh* порождаются вопросно-ответной транзакцией, в которой информационной посылкой активного агента является один из *wh*-вопросов, то они являются вопросно-релятивными и должны рассматриваться как знания субъектом S (реактивным агентом) пропозиции p , являющейся ответом на соответствующий вопрос.

Как было отмечено в подразделе 1.1.7, в эпистемологии не используются понятия декларативных и процедурных знаний, принятые в когнитивной психологии. Однако из контекста публикаций следует, что знания-*that* и знания-*wh* можно трактовать как *декларативные*, поскольку они представляются в виде пропозиций. Таким образом, в общем случае, декларативные знания так, как их определяет, например, Андерсен [31], — это и есть знания-*that* или знания-*wh* так, как их понимает, например, Шаффер [4]. На рис. 2.1 в виде UML диаграммы классов представлена классификация декларативных знаний, ассоциированных с вопросно-ответной транзакцией, принятая в эпистемологии.

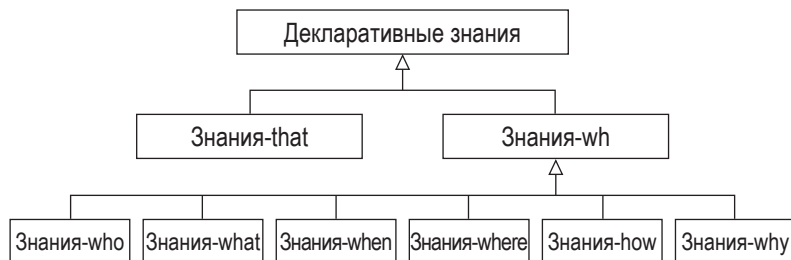


Рис. 2.1. Эпистемологическая классификация декларативных знаний, ассоциированных с вопросно-ответной транзакцией

Различие в объёме понятий «декларативные знания» и «знания-that или знания-wh» заключается в следующем. Когда когнитивные психологи говорят о декларативных знаниях, они не всегда трактуют их лингвистически и представляют в виде пропозиций. Андерсен, например, использует термин «chunk» (кусочек) [31]. Способ представления chunk может быть различным. Например, можем мыслить chunk в виде фрейма Марвина Минского [32]. Когда эпистемологи говорят о декларативных знаниях в виде знаний-wh или знаний-that, они всегда предполагают лингвистическое (пропозициональное) представление знаний. Об этом свидетельствует формула K_{sp} « S знает что p ».

2.2.1. Конвергентные знания

Шаффер показал, что *упрощённое понимание знаний-wh*, как знаний-that (только в виде K_{sp} и без учёта wh-вопроса, породившего p) недостаточно, поскольку создаёт проблему конвергентных знаний [4]. *Конвергентными вопросами* Шаффер называет различные wh-вопросы с одинаковыми истинными ответами, а утверждения о наличии знаний-wh, связанные с конвергентными вопросами, *конвергентными утверждениями о знаниях-wh*.

Проблема конвергентных знаний заключается в том, что согласно упрощённому пониманию знаний-wh все конвергентные

утверждения о знаниях-wh должны рассматриваться как эквивалентные, несмотря на очевидную неправомерность такого тезиса.

Шаффер иллюстрирует проблему конвергентных знаний несколькими примерами.

Представим себе, что бывший президент США Джордж Буш выступает на телевидении. В период президентства Джорджа Буша в США были популярны комедиант Вилл Ферелл, который очень натурально имитировал Джорджа Буша, и поп-дива Жанет Джексон, которую трудно спутать с президентом Бушем. В этом случае конвергентными являются следующие wh-вопросы:

*Que*₁ Кто на телевидении: Буш или Жанет Джексон?

*Que*₂ Кто на телевидении: Буш или Вилл Ферелл?

Приведенные вопросы являются конвергентными, поскольку имеют одинаковый истинный ответ:

*P*₁ Буш на телевидении

Конвергентными утверждениями о знаниях-wh, связанными с приведенными вопросами, являются утверждения:

*K*_{*Que*₁} Я знаю, кто на телевидении, Буш или Жанет Джексон.

*K*_{*Que*₂} Я знаю, кто на телевидении, Буш или Вилл Ферелл.

Из упрощенного понимания знаний-wh следует, что эти утверждения эквивалентны, поскольку каждое из них эквивалентно следующему:

*K*_{*sp*₁} Я знаю, что Буш на телевидении.

Таким образом, если принять упрощенное понимание знаний-wh, то *K*_{*Que*₁} и *K*_{*Que*₂} эквивалентны, откуда можно сделать

вывод, что все конвергентные утверждения о знаниях-wh эквивалентны. Однако, ясно, что такой вывод неверен. Вопросы *Que₁* и *Que₂* это вопросы типа задача. Для ответа на них реактивный агент должен решать задачу распознавания образов. Получение знаний о том, Буш или Жанет Джексон на телевидении, является относительно лёгкой задачей распознавания. Поэтому вопрос о том, Буш это или Жанет Джексон, является относительно лёгким вопросом. Фактически любой человек (с нормальным зрением и минимальной культурой) может отличить Буша от Жанет Джексон. В противоположность этому получение знаний о том Буш или Вилл Ферелл на телевидении, является относительно трудной задачей распознавания. Если подражание достаточно успешно, то вопрос о том, Буш это или Вилл Ферелл, может быть весьма трудным. Возможно, что только жена Джорджа Буша будет в состоянии указать на различия. Таким образом, кто-либо может точно знать Буш или Жанет Джексон на телевидении, но не знать Буш или Вилл Ферелл на телевидении.

Рассмотрим ещё один пример, иллюстрирующий проблему конвергентных знаний. Представим себе, что в саду на дереве сидит щегол. В этом случае конвергентными являются следующие вопросы:

Que₃ Кто в саду, щегол или ворон?

Que₄ Кто в саду, щегол или канарейка?

Que₅ Где щегол, в саду или поблизости?

Приведенные вопросы являются конвергентными, поскольку имеют одинаковый истинный ответ:

p₂ В саду щегол

Конвергентными утверждениями о знаниях-wh, связанными с приведенными вопросами, являются утверждения:

K_{Que₃} Я знаю кто в саду, щегол или ворон.

K_{Que_4} Я знаю кто в саду, щегол или канарейка.

K_{Que_5} Я знаю где щегол, в саду или поблизости.

Упрощённое понимание знаний-wh требует считать, что все эти три утверждения эквивалентны, поскольку каждое из них эквивалентно утверждению:

K_{sp_2} Я знаю, что в саду щегол.

Однако K_{Que_3} , K_{Que_4} , и K_{Que_5} явно не эквивалентны. Получение знаний о том, находится ли в саду щегол или ворон, является относительно лёгкой задачей «птицераспознавания». Фактически любой человек (с нормальным зрением и минимальной компетенцией) может знать, находится ли в саду щегол или ворон. Однако получение знаний о том, находится ли в саду щегол или канарейка — это более тяжёлая задача. Возможно, что только эксперт будет в состоянии описать различие. А знание того, находится ли щегол в саду или поблизости, является задачей, имеющей отношение, скорее, к ландшафту, а не к птицам. Возможно, что только владелец дома сможет ответить на этот вопрос. Таким образом, кто-либо может точно знать, находится ли в саду щегол или ворон, но не знать, находится ли в саду щегол или канарейка и/или не знать, находится ли щегол в саду или поблизости.

Утверждения $K_{Que_1} - K_{Que_5}$ можно рассмотреть как утверждения о знаниях ответов на ли-вопросы в теории Белнапа и Стила [6]. Субъект ли-вопроса содержит список пропозиций, являющихся возможными ответами на вопрос (смотри подраздел 1.2.2). Следовательно, вопрос «Кто на телевидении?», когда он задан в контексте, означающем, что на телевидении Буш или Джексон, имеет субъект, состоящий из двух альтернатив: «Буш или Джексон на телевидении». Таким образом, утверждение K_{Que_1} в форме «Я знаю, кто на телевидении» представляет собой утверждение, произнесённое в контексте, означающем, что на телевидении Буш или Джексон, а утверждение K_{Que_2} в форме «Я знаю, кто на телевидении» — утверждение, произне-

сённое в контексте, означающем, что на телевидении Буш или Ферелл. Проведенные рассуждения могут привести к заключению о том, что если имеется некоторый исходный ли-вопрос, то можно создать конвергентный ему ли-вопрос, замещая в субъекте исходного вопроса хотя бы одну ложную альтернативу на другую. Пусть, например, имеется некоторое утверждение о знании ответа на ли-вопрос, субъект которого содержит две альтернативы a и b . Пусть истинной альтернативой, переходящей в ответ, является альтернатива a , а ложной — альтернатива b . Запишем это в виде

$$K_{Q_i} S \text{ knows}(a \vee b)$$

$$p_i a.$$

Тогда согласно сделанному только что заключению конвергентное утверждение о знании ответа можно записать в виде

$$K_{Q_i} S \text{ knows}(a \vee c),$$

где c — ложная альтернатива, замещающая альтернативу b в утверждении K_{Q_i} .

Но *эта схема явно не правомерна*. Иногда замещение ложной альтернативы ответа, или добавление «коварной» альтернативы ответа, может превратить лёгкий ли-вопрос в сложный. Любой, кто разрабатывал письменные экзамены с множественным выбором, понимает эту возможность. Студенты уверенно отвечают на вопросы, только если альтернативы лёгкие. Следовательно, не все конвергентные утверждения знаний эквивалентны: это не тот случай, что если wh_1 -вопрос и wh_2 -вопрос имеют один и тот же истинный ответ, то субъект обладает знаниями- wh_1 тогда и только тогда, когда он обладает знаниями- wh_2 .

Таким образом, упрощённый взгляд на знание- wh является неверным, а знания-*that* не сводятся к знаниям- wh . Есть нечто большее в знании- wh , чем знание пропозиции, которая оказалась ответом.

2.2.2. Представление знаний-wh, включающее вопрос

Проблема конвергентных знаний имеет простое и естественное решение, которое заключается в дополнении формулы K_{sp} соответствующим wh-вопросом. Иными словами, утверждение о знании-wh должно включать вопрос и обозначать следующее: « S знает-wh тогда и только тогда, когда K_{spQ} , где Q является вопросом, порождающим истинный ответ P ». В этом случае исчезает проблема конвергентных знаний, а утверждение о знании-wh выражает декларативные знания, ассоциированные с вопросно-ответной транзакцией и являющиеся ответом реактивного агента на вопрос активного агента. Отметим, что формула K_{spQ} соответствует вопросно-релятивному пониманию знаний, которое может быть выражено следующим высказыванием: «Знать что-либо означает знать ответ на соответствующий вопрос». Рассмотрим несколько аргументов, подтверждающих необходимость использования K_{spQ} вместо K_{sp} в утверждениях о знаниях-wh.

Во-первых, решается проблема конвергентных вопросов. В том случае, когда Буш на телевидении, знания K_{Que_1} и K_{Que_2} появляются, как знания, выражающие различные пропозиции: K_{spQue_1} и K_{spQue_2} . Знать, что Буш или Жанет Джексон на телевидении, означает знать, что *Буш на телевидении*, как истинный ответ на вопрос *Буш или Жанет Джексон на телевидении?* В то время, как знать Буш или Вилл Ферелл на телевидении, означает знать, что *Буш на телевидении*, как истинный ответ на вопрос *Буш или Вилл Ферелл на телевидении*. Таким же образом, в случае щегла в саду, K_{Que_3} , K_{Que_4} и K_{Que_5} появляются как знания, выражающие различные пропозиции: K_{spQue_3} , K_{spQue_4} и K_{spQue_5} . Включение в формулу вопроса дифференцирует знания.

Во-вторых, формула K_{spQ} неявно предполагает, что реактивный агент обладает экспертизой, необходимой для решения задачи, ассоциированной с вопросом. Для реактивного агента обладание необходимой экспертизой является функцией

вопроса. Например, если имеет место вопрос Que_3 , то реактивный агент должен обладать экспертизой, необходимой для того, чтобы отличить щегла от ворона. Если имеет место вопрос Que_4 , то реактивный агент должен обладать экспертизой, необходимой для того, чтобы отличить щегла от канарейки. И если имеет место вопрос Que_5 , то реактивный агент должен обладать экспертизой, необходимой для того, чтобы отличить сад от окрестностей.

В третьих, имеется несколько лингвистических фактов, подтверждающих необходимость включение вопроса в формулу знаний-wh. Первый лингвистический факт — это справедливость *экзистенциального обобщения* (*existential generalization*) вопроса. Схему экзистенциального обобщения представим в виде следующего правила вывода:

S знает-wh

Существует вопрос, ответ на который знает *S*.

Например, если я знаю где припаркована машина, то из этого следует, что имеется вопрос, ответ на который я знаю (а именно, вопрос о том, где припаркована машина). Точно также, если я знаю сколько времени, то из этого следует, что существует вопрос, ответ на который я знаю (а именно вопрос о том, который сейчас час). Справедливость экзистенциального обобщения, таким образом, указывает на присутствие вопроса в K_{spQ} , как той сущности, которая была обобщена.

Второй лингвистический факт в пользу включения вопроса в формулу K_{spQ} — это справедливость *замещения* (*substitution*) для вопроса. Схему замещения представим в виде следующего правила вывода:

S знает-wh, wh-вопрос является ADJ вопросом

S знает ответ на ADJ вопрос.

ADJ вопросом (adjective question) английского языка называется один из вопросов: «Which one?», «What kind?», «How

many?» «Whose?»). Рассмотрим пример, иллюстрирующий приведенную схему замещения. Если я знаю, когда родился Наполеон, и если вопрос о том, когда родился Наполеон, относится к историческим вопросам, то из этого следует, что я знаю ответ на исторический вопрос. Справедливость замещения, таким образом, указывает на присутствие вопроса в K_{spQ} , как той сущности, которая была замещена.

Третий лингвистический факт в пользу включения вопроса в знания-wh — это возможность *анафорической ссылки* на вопрос. Например: «Я знаю, почему небо голубое. *Это* хитрый вопрос». Или: «Я теперь знаю, почему небо голубое. Я размышлял над *этим* годами». Естественно понимать слова «это» и «этим» как ссылки на вопрос о том, почему небо голубое. Возможность такого понимания указывает на присутствие вопроса в K_{spQ} как сущности, с которой связана анафора.

В заключении отметим, что эпистемологический взгляд на знания в вопросно-ответной транзакции фокусируется на понимании декларативных знаний реактивного агента диалога. Эти знания называются знания-wh и представимы формулой K_{spQ} , которая обозначает следующее: «Реактивный агент S обладает знаниями-wh тогда и только тогда, когда он знает пропозицию p , являющуюся истинным ответом на вопрос Q ». На рис. 2.2. приведена UML диаграмма классов, которая графически представляет формулу знаний-wh в виде K_{spQ} , как два отношения типа ассоциация между классами S , p и Q .

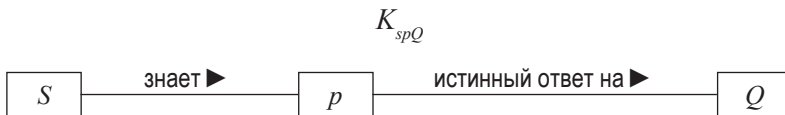


Рис. 2.2. Представление знаний-wh в виде отношений типа ассоциация между классами S , p и Q

3. ПРОБЛЕМНО-НЕЗАВИСИМЫЕ МОДЕЛИ ДИАЛОГА

Раздел посвящен в основном описанию ряда проблемно-независимых моделей диалога, которые могут генерировать протоколы вопросно-ответных диалоговых процессов. Модели инвариантны к проблемной области диалога и могут рассматриваться как формальное представление вопросно-ответного диалога. Формальное представление диалога, в отличие от содержательного, позволяет перейти от неопределённых онтологических моделей к более определённым концептам и отношениям, свойственным всему диалоговому процессу. Формальные модели диалога ценны также прикладным аспектом. Прикладная ценность модели подразумевает возможность её использования в качестве основы для разработки спецификаций программной системы, имитирующей поведение активного агента диалога. Такая программная система находится в постоянном общении с пользователем (реактивным агентом), которое для этого класса программных систем является жизненно важной частью её функционирования.

При разработке формальной модели диалога важным является её адекватность психологии диалога (или адекватность модели процессам восприятия и переработки информации человеком). В случае, когда в основу формальной модели ди-

алога положены «удачные» психологические модели, можно ожидать, что искусственные диалоговые агенты наследуют гибкость и универсальность системы восприятия и переработки информации человеком.

3.1. Когнитивные основы диалогового процесса

Среди множества теорий и моделей, предлагаемых когнитивной психологией и имеющих отношение к диалоговому процессу, остановимся на моделях, описывающих диалог на уровне восприятия, который не зависит от предметной области диалога и от характера проблемы, решаемой в процессе диалогового поведения. Таким образом, психологическая модель диалога, полезная в прикладном аспекте, должна описывать некоторую унифицированную когнитивную «машину диалога».

В современной когнитивной психологии, исповедующей принципы информационной переработки, можно выделить четыре когнитивные подсистемы, принимающие участие в диалоговом общении: (1) сенсорная подсистема, (2) подсистема перцепции (восприятия), (3) подсистема фокусировки внимания, (4) подсистема категоризации и запоминания [33, 34].

Экспериментальные данные показывают, что в перцепции участвуют два встречных ментальных процесса: восходящий процесс (от сенсорной системы к долговременной памяти) и нисходящий процесс (от долговременной памяти к периферии) [35]. Найссер предложил модель, интегрирующую отмеченные процессы и названную циклом перцепции [23]. Цикл Найссера является одной из ключевых моделей, объединяющей многие практические и теоретические результаты в области когнитивной психологии. Цикл проводит весьма условную границу между окружающей средой и её отображением в ментальную систему человека. С точки зрения Найссера, когнитивная структура, являющаяся ключевой в процессе перцепции — это *набор схем, ожидаемых быть воспринятыми, или гипотетический набор схем*. В этом и последующих разделах термин

«схема» понимается так, как это принято в когнитивной психологии: блок знаний, находящийся в долговременной памяти, отражающий объект внешнего мира и участвующий в перцепции. Схемы будем понимать как ментальное представление знаний о мире.

На рис. 3.1 приведено диаграмматическое представление цикла перцепции Найссера. Гипотетический набор схем совместно с перцептивным исследованием определяет информацию, которая будет воспринята и категоризована. Перцепция является конструктивным процессом в том смысле, что на каждом шаге перцепции конструируется очередной гипотетический набор схем. Гипотетический набор схем является частью когнитивной структуры, предложенной Толманом и названной им *когнитивная карта* [36]. Понятие «когнитивная карта» чаще всего используется при моделировании визуальной перцепции. В контексте перцептивного цикла Найссера когнитивную карту будем рассматривать, как ментальное отображение фрагмента окружающей среды.

Образование очередного гипотетического набора схем порождает процесс перцептивного исследования окружающей среды. Процесс перцептивного исследования предполагает перемещение сенсорных органов, головы и всего тела. Наиболее ярким примером перцептивного исследования является ощупывание предмета при его тактильном восприятии. Процесс перцептивного исследования завершается тем, что внимание фокусируется на конкретном сенсорном событии. Сенсорная система формирует информационный образ сенсорного события, который затем распознаётся и воспринимается в виде реальной схемы. В категоризации и восприятии сенсорного события участвует не вся когнитивная карта, а только гипотетический набор схем. Воспринятое сенсорное событие является причиной модификации гипотетического набора схем. Суть процесса модификации текущего гипотетического набора схем заключается в конструировании нового гипотетического набора, используемого на последующем цикле перцепции.



Рис. 3.1. Диаграмматическое представление перцептивного цикла Найссера

Отметим, что гипотетический набор схем формируется из ранее воспринятых и запомненных схем, составляющих когнитивную карту. Таким образом, цикл Найссера предполагает, что процесс перцепции детерминируется двумя факторами: текущим состоянием окружающей среды и прошлым перцептивным опытом человека.

Цикл Найссера хорошо моделирует процесс рутинной перцепции, под которым понимается процесс последовательной фокусировки сенсорной системы на внешних, не представляющих угрозу (не подозрительных) сенсорных событиях. Перцепция подозрительных сенсорных событий, очевидно, требует модификации цикла Найссера. Например, поскольку

«обработка» подозрительных сенсорных событий должна выполняться быстро, то очевидно, что появление подозрительного сенсорного события должно прерывать рутинную перцепцию на любой её стадии.

Диалоговый процесс аналогичен процессу перцептивного взаимодействия человека с окружающей средой. Отличие заключается в том, что в процессе диалога главные компоненты сенсорной системы человека — зрительный и слуховой анализаторы — подключены не к «естественной» окружающей среде, а к «искусственной», формируемой потоками зрительных и звуковых сенсорных событий, генерируемых противоположным агентом диалога. Таким образом, в диалоговом процессе реальная окружающая среда подменяется искусственной. Однако очевидно, что перцепция и последующая переработка как сенсорных событий, порождаемых искусственной средой (диалоговым агентом), так и сенсорных событий, генерируемых естественной окружающей средой, осуществляются по одним и тем же «правилам и законам» и одной и той же цепью подсистем переработки информации человеком.

Отмеченная аналогия цикла Найссера и диалогового процесса, а также независимость цикла Найссера от «предметной области» перцепции послужили причиной выбора цикла перцепции Найссера в качестве когнитивной основы проблемно-независимых формализованных моделей диалога, рассматриваемых в данном разделе. Используя онтологию диалогового процесса, введенную ранее, преобразуем перцептивный цикл Найссера в *когнитивный цикл вопросно-ответного диалогового процесса* [37–40].

Диаграмма, иллюстрирующая когнитивный цикл вопросно-ответного диалогового процесса, приведена на рис. 3.2. Этот цикл моделирует поведение активного агента в диалоге с фиксированными ролями. Фиксирование ролей агентов означает запрет на смену ролей агентов в ходе вопросно-ответного диалога. Один из агентов всегда задаёт вопросы, а другой — всегда отвечает на них. Вопросно-ответная транзакция формируется при однократном прохождении всех элементов

диалогового цикла. Будем считать для удобства, что диалоговый цикл начинается с момента, когда активный агент оперирует относительно небольшим набором ответов, названным *набором ожидаемых ответов*, а реальный ответ реактивного агента совпадает с одним из ответов из набора ожидаемых ответов активного агента. Набор ожидаемых ответов является частным случаем гипотетического набора схем цикла Найссера (см. рис. 3.1). Таким образом, одной из основных когнитивных структур, которой оперирует активный агент, является схема-ответ, или ментальное представление ответа.

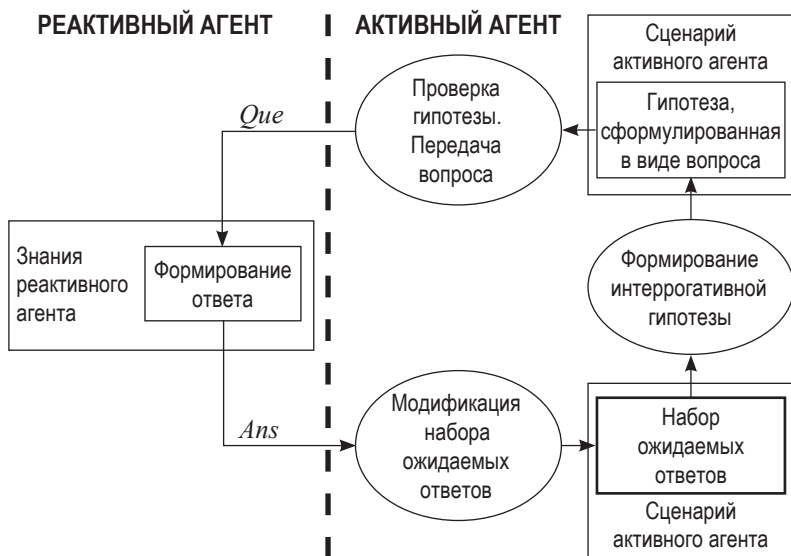


Рис. 3.2. Когнитивный цикл вопросно-ответного диалогового процесса

Набор ожидаемых ответов встроен в более обширную когнитивную структуру, названную *сценарием активного агента* и являющуюся аналогом когнитивной карты в цикле Найссера. Предполагается, что сценарий активного агента включает все необходимые для данного диалога схемы-ответы и что набор

ожидаемых ответов — это некоторое подмножество ответов из сценария, которое соответствует текущему циклу диалогового процесса.

После восприятия и распознавания реального ответа, полученного от реактивного агента, осуществляется модификация текущего набора ожидаемых ответов, целью которой является подготовка набора ожидаемых ответов для следующего цикла. В результате модификации формируется новый набор ожидаемых ответов, включающий ответ, ожидаемый на следующем цикле. В процессе модификации текущий набор ожидаемых ответов может быть заменён полностью или частично.

Восприятие и распознавание реального ответа порождает два процесса. Один из процессов — это описанный только что процесс модификации набора ожидаемых ответов, а другой — процесс генерации очередного вопроса. Последующий вопрос детерминирован текущим ответом, а то, какой вопрос будет сформирован активным агентом в следующем цикле, определяется ответом, полученным от реактивного агента в текущем цикле. *Для продолжения диалога* активному агенту на каждом диалоговом цикле (каждом шаге диалога) нужна новая порция знаний. Интеррогативная гипотеза — это гипотеза о том, как получить эти знания от реактивного агента при помощи вопроса.

Одним из отличительных свойств диалогового процесса является его целенаправленность. За внешним процессом обмена вопросами и ответами всегда стоит процесс достижения цели. По отношению к активному агенту под целью можно понимать или получение от реактивного агента некоторого факта в виде ожидаемого целевого ответа, (примером может служить допрос подозреваемого), или реализация диалогового процесса по некоторой целевой «траектории» (примером может служить обучающий диалог студента с компьютерным тьютором). Целенаправленный характер диалога подразумевает, что вопросы активного агента формируются не произвольным образом, а целенаправленно. Управлять этим процессом активный агент может только построением интеррогативной гипотезы, форму-

лировкой релевантного вопроса и испытанием гипотезы передачей её реактивному агенту.

Схема-вопрос, или ментальное представление структуры вопроса, является одной из основных когнитивных структур, которыми оперирует активный агент. Предполагается, что схемы-вопросы, необходимые для конкретного диалога, также встроены в сценарий активного агента.

Концепция когнитивной карты и её частного случая — сценария диалога — ясно указывают на то, что активный агент-человек не «вычисляет» очередной вопрос с помощью некоторого «супералгоритма», а, скорее, разыскивает его в памяти вопросов, используя в качестве *метода доступа к памяти вопросов* метод достижения цели диалога. Таким образом, *метод доступа к памяти вопросов является хранилищем метода достижения цели диалога, или диалоговым методом решения проблемы*. Концепты «память вопросов» и «диалоговый метод решения проблемы» входят в объем понятия «сценарий» активного агента и детализируют его структуру.

Как интроспекция, так и анализ реальных протоколов вопросно-ответных диалоговых процессов, обнаруживают, что имеет место неоднозначная связь между текущим ответом реактивного агента и последующим вопросом, генерируемым активным агентом. Иными словами, в различных диалоговых транзакциях на один и тот же ответ, полученный от реактивного агента, активный агент может генерировать различные вопросы. Поэтому при разработке диалоговых методов решения задач для моделирования отмеченной неоднозначности предлагается учитывать следующие три принципа.

Принцип «глубины диалога». Принцип «глубины диалога» означает, что диалоговый метод при формировании очередного вопроса должен учитывать как воспринятый ответ, так и индекс диалогового цикла. Здесь термин «индекс» используется как синоним термина имя, или идентификатор. В различных диалоговых циклах на один и тот же ответ, полученный от реактивного агента, могут быть сформированы разные вопросы.

Принцип «истории ответов». Принцип «истории ответов» означает, что диалоговый метод при формировании очередного вопроса должен учитывать как воспринятый ответ, так и историю воспринятых ответов. В одном и том же диалоговом цикле на один и тот же ответ, полученный от реактивного агента, могут быть сформированы разные вопросы в зависимости от того, какие ответы были получены в предыдущих циклах.

Принцип «истории вопросов». Принцип «истории вопросов» означает, что диалоговый метод при формировании очередного вопроса должен учитывать как воспринятый ответ, так и историю ранее сформированных вопросов. В одном и том же диалоговом цикле на один и тот же ответ могут быть сформированы разные вопросы в зависимости от того, какие вопросы были сформированы в предыдущих циклах.

Диалоговый цикл, приведенный на рис. 3.2, хорошо описывает *«гармоничный диалог»*, соответствующий циклу Найссера для случая рутинной перцепции. Под гармоничным диалогом будем понимать такой диалог, когда оба агента удовлетворены своими ролями и не хотят их изменять. Однако гармоничный диалог не всегда возможен. Каким образом можно учесть возможность нарушения гармоничного диалога и изменения текущих ролей диалоговых агентов на противоположные? Инициатором смены ролей обычно выступает реактивный агент, а признаком момента смены ролей является генерация и передача активному агенту информационного сообщения, имеющего статус вопроса. Поэтому одним из способов учёта возможности смены ролей может быть включение в список ожидаемых ответов «детектора вопроса». Детектор вопроса может быть включён в каждый набор ожидаемых ответов (тогда агенты будут иметь возможность сменить роли в любом диалоговом цикле) или только в некоторые наборы ожидаемых ответов (тогда роли агентов могут быть изменены только в некоторых, заранее определенных, диалоговых циклах).

В последующих подразделах рассмотрены формальные *проблемно-независимые модели вопросно-ответного диалога, которые являются моделями диалогового метода доступа к*

памяти вопросов, либо моделями диалогового метода решения проблем. Эти модели интерпретируют поведение активного агента вопросно-ответного диалога, как функционирование некоторой «машины диалога». Целью моделирования является получение структуры «машины диалога», инвариантной к цели и предметной области диалогового процесса. Модели, представленные в подразделах 3.2 и 3.3, явно учитывают только принцип «глубины диалога», который для случая моделирования метода доступа к памяти вопросов означает, что при формировании индекса следующего вопроса учитываются только воспринятый ответ и индекс текущего диалогового цикла.

3.2. Конечно-автоматная модель вопросно-ответного диалога

Конечно-автоматная модель диалога моделирует гармоничный диалог и базируется на двух предположениях: (1) проблема, решаемая в ходе вопросно-ответного диалога, может быть решена за конечное количество диалоговых циклов (шагов диалога), а к началу решения проблемы детерминированы все вопросы и ответы, необходимые для формирования диалоговых шагов; (2) модификация текущего набора ожидаемых ответов всегда порождает новый набор, отличающийся от текущего и имеющий уникальный индекс.

Конечно-автоматная модель диалога использует определение автомата Мили и в общем случае может быть описана следующей пятеркой [41]:

$$DiAM = (QUE, ANS, S, f, g). \quad (3.1)$$

В формуле (3.1) $DiAM$ обозначает диалоговый метод доступа к памяти вопросов.

$$QUE = \{Que_i\}; i = 1, \dots, k, \quad (3.2)$$

где:

QUE — множество индексов всех вопросов активного агента, которые, согласно предположению 1, необходимы и достаточны для решения проблемы (достижения цели диалога). Спецификации вопросов находятся в памяти вопросов.

$$ANS = \{Ans_i\}; i = 1, \dots, l, \quad (3.3)$$

где:

ANS — множество всех ответов реактивного агента, которые, согласно предположению 1, необходимы и достаточны для достижения цели диалога.

$$S = \{S_i\}; i = 1, \dots, m, \quad (3.4)$$

где:

S — множество стабильных состояний сценария диалога. Стабильное состояние — это, по сути, состояние ожидания активным агентом ответа от реактивного агента. Естественно предположить, что количество элементов множества S равно количеству диалоговых шагов, необходимых для решения проблемы при помощи $DIAM$.

$$f : S \times ANS \rightarrow S, \quad (3.5)$$

где:

f — функция перехода, переводящая сценарий в последующее стабильное состояние в зависимости от индекса текущего шага и ответа, полученного от реактивного агента.

$$g : S \times ANS \rightarrow QUE, \quad (3.6)$$

где:

g — функция выхода, детерминирующая индекс последующего вопроса активного агента в зависимости от индекса текущего шага и ответа, полученного от реактивного агента.

Функции перехода и выхода соответствуют классическому определению автомата Мили. При моделировании метода доступа к памяти вопросов более естественной является функция

шага. Поэтому в дальнейшем вместо двух отображений f и g будем пользоваться одним:

$$\varphi : S \times ANS \rightarrow S \times QUE, \quad (3.7)$$

где:

φ — функция шага детерминирует пару: индекс нового вопроса и индекс следующего шага в зависимости от текущего ответа и индекса текущего шага. При использовании функции шага конечно-автоматная модель диалогового процесса может быть записана в виде

$$DiAM = (QUE, ANS, S, \varphi). \quad (3.8)$$

При выполнении шага автомат, имитирующий цикл диалога, выполняет итерации, состоящие из следующей последовательности действий:

- воспринимается текущий ответ;
- распознаётся текущий ответ;
- детерминируется индекс следующего вопроса.

Если никак не структурировать множество ответов ANS , то это означает, что для распознавания текущего ответа диалоговый метод $DiAM$ должен на каждом шаге оперировать всем множеством ANS . Для обширных сценариев мощность этого множества может быть значительной. Для уменьшения мощности множества ответов, с которым на каждом шаге оперирует $DiAM$, введём в модель множество ответов, распознаваемых на i -м шаге. Это множество содержит только те

$$Ans \in ANS,$$

которые ожидаются на i -м шаге, необходимы для реализации диалогового метода и, следовательно, должны быть распознаемы. Все остальные $Ans \in ANS$ отнесём к классу ответов, не распознаваемых на i -м шаге.

Введём обозначения:

R^i — множество ответов, распознаваемых на i -м шаге;

NR^i — множество ответов, не распознаваемых на i -м шаге.

К множеству ответов, не распознаваемых на i -м шаге, относятся все ответы, принадлежащие множеству ANS и не принадлежащие множеству R^i , т.е.

$$NR^i = ANS - R^i. \quad (3.9)$$

Однако множество ответов, не распознаваемых на i -м шаге, необходимо рассматривать шире и понимать как множество любых мыслимых ответов реактивного агента, не принадлежащих множеству R^i , т.е.

$$NR^i = U - R^i, \quad (3.10)$$

где U — универсум ответов, т.е. множество ответов, которые может породить реактивный агент. Ясно, что $ANS \subseteq U$.

Естественно предположить, что появление любого ответа, принадлежащего множеству NR^i , приводит к тому, что $DiAM$ генерирует один и тот же индекс последующего вопроса. Ибо в противном случае это означало бы, что ответы относятся к классу распознаваемых ответов. Поэтому в функции шага φ для каждого шага диалога весь класс нераспознаваемых ответов следует моделировать одним нераспознаваемым ответом.

Проведенные рассуждения позволяют вернуться к циклу диалогового процесса, приведенному на рис. 3.2, и уточнить структуру набора ожидаемых ответов. Таким образом, на каждом шаге диалога набор ответов, которые ожидает активный агент, может включать ответы из следующих классов ответов:

- класс распознаваемых ответов;
- класс нераспознаваемых ответов (моделируется одним ответом);
- класс детекторов вопроса (для определения потребности в смене ролей).

Простой пример, рассмотренный ниже, преследует две цели: во-первых, оценить, насколько предложенная конечно-автоматная модель диалога (формулы (3.1)–(3.10)) полезна при конструировании диалогового метода доступа к памяти вопросов, и, во-вторых, обеспечить переход к последующей модели вопросно-ответного диалога в виде сети Петри. Пример включает несколько шагов.

На шаге 1 активный агент формирует вопрос, имеющий индекс Que_1 , и ожидает получить следующие ответы:

Ans_1 — запрос на получение твёрдой копии вопроса;

Ans_3 — ответ, подразумевающий продолжение диалога и формирование вопроса с индексом Que_3 ;

Ans_4 — ответ, подразумевающий продолжение диалога и формирование вопроса с индексом Que_4 ;

Ans_5 — нераспознаваемый ответ (любой ответ реактивного агента, отличный от Ans_1 , Ans_3 , Ans_4). Получение активным агентом ответа Ans_5 означает, что необходимо продолжить диалог и сформировать вопрос с индексом Que_2 .

На шаге 2 активный агент формирует вопрос, имеющий индекс Que_2 , представляющий собой переформулированный вопрос с индексом Que_1 . На этом шаге активный агент ожидает получить тот же набор ответов, что и на первом шаге, а также ответ Ans_2 , являющийся запросом на возвращение на шаг 1 и повторение вопроса с индексом Que_1 .

Метод доступа к памяти вопросов для приведенного примера вопросно-ответного диалога может быть описан следующим образом:

$$ANS = \{Ans_1, Ans_2, Ans_3, Ans_4, Ans_5\}, \quad (3.11)$$

$$QUE = \{Que_1, Que_2, Que_3, Que_4\}, \quad (3.12)$$

$$S = \{s_1, s_2, s_3, s_4\}, \quad (3.13)$$

$$\begin{aligned}
\varphi : (s_1, Ans_1) &\rightarrow (s_1, Que_1), \\
(s_1, Ans_3) &\rightarrow (s_3, Que_3), \\
(s_1, Ans_4) &\rightarrow (s_3, Que_4), \\
(s_1, Ans_5) &\rightarrow (s_3, Que_2), \\
(s_2, Ans_1) &\rightarrow (s_2, Que_2), \\
(s_2, Ans_2) &\rightarrow (s_1, Que_1), \\
(s_2, Ans_3) &\rightarrow (s_3, Que_3), \\
(s_2, Ans_4) &\rightarrow (s_4, Que_4), \\
(s_1, Ans_5) &\rightarrow (s_2, Que_2).
\end{aligned}
\tag{3.14}$$

Функция шага φ , представленная отображениями (3.14), с одной стороны, описывает «логику» доступа к памяти вопросов, а с другой — «логику» диалогового метода решения некоторой проблемы. Функцию шага удобно представить в табличной форме. В таблице на рис. 3.3 содержится та же информация, что и в отображениях (3.14).

Табличное представление функции шага удобно тем, что позволяет легко перейти к графическому представлению вопросно-ответного диалога в форме диаграммы состояний [41]. На рис. 3.4 приведено графическое представление функции шага диалога φ в виде диаграммы состояний. В этой диаграмме множество вершин соответствует множеству шагов диалога (или множеству состояний автомата), а каждая ветвь — одному из возможных вариантов развития диалога, определяемого ответом реактивного агента. Ветви помечены парами

$$Ans_i, Que_j,$$

где:

Ans_i — один из ответов, ожидаемых на данном шаге;

Que_j — индекс вопроса, генерируемого на данном шаге.

Как отмечено ранее, функция шага может рассматриваться как способ описания «логики» диалогового метода решения задачи. В этом смысле диаграмма состояний, пример которой

приведен на рис. 3.4, может рассматриваться как диаграмма, описывающая целенаправленное поведение активного агента.

Текущее состояние (текущий шаг)	Ожидаемые ответы реактивного агента	Индекс вопроса активного агента	Следующее состояние (следующий шаг)
s_1	Ans_1	Que_1	s_1
	Ans_3	Que_3	s_3
	Ans_4	Que_4	s_4
	Ans_5	Que_2	s_2
s_2	Ans_1	Que_2	s_2
	Ans_2	Que_1	s_1
	Ans_3	Que_3	s_3
	Ans_4	Que_4	s_4
	Ans_5	Que_2	s_2

Рис. 3.3. Табличное представление функции φ для примера, описанного в тексте

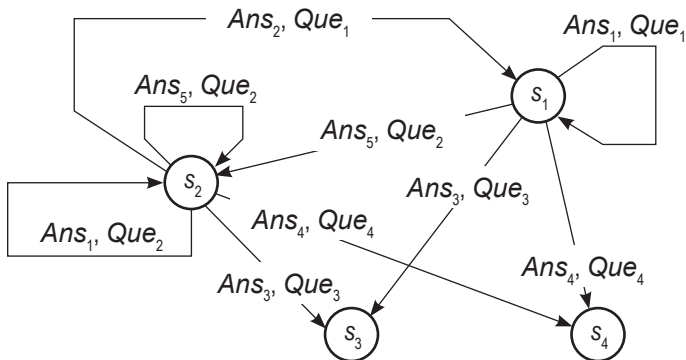


Рис. 3.4. Диаграмма состояний диалогового метода доступа для примера, описанного в тексте

3.2.1. Проблема моделирования неискренности реактивного агента

В понятийном базисе, которым мы пользовались ранее, и в конечно-автоматной модели диалога отсутствует компонент, явно учитывающий такое антропоморфное свойство реального вопросно-ответного диалога, как *неискренность реактивного агента*. В реальном вопросно-ответном диалоге ответы реактивного агента не всегда правдивы, особенно в тех случаях, когда вопросы активного агента касаются нежелательных, запретных или опасных тем. Как правило, одной из целей реактивного агента в ходе диалогового процесса является уклонение от таких тем и сокрытие информации.

Моделирование неискренности реактивного агента даёт возможность строить диалоговые методы, позволяющие решить проблему даже в том случае, когда части ответов реактивного агента нельзя доверять. Это делает модель более адекватной вопросно-ответным диалогам, наблюдаемым в реальной жизни.

Для учёта неискренности реактивного агента предлагается использовать основные положения теории нечётких множеств [42, 43]. Отметим, что прикладная ценность теории нечётких множеств заключается не столько в предлагаемом ею определении и формализме нечёткого множества, сколько в *интерпретации функции степени принадлежности*.

Теория нечётких множеств является следующим (вслед за формальной логикой) шагом в борьбе с нечёткостью и расплывчатостью человеческих знаний [42]. Авторы теории нечётких множеств предложили в качестве основного её постулата отказаться от принципа классической теории множеств, согласно которому элемент может либо принадлежать, либо не принадлежать множеству. Вместо этого для каждого множества вводится характеристическая функция, которая, будучи непрерывной и принимая значения на сегменте $[0,1]$, характеризует степень принадлежности элемента множеству.

Нечёткое множество определяется как

$$A = \{x, \mu_A(x)\}$$

и представляет собой совокупность упорядоченных наборов пар элементов $x \in X$ и соответствующих им степеней принадлежности

$$\mu_A(x): X \rightarrow [0,1].$$

Трансформируем предложенную детерминированную автоматную модель *DiAM* в нечёткую, используя приведенное определение нечёткого множества.

Модель включает три множества — потенциальных кандидатов на использование понятия нечёткости. Это множество индексов всех вопросов активного агента *QUE*, множество всех ответов реактивного агента *ANS* и множество стабильных состояний сценария диалога, или множество шагов диалога *S*.

Стабильное состояние s_i представляет собой, по сути, индекс (имя) шага диалога, и поскольку бессмысленно говорить о степени принадлежности к имени, то множество состояний *S* является чётким.

В детерминированной модели $QUE = \{Que_i\}; i = 1, \dots, k$ представляет собой множество индексов всех вопросов, необходимых и достаточных для достижения цели/целей диалога. С точки зрения модели, вопросы Que_i являются равноценными и одинаково приближают диалог к его цели. Однако в реальном диалоговом процессе дело обстоит иначе, и тот или иной вопрос может (в зависимости от степени прямоты вопроса) способствовать достижению цели диалога в большей или меньшей степени.

Следовательно, представляется целесообразным следующая интерпретация понятия нечёткости по отношению к множеству вопросов. Определим множество вопросов *QUE* как нечёткое следующим образом:

$$QUE = \{Que_j, \mu_Q(Que_j)\}; j = 1, \dots, k, \quad (3.15)$$

$$\mu_Q : QUE \rightarrow [0, 1], \quad (3.16)$$

где:

$\mu_Q(Que_j)$ — функция прямоты вопроса Que_j . Численное значение функции прямоты вопроса отражает «близость» текущего вопроса к прямому вопросу.

Прямым вопросом будем называть вопрос, предполагающий, что ответ содержит решение проблемы. Таким образом, предполагается, что цель диалога может быть достигнута за один шаг при помощи единственного прямого вопроса. Естественно, что при этом активный агент должен «абсолютно» доверять ответам реактивного агента. В противном случае активный агент при помощи серии не прямых вопросов накапливает некоторое количество фактов, используемых как исходные данные для вывода (или расчёта) искомого решения.

Интуитивно ясно, что чем меньше степень прямоты вопросов, которые использует активный партнёр, тем дольше будет длиться диалоговый процесс, и наоборот. Ясно также, что одной и той же цели можно достичь за различное количество шагов в зависимости от степени прямоты используемых вопросов. Поэтому функция прямоты вопроса μ_Q характеризует также «скорость», с которой диалоговый процесс приближается к цели.

Степень прямоты вопроса может трактоваться, как «расстояние» между прямым и непрямым вопросами, либо как «расстояние» до цели диалога. Таким образом, чем меньше активный агент доверяет ответам реактивного агента, тем менее прямые вопросы он задаёт и, следовательно, тем более длительным должен быть «путь к цели», или фактический диалоговый процесс. Если поведение реактивного агента неодинаково на разных фазах диалога, то степень доверия к его ответам тоже различается. Активный агент может учитывать этот факт, изменяя степень прямоты своих вопросов.

Для детерминированной конечно-автоматной модели вопросно-ответного диалога множество $ANS = \{Ans_i\}; i = 1, \dots, l$ представляет собой множество всех ответов, необходимых и достаточных для достижения цели (целей) диалога. Важной характеристикой ответов реактивного агента является их правдивость. Активный агент, зная, насколько правдив тот или иной ответ, может варьировать степень прямоты своих вопросов.

Определим множество ответов ANS как нечёткое следующим образом:

$$ANS = \{Ans_j, \mu_A(Ans_j)\}; j = 1, \dots, l, \quad (3.17)$$

$$\mu_A : ANS \rightarrow [0, 1], \quad (3.18)$$

где:

$\mu_A(Ans_j)$ — функция ожидаемой правдивости ответа Ans_j . Функция μ_A названа функцией ожидаемой правдивости ответа потому, что в реальном вопросно-ответном диалоговом процессе часто сложно точно знать, насколько правдив полученный ответ и активный агент, в этих случаях, может лишь оценивать эту величину.

Таким образом, нечёткая автоматная модель диалогового метода доступа может быть описана семёркой [44,45].

$$DiAM = (QUE, ANS, S, \mu_Q, \mu_A, f, g), \quad (3.19)$$

где:

$$QUE = \{Que_j, \mu_Q(Que_j)\}; j = 1, \dots, k, \quad (3.20)$$

QUE — множество всех непрямых вопросов активного агента, необходимых и достаточных для достижения цели диалога при помощи логического вывода или расчётов.

$$ANS = \{Ans_j, \mu_A(Ans_j)\}; j = 1, \dots, l, \quad (3.21)$$

где:

ANS — множество всех неправдивых ответов реактивного агента, необходимых и достаточных для достижения цели при помощи логического вывода или расчётов.

$$S = \{S_i\}; i = 1, \dots, m, \quad (3.22)$$

где:

S — множество шагов диалога.

$$\mu_Q : QUE \rightarrow [0, 1], \quad (3.23)$$

где:

μ_Q — функция прямоты вопроса активного агента диалога.

$$\mu_A : ANS \rightarrow [0, 1], \quad (3.24)$$

где:

μ_A — функция ожидаемой правдивости ответа реактивного агента диалога.

$$f : S \times ANS \rightarrow S, \quad (3.25)$$

где:

f — функция перехода, переводящая $DiAM$ в новый шаг в соответствии с ответом, полученным от реактивного агента.

$$g : S \times ANS \rightarrow QUE, \quad (3.26)$$

где:

g — функция выхода, детерминирующая индекс последующего вопроса активного агента на основании индекса текущего шага и текущего ответа, полученного от реактивного агента.

Поскольку, как отмечалось выше, значения функций μ_A и μ_Q для, соответственно, предыдущего ответа и очередного вопроса не независимы, необходимо ввести дополнительный формализм, учитывающий их корреляцию:

$$\mu_A(Ans_j) = k * \mu_Q(Que_{j+1}), k \approx 1. \quad (3.27)$$

Нечёткая автоматная модель вопросно-ответного диалога предполагает использование нескольких стратегий при конструировании диалогового метода решения проблемы.

Стратегия 1. Активный агент строит своё поведение, исходя из предположения, что в ходе диалога степень ожидаемой правдивости ответов реактивного агента остается неизменной

и, следовательно, степень правдивости его вопросов также остается неизменной.

Стратегия 2. Активный агент строит своё поведение, исходя из предположения, что в ходе диалога степень ожидаемой правдивости ответов реактивного агента изменяется. Тогда на каждом шаге связь между ожидаемой правдивостью ответа и степенью прямоты очередного вопроса определяется некоторой зависимостью. Например, линейной зависимостью вида (3.27) с коэффициентом, близким к единице (чем меньше ожидаемая правдивость ответа, тем менее прямым должен быть очередной вопрос и наоборот).

3.3. Модель вопросно-ответного диалога в виде сети Петри

Конечно-автоматная модель, описанная в подразделе 3.2, имеет свою сферу применимости и может использоваться для моделирования и специфицирования ряда прикладных диалоговых систем. Так, при помощи этой модели легко описывается метод внутреннего программирования, предложенный Краудером и используемый при проектировании систем программированного обучения [46]. Однако эта модель обладает рядом недостатков, проявляющихся на этапе её компьютерной реализации.

Эффективным подходом к компьютерной реализации диалоговых систем является *даталогический подход* (data-driven approach), который для случая реализации диалогового метода доступа *DiAM* предполагает реализацию функции шага не в виде программного кода, а в виде отображения в базу данных. [47–49]. Решающим преимуществом даталогического подхода является возможность лёгкой модификации диалогового метода *DiAM* при помощи полноэкранных средств редактирования. Таким образом, даталогический подход исключает процесс компиляции исходных текстов при модификации программной системы и позволяет выполнить модификацию си-

лами автора диалогового метода. С позиций даталогического подхода формализм, принятый для описания диалогового метода, необходимо отождествлять с концептуальной моделью некоторой базы данных. Как отмечено в [50], проблематика концептуальной модели связана с таким представлением модели, которое, с одной стороны, наиболее естественно отражает предметную область (в нашем случае диалоговый метод доступа к памяти вопросов), а, с другой, — может поддерживаться компьютерными средствами. Иными словами, необходим формализм, легко трансформируемый в схему данных. Опыт проектирования диалоговых приложений позволяет сделать вывод, что отмеченным свойством в большей степени обладает не диаграмма состояний диалогового метода доступа *DiAM* (пример которой приведен на рис. 3.4), а граф, моделирующий метод доступа *DiAM* в виде сети Петри [51].

Петри-модель метода доступа к памяти вопросов так же, как и автоматная модель, базируется на двух предположениях: (1) проблема, решаемая в процессе диалога, может быть решена за конечное количество шагов диалога, а к началу решения проблемы детерминированы все вопросы и ответы, необходимые для её решения; (2) модификация текущего набора ожидаемых ответов реактивного агента в любом случае порождает набор ответов, отличающийся от текущего и, следовательно, имеющий уникальный индекс.

Петри-модель диалогового метода опишем четвёркой

$$DiAM = (QUE, ANS, NextQue, NextAns), \quad (3.28)$$

$$QUE = \{Que_i\}; i = 1, \dots, n, \quad (3.29)$$

где:

QUE — комплект индексов всех вопросов активного агента, которые, согласно предположению 1, необходимы и достаточны для решения проблемы. Комплект (3.29) составлен из элементов множества (3.2). Здесь под комплектом понимается набор элементов, в котором в отличие от множества допускается присутствие нескольких экземпляров одного и того же элемента [51].

Если $\#(x, B)$ функция, которая определяет количество экземпляров элемента x в комплекте B , то мощность комплекта вопросов определяется формулой $n = |QUE| = \sum_i \#(Que_i, QUE)$.

$$ANS = \{Ans_i\}; i = 1, \dots, p, \quad (3.30)$$

где:

ANS — комплект всех ответов реактивного агента, составленный из элементов множества (3.3). Мощность комплекта ответов определяется формулой

$$p = |ANS| = \sum_i \#(Ans_i, ANS).$$

Комплекты вопросов и ответов не пересекаются, т.е. имеет место $QUE \cap ANS = 0$.

$$NextQue : ANS \rightarrow QUE, \quad (3.31)$$

где:

$NextQue$ — функция индексов последующих вопросов, которая задаёт отображение комплекта ожидаемых ответов в комплект индексов вопросов. Специфика вопросно-ответного диалогового процесса двух агентов, заключающаяся в том, что активный агент на каждом шаге диалога генерирует и передаёт реактивному агенту только один вопрос, накладывает на функцию $NextQue$ следующее ограничение. Для каждого ответа из комплекта ожидаемых ответов функция индексов последующих вопросов детерминирует индекс одного единственного вопроса из комплекта вопросов.

$$NextAns : QUE \rightarrow ANS, \quad (3.32)$$

где:

$NextAns$ — функция последующих ответов, задающая отображение комплекта индексов вопросов в комплект ожидаемых ответов.

Для иллюстрации Петри-модели представим пример, описанный в подразделе 3.2 в виде Петри-модели:

$$ANS = \{Ans_1, Ans_2, Ans_3, Ans_4, Ans_5\}, \quad (3.33)$$

$$QUE = \{Que_1, Que_2, Que_3, Que_4\}, \quad (3.34)$$

$$\begin{aligned} NextAns(Que_1) &= \{Ans_1, Ans_3, Ans_4, Ans_5\}, \\ NextAns(Que_2) &= \{Ans_1, Ans_2, Ans_3, Ans_4, Ans_5\}, \\ NextAns(Que_3) &= \{\}, \\ NextAns(Que_4) &= \{\}, \end{aligned} \quad (3.35)$$

$$\begin{aligned} NextQue(Ans_1) &= \{Que_1\}, \\ NextQue(Ans_2) &= \{Que_1\}, \\ NextQue(Ans_3) &= \{Que_3\}, \\ NextQue(Ans_4) &= \{Que_4\}, \\ NextQue(Ans_5) &= \{Que_2\}. \end{aligned} \quad (3.36)$$

Формулы (3.35) и (3.36) можно использовать для табличного задания функций $NextAns$ и $NextQue$, соответственно. Правая часть i -й строки функции $NextAns$ задает множество ответов реактивного агента, ожидаемых на i -м шаге:

$$NextAns(Que_i) = R^i \cup NR^i. \quad (3.37)$$

Правая часть i -й строки функции $NextQue$ состоит из одного элемента комплекта QUE , что является следствием отмеченного ограничения, заключающегося в том, что для каждого ответа из комплекта ожидаемых ответов функция индексов последующих вопросов детерминирует индекс одного единственного вопроса.

На рис. 3.5 приведен граф Петри-модели диалогового метода доступа, соответствующий рассмотренному примеру. При графической интерпретации Петри-модели диалогового метода доступа комплект индексов вопросов QUE и комплект ответов ANS представлены множествами позиций (кружки) и переходов (прямоугольники) соответственно. Множество ожидаемых на i -м шаге ответов $R^i \cup NR^i$ представлено совокупностью переходов инцидентных i -й позиции и соединённых с ней исходящими ветвями.

Для Петри-модели шаг является естественным «строительным блоком» диалогового метода доступа и для i -го шага может быть определён, как совокупность

$$Que_i, (R^i \cup NR^i). \quad (3.38)$$

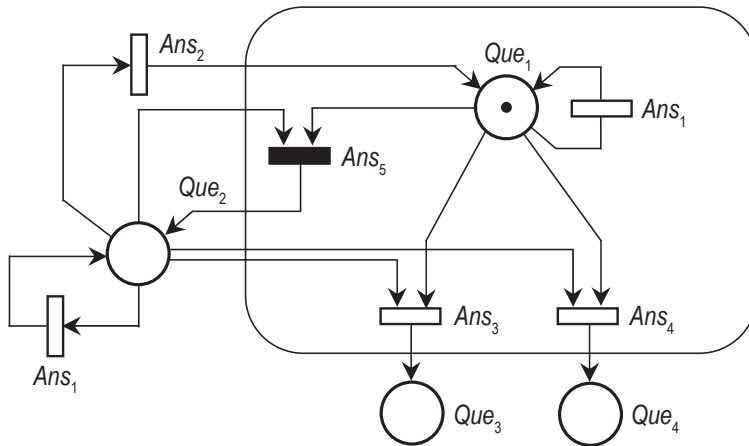


Рис. 3.5. Граф Петри-модели диалогового метода доступа для примера, описанного в тексте. Выделены элементы сети, объединённые понятием «шаг диалога». Зачернённый переход моделирует множество нераспознаваемых ответов

При графической интерпретации Петри-модели шаг диалога представляет собой некоторую позицию и совокупность переходов, соединённых с ней исходящими ветвями. На рис. 3.5 выделены элементы, входящие в объём понятия «шаг диалога».

Граф Петри-модели маркирован единственной фишкой, отмечающей текущий активный шаг. Начальная разметка в этом случае — это указание на первый шаг, с которого начинается реализация диалогового метода. Условием запуска перехода является совпадение ответа, кодируемого данным переходом,

и ответа, полученного от реактивного агента. Иными словами, переход запускается и «пропускает» фишку, если от реактивного агента получен ответ, совпадающий с ответом, кодируемым данным переходом.

С точки зрения существующей классификации, сеть Петри, граф которой приведен на рис. 3.5, является ординарной (все дуги имеют одну и ту же кратность, равную единице) и не автоматной, по крайней мере, согласно определениям, приведенным в [52, 53], поскольку предлагаемая Петри-модель допускает несколько входов для одного и того же перехода.

Теория сетей Петри предоставляет удобный аппарат для отслеживания динамики моделируемого процесса, состоящий из маркировки позиций и условий запуска переходов. Поскольку в Петри-модели вопросно-ответного диалога имеется единственная фишка, то маркировка может быть представлена вектором, количество компонентов которого равно количеству элементов комплекта индексов вопросов — числу n :

$$\mu = (\mu_1, \mu_2, \dots, \mu_n). \quad (3.39)$$

Каждый компонент вектора μ принимает значения на двух-элементном множестве $\{0, 1\}$. $\mu_i = 0$ в том случае, если фишка отсутствует в i -й позиции, соответствующей i -му элементу комплекта QUE . $\mu_i = 1$ в том случае, если фишка находится в i -й позиции, соответствующей i -му элементу комплекта QUE .

Диалоговый процесс сопровождается перемещением фишки по сети без её уничтожения или размножения. Следовательно, в любой момент диалога имеет место $\sum_i \mu_i = 1$.

Таким образом, в процессе вопросно-ответного диалога диалоговый метод характеризуется непрерывно изменяющимся вектором μ , который указывает на текущий активный шаг диалога. Для сети, приведенной на рис. 3.5, вектор μ указывает на шаг номер один и имеет следующее значение: $\mu = (1, 0, 0, 0)$.

Для того чтобы фишка переместилась в последующую позицию, необходимо, чтобы запустился один из переходов шага.

Представление диалогового метода в виде сети Петри моделирует его пространственную структуру, но не отражает динамику диалогового процесса, задаваемую в общем случае когнитивным циклом вопросно-ответного диалогового процесса, изображенном на рис. 3.2. Дополним статическое описание диалогового метода компонентами, моделирующими динамику диалогового процесса. С этой целью рассмотрим, каким образом в рамках предложенной Петри-модели реализуется цикл вопросно-ответного диалогового процесса. Моделирование цикла диалогового процесса будем рассматривать, исходя из предположения, что функции *NextAns* и *NextQue* представлены таблично. Такое предположение естественно, поскольку трудно представить другой, практически значимый способ описания отмеченных функций.

Отдельная строка таблицы функции *NextAns* соответствует одному шагу диалога. Каждая строка этой таблицы включает индекс вопроса, генерируемого на данном шаге, за которым следует множество ответов, ассоциированных с данным шагом. Таким образом, таблица функции *NextAns* состоит из n строк, где n — количество элементов комплекта индексов вопросов *QUE*. На рис. 3.6 приведена структура таблицы функции *NextAns*. Поскольку множество ответов, ассоциированных с шагом, индивидуально для каждого шага и определяется количеством ожидаемых на данном шаге ответов, то часть ячеек в таблице на рис. 3.6 не заполнена.

Que_1	<i>Ans</i>	<i>Ans</i>	...	<i>Ans</i>			
Que_2	<i>Ans</i>	<i>Ans</i>	...	<i>Ans</i>	<i>Ans</i>		
Que_i	<i>Ans</i>	<i>Ans</i>	...	Ans_i	...	<i>Ans</i>	
Que_n	<i>Ans</i>	<i>Ans</i>	...	<i>Ans</i>	...	<i>Ans</i>	

Рис. 3.6. Структура таблицы функции *NextAns*.
Заштрихованы незаполненные ячейки

Таблица функции *NextQue* имеет более простую структуру. Она состоит из p строк, где p — количество элементов комплекта ответов *ANS*. Каждая строка таблицы функции *NextQue* содержит два элемента: ожидаемый ответ и связанный с ним последующий вопрос.

Введем в модель унифицированный процесс *DiMC*, моделирующий циклически повторяющуюся динамику диалога:

$$DiMC(\mu, NextAns, NextQue). \quad (3.40)$$

Унифицированный процесс *DiMC* управляет переходом от предыдущего шага диалога к последующему, что эквивалентно открытию одного из переходов шага и перемещению фишки. Унифицированность процесса *DiMC* означает, что для выполнения любого шага диалога необходимо выполнить одну и ту же последовательность действий и что эта последовательность действий не зависит от предметной области диалогового процесса. На содержательном уровне отдельную итерацию цикла *DiMC* опишем следующим образом.

1. При помощи вектора маркировки μ определяется текущий активный шаг диалогового процесса, который соответствует номеру строки в таблице функции *NextAns*.
2. Определяется и передаётся в память вопросов индекс текущего вопроса. При помощи спецификаций вопроса, находящихся в памяти вопросов, активный агент генерирует и передаёт реактивному агенту текущий вопрос. Активный агент переходит в состояние ожидания ответа реактивного агента.
3. Активный агент воспринимает текущий ответ реактивного агента.
4. Обрабатывается текущий ответ реактивного агента. Последовательно сравнивается воспринятый ответ со всеми, ожидаемыми на данном шаге ответами, хранящимися в ячейках выбранной строки таблицы функции *NextAns*. Для совпавшего ответа (например, Ans_j) определяется строка таблицы функции *NextQue*.

5. При помощи таблицы функции *NextQue* определяется компонент вектора μ , который детерминирует следующий шаг диалога. Значение этого компонента изменяется с нулевого на единичное, и процесс возвращается к пункту 1.

На рис. 3.7 показана последовательность взаимодействия вектора маркировки μ и таблиц функций *NextAns* и *NextQue* при выполнении процесса *DiMC* на стороне активного агента.

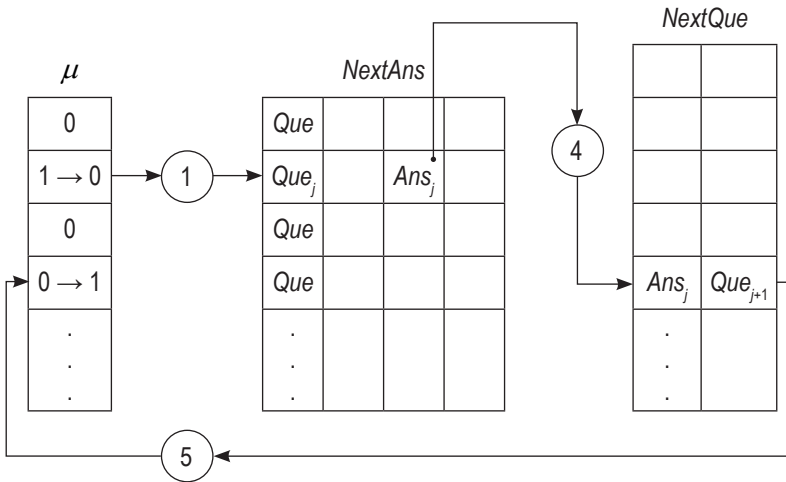


Рис. 3.7. Использование вектора маркировки μ и таблиц функций *NextAns* и *NextQue* процессом *DiMC*. Номера в окружностях соответствуют пунктам содержательного описания *DiMC*

Среди действий, итерационно выполняемых унифицированным процессом *DiMC*, есть два действия, требующие более детального рассмотрения и уточнения: (1) генерация текущего вопроса; (2) обработка текущего ответа. Поскольку *DiMC* является унифицированным процессом, который «единообразно» обрабатывает любой шаг, то и генерация вопроса, и обработка ответа осуществляются по одним и тем же правилам для любого шага диалога. Эти правила определяют некую стандартную модель функционирования активного агента диалога.

Иными словами, *DiMC* неким стандартным способом генерирует вопрос и обрабатывает ответ. Ясно, что даже при достаточно обширном стандарте могут встретиться случаи, когда стандартных средств недостаточно. Например, при генерации вопроса в том случае, когда он представлен не вербально, невозможно учесть всё многообразие невербальной презентации информации, а при обработке ответа может понадобиться специфическая численная обработка. Поэтому кроме базового процесса *DiMC* в модель целесообразно ввести «внешние», по отношению к *DiMC*, процессы. Эти процессы должны вызываться из *DiMC* и возвращать ему управление после завершения. Назовём такие процессы процессами-демонами и введём в модель два класса процессов-демонов:

QueDemon и *AnsDemon*,

которые осуществляют «дообработку» вопроса или ответа на любом шаге диалогового процесса в том случае, когда стандартных средств *DiMC* недостаточно.

3.3.1. Применимость модели вопросно-ответного диалога в виде сети Петри

Рассмотрим применимость модели вопросно-ответного диалога в виде сети Петри. С этой целью разделим диалоговые методы на два класса: легко реализуемые в рамках Петри-модели и трудно реализуемые в рамках Петри-модели.

К легко реализуемым диалоговым методам отнесём методы, в которых *последующие вопросы полностью детерминуются предыдущими ответами*. Примерами таких методов могут служить методы диалоговых лингводидактических систем, предназначенные для обучения иностранным языкам. В лингводидактических системах логика сценария диалога построена таким образом, что активный агент (учитель) передаёт реактивному агенту (ученику) серию обучающих стимулов,

имеющих статус вопроса. Каждый последующий обучающий стимул полностью детерминируется предыдущим ответом реактивного агента.

К трудно реализуемым диалоговым методам отнесём методы, в которых *последующие вопросы не зависят от предыдущих ответов*. Примерами таких методов могут служить методы психодиагностических тестов. Как правило, на первом этапе психодиагностического тестирования (в процессе диалогового взаимодействия с субъектом диагностики) формируется первичный набор данных в виде «истории» ответов субъекта, который затем используется для формирования заключения. Примером психодиагностического теста может служить методика прогрессивных матриц Равена, используемая для тестирования интеллектуальных способностей [54]. Методика прогрессивных матриц Равена предполагает, что на каждом шаге тестирования субъекту тестирования предъявляются невербальные стимулы, имеющие статус вопроса, а последовательность вопросов, передаваемых реактивному агенту (субъекту тестирования), формируется активным агентом без учёта ответов реактивного агента.

Рассмотрим пример, который позволит проанализировать причины трудной реализации психодиагностических тестов при помощи Петри модели. Для упрощения модели полагаем, что диалоговый метод психодиагностического теста реализуется при помощи трёх шагов. На первом шаге активный агент формирует и передаёт реактивному агенту вопрос Que_1 на который ожидает получить один из двух ответов: Ans_1 , Ans_2 . На втором шаге, вне зависимости от ответа, полученного на первом шаге, активный агент передаёт реактивному агенту вопрос Que_2 , на который ожидает получить один из ответов: Ans_3 , Ans_4 . На третьем шаге, вне зависимости от ответа, полученного на втором шаге, активный агент передаёт реактивному агенту вопрос Que_3 , на который ожидает получить один из ответов: Ans_5 , Ans_6 . Решение задачи (заключение) формируется в зависимости от комбинации ответов, полученных на первом, втором и третьем шагах диалогового процесса.

На рис. 3.8 приведена сеть Петри для описанного примера, моделирующая диалоговый метод, соответствующий описанному психодиагностическому тесту. Нераспознаваемые ответы реактивного агента опущены, чтобы не загромождать рисунок.

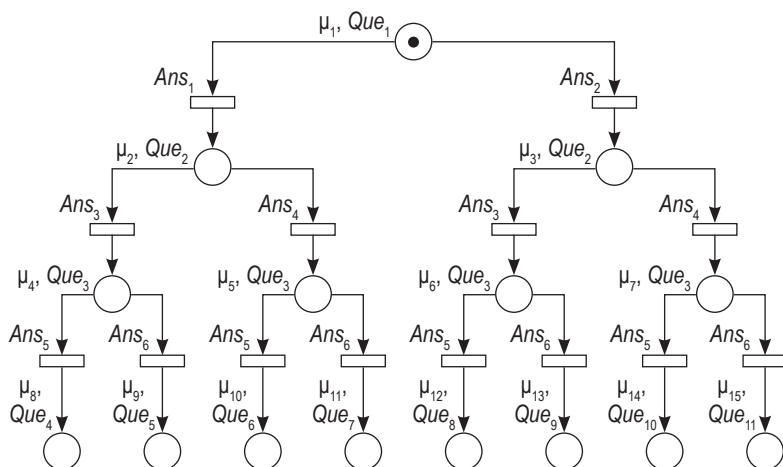


Рис. 3.8. Сеть Петри трудно реализуемого диалогового метода для примера, описанного в тексте

В нижней части рис. 3.8 изображены целевые позиции, которые соответствуют окончанию диалогового процесса и формированию заключения. Поэтому Que_4 — Que_{11} обозначают не вопросы, а заключения. В рассматриваемом примере сложность реализации диалогового метода проявляется в быстром увеличении количества шагов диалога в зависимости от количества независимых вопросов и количества ожидаемых ответов для каждого из вопросов.

Для случая, приведенного на рис. 3.8, количество шагов диалогового метода может быть подсчитано по формуле

$$n = 2^0 + 2^1 + 2^2 + 2^3. \quad (3.41)$$

В общем случае, если на каждый вопрос ожидается одинаковое количество ответов a , т.е. для каждого шага i имеет место

$$\left| R^i \cup NR^i \right| = const = a, \quad (3.42)$$

где:

$\left| R^i \cup NR^i \right|$ мощность множества ответов, ожидаемых на i -м шаге, то количество шагов диалогового метода может быть подсчитано по формуле

$$n = a^0 + a^1 + a^2 + \dots + a^\alpha, \quad (3.43)$$

где α — общее количество независимых последовательных вопросов теста.

Трудность реализации диалогового метода, решающего психодиагностическую задачу при помощи сети Петри, иллюстрируется числом шагов, необходимых для построения полного дерева. Так, методика прогрессивных матриц Равена включает 60 матриц-вопросов, на каждый из которых диалоговый метод ожидает только два ответа (правильный и неправильный). Общее количество шагов диалогового метода может быть посчитано по формуле (3.43) и определяется числом $2^{61}-1$. Причина, породившая древовидную форму сети Петри и, как следствие, приведшая к большому количеству шагов, заключается в том, что модель позволяет учесть «историю» ответов, необходимую для получения заключения, только одним способом — формированием в сети Петри *уникальных траекторий прохождения фишки*.

Модель диалогового метода, решающего психодиагностическую задачу, можно существенно упростить, если: (1) ввести в модель *память истории ответов PathMem*; (2) ввести в модель переходы, условия запуска которых учитывают как текущий ответ, полученный от реактивного агента, так и содержимое памяти истории ответов. При графическом представлении диалогового метода переход, соответствующий одному из ожидаемых ответов и учитывающий историю ответов, имеет одну входящую ветвь (от предыдущего вопроса) и несколько исходящих ветвей (к последующим вопросам или заключениям). Таким образом, в модели имеется не одно условие запуска переходов, общее для всех переходов сети, а два условия запуска переходов, соответствующие двум типам переходов (пере-

ход, не учитывающий историю ответов, и переход, учитывающий историю ответов).

Наличие памяти истории ответов означает учёт сформулированного в подразделе 3.1 принципа истории ответов, который означает, что диалоговый метод при формировании очередного вопроса должен учитывать как только что воспринятый ответ, так и ранее воспринятые ответы. В память истории ответов *PathMem* последовательно записываются все ответы реактивного агента, начиная с первого шага диалога. Память *PathMem* может быть реализована различным способом. Например, в виде очереди с FIFO (First In First Out) правилом обслуживания. В этом случае в «голове» очереди будет находиться самый первый ответ реактивного агента.

На рис. 3.9. изображена Петри модель для описанного примера диалогового метода, решающего психодиагностическую задачу с учётом памяти истории ответов.

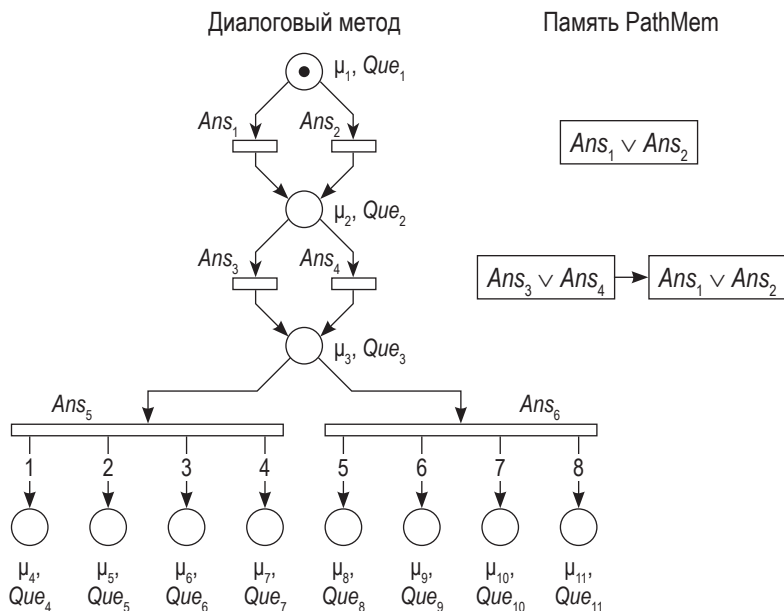


Рис. 3.9. Сеть Петри модели трудно реализуемого диалогового метода при наличии памяти истории ответов *PathMem*

В Петри-модели диалогового метода, изображенного на рис. 3.9, имеются переходы двух типов:

- Переходы Ans_1 — Ans_4 , имеющие одну входящую ветвь от предыдущего вопроса и одну исходящую ветвь к последующему вопросу.
- Переходы Ans_5 и Ans_6 имеющие одну входящую ветвь от предыдущего вопроса и несколько исходящих ветвей к последующим вопросам.

Условия запуска этих переходов различны и могут быть описаны при помощи условных предложений (продукционных правил). Количество условных предложений, описывающих условия запуска перехода, равно количеству ветвей, исходящих из перехода. Например, условие срабатывания перехода, соответствующего Ans_1 на рис. 3.9, может быть записано в виде одного условного предложения:

ЕСЛИ текущий ответ Ans_1
ТО перейти к вопросу Que_2 (3.44)

Для перехода, соответствующего Ans_5 на рис. 3.9, условия срабатывания описываются четырьмя условными предложениями:

ЕСЛИ текущий ответ Ans_5 И $PathMem = \{Ans_1, Ans_3\}$
ТО перейти к вопросу Que_4

ЕСЛИ текущий ответ Ans_5 И $PathMem = \{Ans_1, Ans_4\}$
ТО перейти к вопросу Que_5 (3.45)

ЕСЛИ текущий ответ Ans_5 И $PathMem = \{Ans_2, Ans_3\}$
ТО перейти к вопросу Que_6

ЕСЛИ текущий ответ Ans_5 И $PathMem = \{Ans_2, Ans_4\}$
ТО перейти к вопросу Que_7

Модель вопросно-ответного диалога в виде сети Петри имеет более широкую сферу применимости, чем конечно-ав-

томатная модель. Её преимущество в том, что она моделирует стандартный цикл диалогового процесса при помощи унифицированного процесса *DiMC*, а также возможность нестандартной обработки вопросов и ответов при помощи уникальных процессов *QueDemon* и *AnsDemon*. Модель диалога в виде сети Петри также позволяет учитывать принцип истории ответов реактивного агента.

Как показал проведенный анализ, легко реализуемыми в рамках Петри-модели диалоговыми методами являются методы, в которых существует *тесная логическая связь между предыдущим ответом реактивного агента и последующим вопросом активного агента*. В тех случаях, когда последовательность вопросов не зависит от предыдущих ответов, модель должна включать память истории ответов, в которой накапливаются предыдущие ответы, а заключения формируются с учётом содержимого этой памяти.

3.4. Сетевая даталогическая модель вопросно-ответного диалога

Модель, описанная в данном подразделе, развивает главные идеи двух предыдущих моделей, но базируется не на математической структуре (конечный автомат Мили или сеть Петри), а на идее представления диалогового метода *в виде сети, состоящей из разнотипных узлов и последующим отображением этой сети в базу данных*. В предлагаемой модели существенно следующее:

- учитываются два из трёх, сформулированных в подразделе 3.1, принципов организации вопросно-ответного диалога: (1) принцип глубины диалога и (2) принцип истории ответов реактивного агента;
- развивается даталогический подход и а priori предполагается, что диалоговый метод и спецификации вопросов активного агента представляются данными,

хранящимися в некоторой базе данных, а не программными кодами;

- развивается сетевое понимание структуры диалогового метода.

Как и ранее (смотри подраздел 3.1), предполагаем, что модель включает сценарий активного агента диалога, который состоит из памяти вопросов, хранящей спецификации вопросов, и диалогового метода доступа к памяти вопросов *DiAM*. Также, предполагаем, что модель включает память истории ответов реактивного агента *PathMem*, в которой накапливается последовательность ответов реактивного агента.

DiAM имеет сетевую организацию и состоит из связанных между собой узловых элементов. Отдельный узел *DiAM* соответствует одному шагу диалога. В каждый момент времени активным является только один узел. Определение последующего узла-шага и его активизацию осуществляет унифицированный процесс *DiMC*, а если это необходимо, то и процесс-демон *AnsDemon*. Связи между узловыми элементами являются направленными, а с каждым узлом ассоциировано как множество возможных предыдущих шагов (входящие ветви), так и множество возможных последующих шагов диалога (исходящие ветви).

Сеть *DiAM* неоднородна в том смысле, что она строится из узлов разных типов. Узлы отличаются назначением, описанием и процедурой навигации к последующим узлам. Идея неоднородности сети является принципиальной и означает, во-первых, что сеть состоит из разнотипных узловых элементов, а во-вторых, что возможна различная типология узловых элементов. В описанной модели используется классификация, которая делит узловые элементы на три типа [55].

- *Узел прямой навигации* (узел типа *Node1*). Это узел, который соответствует шагу диалога, предполагающему продолжение диалогового процесса. Для узла типа *Node1* индекс следующего шага определяется только с учётом принципа глубины (индекс следующего шага детерминирован только текущим ответом реактивного агента).

- *Узел условной навигации* (узел типа *Node2*). Это узел, который также соответствует шагу диалога, предполагающему продолжение диалога. Для узла типа *Node2* индекс следующего шага определяется с учётом принципа глубины и принципа истории ответов (индекс следующего шага детерминируется и текущим ответом реактивного агента, и содержимым памяти истории ответов).
- *Целевой узел* (узел типа *Node3*). Это узел, которым завершается диалоговый процесс решения проблемы.

Для каждого узла типов *Node1* и *Node2* существует своя процедура навигации, при помощи которой определяется индекс следующего узла.

Навигация для узла типа *Node1* осуществляется следующим образом. Текущий ответ реактивного агента рассматривается как «закодированный индекс» следующего шага. Поэтому процесс его распознавания (либо классификации как нераспознанного) рассматривается как процесс табличной перекодировки воспринятого ответа в индекс следующего шага. Количество исходящих ветвей для узла с прямой навигацией равно количеству ожидаемых ответов реактивного агента для шага диалога, соответствующего узлу с прямой навигацией. Процедура навигации для узла с прямой навигацией реализуется унифицированным процессом *DiMC* и является его частью. Отметим, что существенным отличием диалогового процесса от других видов интерактивного взаимодействия (например, интерактивного взаимодействия с базой данных) является тесная логическая связь не только между вопросом и ответом на этот вопрос, но и *между текущим ответом и последующим вопросом*. Поэтому следует ожидать, что узлы с прямой навигацией должны быть основными строительными элементами диалогового метода.

Процедура навигации для узла типа *Node2* должна учитывать как текущий ответ реактивного агента, так и содержимое памяти *PathMem*. Количество исходящих ветвей для узла с условной навигацией равно количеству ожидаемых ответов

реактивного агента для шага, соответствующего узлу с условной навигацией, умноженное на количество состояний памяти *PathMem*. Это может быть очень большое число. Для определения количества возможных состояний памяти *PathMem* необходимо перемножить количество ожидаемых ответов для всех шагов, учтённых в памяти истории ответов. Например, если в памяти истории ответов хранятся ответы, полученные на десяти предшествующих шагах, а на каждом из предшествующих шагов ожидалось по три ответа, то количество состояний памяти *PathMem* равно 3^{10} . Процедуру навигации для узлов с условной навигацией для памяти истории ответов, хранящей всю историю ответов, трудно (если вообще возможно) реализовать при помощи унифицированного процесса *DiMC*. Поэтому алгоритм навигации для узлов с условной навигацией целесообразно кодировать в процедуре-демоне *AnsDemon*. Сделать процедуру навигации для узла с условной навигацией частью унифицированного процесса *DiMC* можно в том случае, когда память истории ответов хранит ответ, полученный только на одном из предшествующих шагов. В естественном диалоге, когда оба агента являются личностями, обычно так и происходит. Активный агент может запомнить ответ реактивного агента на некоторый (один) вопрос и далее, в процессе диалога, задать тот же вопрос, немного изменив его формулировку. При выборе варианта продолжения диалога активный агент будет учитывать как ранее полученный ответ, так и текущий ответ.

Как было отмечено введенная классификация узлов не является единственной. При практической реализации модели может быть полезна, например, классификация, учитывающая тип вопроса, ассоциированного с шагом. Важно, что сетевой подход к моделированию диалогового метода позволяет вводить сколь угодно сложную классификацию узлов в отличие, например, от сети Петри, которая, являясь двудольным графом, предполагает наличие только двух типов вершин.

Даталогический подход означает, что сеть *DiAM* отображается в базу данных. Для того чтобы такое отображение было возможным, необходимо вначале разработать даталогическое

представление каждого из введенных типов узлов. Здесь под даталогическим представлением узла понимаются данные, при помощи которых описывается узел. На рис. 3.10 в графическом виде приведено даталогическое представление узла типа *Node1*.

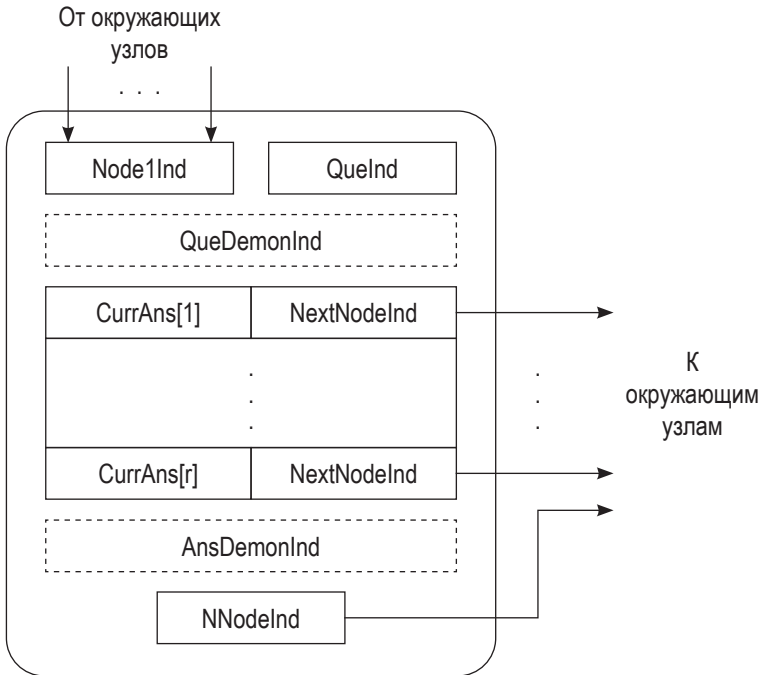


Рис. 3.10. Даталогическое представление узла типа *Node1*.
Пунктиром отмечены необязательные элементы

На рис. 3.10 использованы следующие обозначения:
Node1Ind — индекс узла, в котором отражён его тип.
QueInd — индекс вопроса, ассоциированного с данным узлом.
QueDemonInd — индекс необязательного процесса-демона, который может участвовать в генерации вопроса для данного узла.
CurrAns[1]...CurrAns[r] — текущие ожидаемые ответы.

NextNodeInd — индекс последующего узла для распознанного ответа.

NNodeInd — индекс последующего узла для нераспознанных ответов.

AnsDemonInd — индекс необязательного процесса-демона, который может участвовать в трансформации текущего ответа в индекс последующего узла.

Даталогическое описание узла типа *Node1*, приведенное на рис. 3.10, интерпретируется унифицированным процессом *DiMC*, который также реализует процедуру навигации для узла типа *Node1*.

В подразделе 3.3 отдельная итерация процесса *DiMC* была описана в терминах Петри-модели. Для случая сетевой модели отдельную итерацию *DiMC* удобно представить в виде следующих, последовательно выполняемых трёх фаз.

Фаза генерации вопроса. В памяти вопросов отыскивается даталогическое описание текущего вопроса с индексом *QueInd*. Используя это описание, *DiMC* генерирует текущий вопрос. Если даталогическое описание текущего узла включает *QueDemonInd*, то этот процесс вызывается и участвует в генерации текущего вопроса.

Фаза восприятия ответа. Воспринимается текущий ответ. Воспринятый ответ записывается в память истории ответов *PathMem*.

Фаза навигации. Определяется индекс следующего узла. Если даталогическое описание узла включает *AnsDemonInd*, то этот процесс загружается и участвует в определении индекса следующего узла.

С учётом введенных обозначений «логику» процедуры навигации для узла типа *Node1* запишем в виде условных предложений:

```

ЕСЛИ распознан CurrAns[i]
ТО перейти к узлу NextNodeInd
ИНАЧЕ
ЕСЛИ не распознан CurrAns[i]
ТО перейти к узлу NNodeInd

```

(3.46)

Идея учёта истории ответов для узлов с условной навигацией предполагает учёт «произвольной» истории ответов. В этом случае память истории ответов *PathMem* может представлять собой очередь с правилом обслуживания типа FIFO (First In First Out) и возможностью прямого доступа к элементам очереди. Иными словами, индексы следующих узлов для каждого распознанного текущего ответа должны определяться при условии, что на некотором множестве предшествующих шагов были восприняты конкретные ответы. Как уже отмечалось, в связи с проблематичностью даталогического представления узла с условной навигацией, учитывающего «произвольную» историю ответов, рассмотрим случай, когда узел типа *Node2* для каждого распознанного текущего ответа учитывает набор ожидаемых ответов *только для одного произвольного предшествующего шага PathMemInd*. На рис. 3.11 в графическом виде приведено даталогическое представление узла типа *Node2*.

На рис. 3.11 использованы следующие дополнительные обозначения:

$PrevAns[1], \dots, PrevAns[r]$ — ожидаемые ответы для предшествующего шага *PathMemInd* ;

$CurrAns[1], \dots, CurrAns[r]$ — ожидаемые ответы для текущего шага.

Интерпретация данных для узла типа *Node2* отличается от интерпретации данных для узла типа *Node1* только фазой навигации. С учётом введенных обозначений «логику» навигации для узла типа *Node2* запишем в виде

ЕСЛИ распознан $CurrAns[i]$
 И распознан $PrevAns[j]$ на шаге *PathMemInd*
 ТО перейти к узлу $NextNodeInd[i, j]$
 ИНАЧЕ
 ЕСЛИ не распознан $CurrAns[i]$
 ТО перейти к узлу $NNodeInd$ (3.47)

На рис. 3.12 приведено даталогическое представление узла типа *Node3*. Поскольку узел типа *Node3* является целевым и

переход на этот узел эквивалентен завершению работы диалогового метода, то, во-первых, его описание не предполагает перехода к последующим узлам, во-вторых, *QueInd* является индексом утверждения, а не вопроса. Отметим, что при практической реализации модели в виде программной системы достижение узла типа *Node3* может означать переход в программную среду более высокого уровня, например в среду операционной системы.

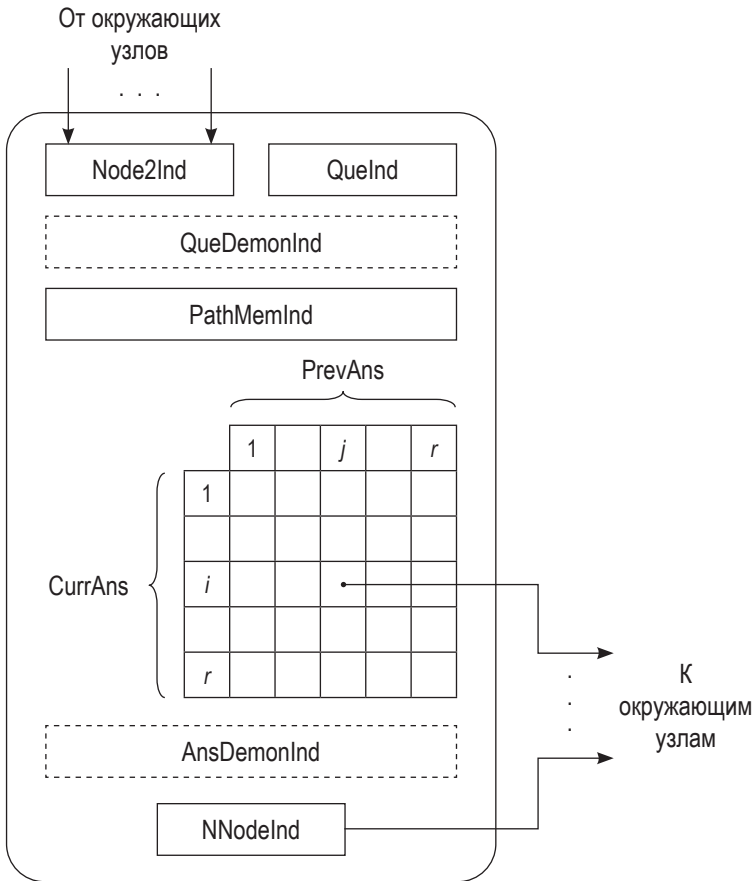
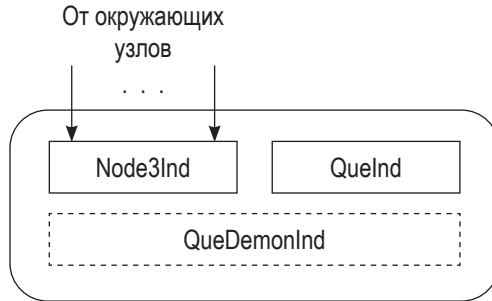


Рис. 3.11. Дatalogическое представление узла типа *Node2*.
Пунктиром отмечены необязательные описатели

Рис. 3.12. Даталогическое представление узла типа *Node3*

3.4.1. Отображение сетевой модели в реляционную базу данных

Опишем сеть *DiAM*, состоящую из узловых элементов типа *Node1*, *Node2* и *Node3*, в виде абстрактной реляционной базы данных, используя нотацию для схем отношений, предложенную в [56]. Формальное представление реляционной базы данных, согласно этой нотации, включает список имён отношений, составляющих базу. Каждое отношение представлено списком имён атрибутов. Атрибуты, формирующие ключ отношения, выделяются подчёркиванием. Значения атрибутов ключа позволяют однозначно идентифицировать любой кортеж (строку) отношения. Для каждого атрибута специфицируется его домен (область значений).

Задача проектирования реляционной базы данных является, по сути, задачей информационного моделирования и не имеет единственного решения, поскольку одну и ту же систему можно представить в виде нескольких наборов отношений. Задача информационного моделирования сети *DiAM* заключается в описании сети, состоящей из узловых элементов типов *Node1*, *Node2* и *Node3* в виде набора отношений. Ясно, что можно предложить несколько вариантов такого описания. Приведенная база не претендует на оптимальность, она скорее, иллюстрирует конструктивность подхода.

Сеть *DiAM* будем моделировать следующими пятью отношениями.

$$DiAM = \{Step, Navig1, Navig2, QueDemon, AnsDemon\}, \quad (3.48)$$

где:

Step — отношение, содержащее данные, описывающие шаги диалога (узлы *DiAM*).

Navig1 — отношение, содержащее данные, необходимые для навигации из узла типа *Node1*.

Navig2 — отношение, содержащее данные, необходимые для навигации из узла типа *Node2*.

QueDemon — отношение, содержащее данные, необходимые для вызова процесса *QueDemon*.

AnsDemon — отношение, содержащее данные, необходимые для вызова процесса *AnsDemon*.

Представим каждое из отношений базы данных (3.48) в виде набора атрибутов:

$$Step = \{NodeInd, NodeType, QueInd, NNodeInd\} \quad (3.49)$$

$$dom(NodeInd) = \{1, \dots, NodeQ\};$$

NodeQ — общее количество узлов *DiAM*;

$$dom(NodeType) = \{Node1, Node2, Node3\};$$

$$dom(QueInd) = \{1, \dots, QueQ\};$$

QueQ — общее количество различных вопросов, $NodeQ \geq QueQ$;

$$dom(NNodeInd) = \{1, \dots, NodeQ\}.$$

Один кортеж отношения *Step* описывает один узел *DiAM*, поэтому количество кортежей отношения *Step* равно *NodeQ*.

$$Navig1 = \{Node1Ind, CurrAns, NextNodeInd\} \quad (3.50)$$

$$dom(Node1Ind) = \{1, \dots, Node1Q\};$$

$Node1Q$ — общее количество узлов с прямой навигацией;

$$dom(CurrAns) = \{CurrAns_1, \dots, CurrAns_r\};$$

r — количество распознаваемых ответов для узла с прямой навигацией;

$$dom(NextNodeInd) = \{1, \dots, NodeQ\}.$$

Один кортеж отношения $Navig1$ описывает одну связь узла с прямой навигацией с окружающими узлами. Общее количество кортежей в отношении $Navig1$ равно r .

$$Navig2 = \{Node2Ind, CurrAns, PathMemInd, PrevAns, NextNodeInd\} \quad (3.51)$$

$$dom(Node2Ind) = \{1, \dots, Node2Q\};$$

$Node2Q$ — общее количество узлов с условной навигацией;

$$dom(CurrAns) = \{CurrAns_1, \dots, CurrAns_r\};$$

$$dom(PathMemInd) = \{1, \dots, N\};$$

N — количество элементов в памяти истории ответов.

$$dom(PrevAns) = \{PrevAns_1, \dots, PrevAns_r\};$$

$$dom(NextNodeInd) = \{1, \dots, NodeQ\}.$$

Один кортеж отношения $Navig2$ описывает одну связь узла типа $Node2$ с окружающими узлами. Общее количество кортежей в отношении $Node2$ равно r^2 .

$$QueDemon = \{NodeInd, QueDemonInd\} \quad (3.52)$$

$$dom(NodeInd) = \{1, \dots, NodeQ\};$$

$$dom(QueDemonInd) = \{1, \dots, QueDemonQ\};$$

$QueDemonQ$ — общее количество различных процессов $QueDemon$.

$$AnsDemon = \{ \underline{NodeInd}, AnsDemonInd \} \quad (3.53)$$

$$dom(NodeInd) = \{1, \dots, NodeQ\};$$

$$dom(AnsDemonInd) = \{1, \dots, AnsDemonQ\}$$

$AnsDemonQ$ — общее количество различных процессов $AnsDemon$.

Можно показать, что Петри-модель, описанная в подразделе 3.3, является частным случаем реляционной модели. В 3.3.1 рассмотрена сфера применимости Петри-модели на примере трудно реализуемого диалогового метода, относящегося к проблематике психодиагностики. Диалоговый метод был вначале описан без использования памяти истории ответов (рис. 3.8), а, затем с памятью истории ответов, хранящей ответы, полученные на предыдущих шагах (рис. 3.9).

Проведём аналогичное исследование этого же примера, используя полученную сетевую модель. Отметим, что в описанной выше модели узел с условной навигацией (рис. 3.11) учитывает ответ, полученный только на одном из предыдущих шагах. Полагаем, что этим шагом является последний активный шаг. Рассмотрим два варианта сетевого описания *DiAM*: (1) диалоговый метод не использует память истории ответов и состоит только из узлов с прямой навигацией (узлов типа *Node1*) и целевых узлов (узлов типа *Node3*); (2) диалоговый метод использует память истории ответов и состоит из узлов всех трёх типов.

Напомним, что суть рассматриваемого примера трудно реализуемого диалогового метода для психодиагностического теста заключается в следующем. На первом шаге реактивному агенту передаётся вопрос Que_1 . На этот вопрос ожидается два ответа: Ans_1 , Ans_2 . На втором шаге, вне зависимости от полу-

ченного ответа, реактивному агенту передаётся вопрос Que_2 , на который также ожидается два ответа: Ans_3 , Ans_4 . На третьем шаге, вне зависимости от полученных ранее ответов, реактивному агенту передаётся вопрос Que_3 на который ожидаются ответы Ans_5 , Ans_6 . В зависимости от комбинации ответов на все три вопроса выбирается одно из восьми заключений. Будем считать, что в случае, когда на любой из вопросов получен нераспознаваемый ответ, вопрос повторяется. На рис. 3.13 приведено графическое изображение сети $DiAM$, которая состоит только из узлов типа $Node1$ и $Node3$.

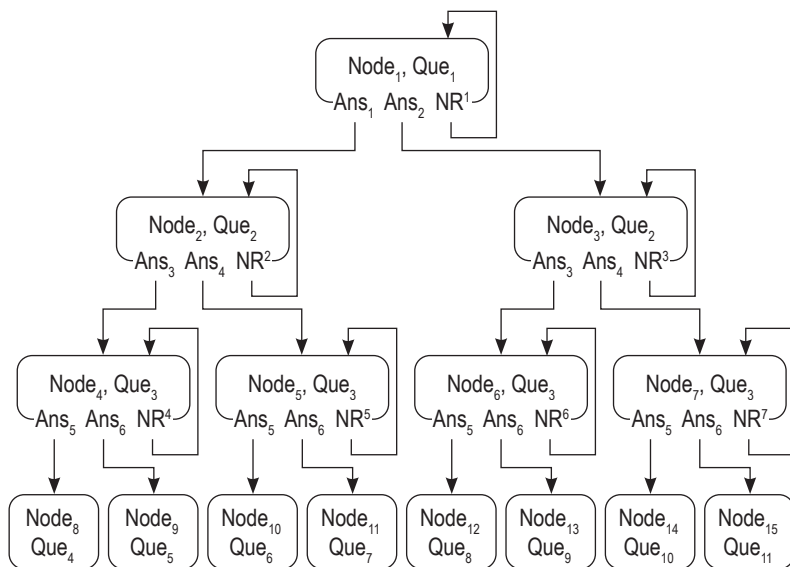


Рис. 3.13. Графическое представление сети $DiAM$ для примера, описанного в тексте. Сеть состоит только из узлов $Node1$ и $Node3$

В нижней части сети, изображённой на рис. 3.13, расположены целевые узлы. Поэтому $Que_4 - Que_{11}$ обозначают не вопросы, а утверждения (заключения). Утверждения формулируются в зависимости от комбинации ответов, полученных на вопросы

Que_1 , Que_2 и Que_3 . Так, заключение Que_4 соответствует следующей последовательности ответов: Ans_1 , Ans_3 , Ans_5 .

Рассмотрим реляционную базу, моделирующую сеть $DiAM$, изображенную на рис. 3.13. Поскольку в сети использованы только узлы с прямой навигацией и целевые узлы, для которых навигация отсутствует, для описания реляционной базы достаточно отношений $Step$ и $Navig1$.

На рис. 3.14 приведена реляционная база данных, описывающая сетевую модель $DiAM$, для рассматриваемого примера диалогового метода в виде отношений $Step$ и $Navig1$. Для упрощения принято, что процессы-демоны не используются.

Предположим теперь, что диалоговый метод использует память истории ответов, в которой хранится последний из полученных ответов и в сеть включены узлы с условной навигацией. Как и прежде, полагаем, что в случае, когда на вопрос получен нераспознаваемый ответ, вопрос повторяется. Однако с целью упрощения рисунка ветви, соответствующие нераспознаваемым ответам, не будем изображать. На рис. 3.15 приведено графическое изображение сети $DiAM$, состоящей из узлов типов $Node1$, $Node2$ и $Node3$.

Сеть диалогового метода, изображенная на рис. 3.15, состоит из одного узла с прямой навигацией (узел $Node_1$), трёх узлов с условной навигацией (узлы $Node_2$, $Node_3$, $Node_4$) и восьми целевых узлов (узлы $Node_5$ — $Node_{12}$).

Ветви, исходящие из узлов с условной навигацией, помечены значениями ответов, находящихся в памяти истории ответов. Как отмечалось, рассматриваем память истории ответов, в которой хранится только *один ответ, полученный на предыдущем шаге диалога*.

На рис. 3.16 приведена реляционная база данных, описывающая сетевую модель $DiAM$, для рассматриваемого примера диалогового метода в виде отношений $Step$, $Navig1$ и $Navig2$. Как и прежде, принято, что процессы-демоны не используются.

Резюмируем рассмотренные варианты конструирования трудно реализуемого диалогового метода. Существует прин-

Step			
NodeInd	NodeType	QueInd	NNodeInd
1	Node1	Que ₁	1
2	Node1	Que ₂	2
3	Node1	Que ₂	3
4	Node1	Que ₃	4
5	Node1	Que ₃	5
6	Node1	Que ₃	6
7	Node1	Que ₃	7
8	Node3	Que ₄	отсутствует
9	Node3	Que ₅	отсутствует
10	Node3	Que ₆	отсутствует
11	Node3	Que ₇	отсутствует
12	Node3	Que ₈	отсутствует
13	Node3	Que ₉	отсутствует
14	Node3	Que ₁₀	отсутствует
15	Node3	Que ₁₁	отсутствует

Navig1		
Node1Ind	CurrAns	NextNodeInd
1	Ans ₁	2
1	Ans ₂	3
2	Ans ₃	4
2	Ans ₄	5
3	Ans ₃	6
3	Ans ₄	7
4	Ans ₅	8
4	Ans ₆	9
5	Ans ₅	10
5	Ans ₆	11
6	Ans ₅	12
6	Ans ₆	13
7	Ans ₅	14
7	Ans ₆	15

Рис. 3.14. Реляционная база данных для сетевой модели *DiAM*, изображенной на рис. 3.13

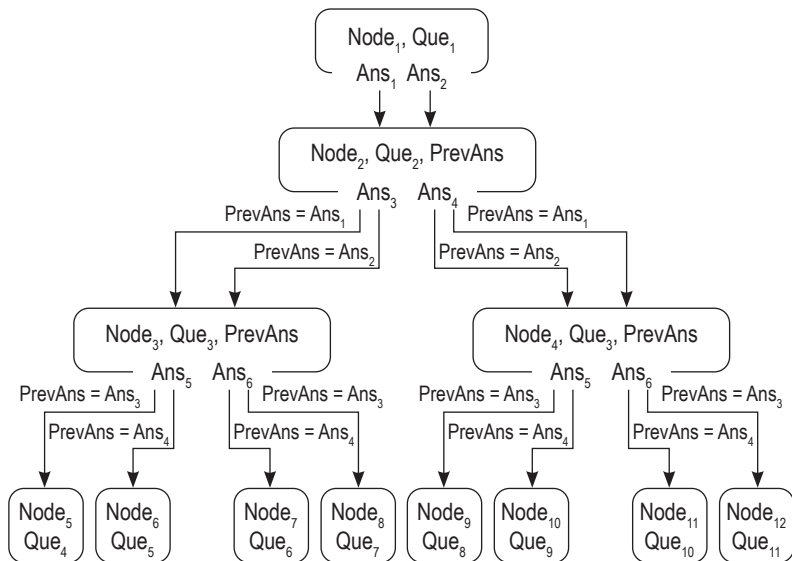


Рис. 3.15. Графическое представление сети *DiAM* для примера, описанного в тексте. Сеть состоит из узлов типов *Node1*, *Node2* и *Node3*

ципиальная возможность двойного описания одного и того же трудно реализуемого диалогового метода как с использованием памяти истории ответов и узлов с условной навигацией, так и без использования узлов этого типа. В случае, когда сеть *DiAM* строится без использования узлов с условной навигацией, в сети необходимо формировать уникальные, непересекающиеся траектории, соединяющие начальный узел и целевые узловы элементы. В случае, когда сеть трудно реализуемого диалогового метода строится с использованием истории ответов и узлов с условной навигацией, общее количество узлов и особенно узлов, предполагающих навигацию (узлов типа *Node1* и *Node1*), резко уменьшается. В рассмотренных примерах сокращение общего количества узлов составило примерно 20%. Однако учёт истории ответов и использование узлов с условной навигацией *не упрощает реляционную базу данных*.

Step			
NodeInd	NodeType	QueInd	NNodeInd
1	Node1	Que ₁	1
2	Node2	Que ₂	2
3	Node2	Que ₃	3
4	Node2	Que ₃	4
5	Node3	Que ₄	отсутствует
6	Node3	Que ₅	отсутствует
7	Node3	Que ₆	отсутствует
8	Node3	Que ₇	отсутствует
9	Node3	Que ₈	отсутствует
10	Node3	Que ₉	отсутствует
11	Node3	Que ₁₀	отсутствует
12	Node3	Que ₁₁	отсутствует

Navig1		
Node1Ind	CurrAns	NextNodeInd
1	Ans ₁	2
1	Ans ₂	2

Navig2				
Node2Ind	CurrAns	PathMemInd	PrevAns	NextNodeInd
2	Ans ₃	1	Ans ₁	3
2	Ans ₃	1	Ans ₂	3
2	Ans ₄	1	Ans ₁	4
2	Ans ₄	1	Ans ₂	4
3	Ans ₅	1	Ans ₃	5
3	Ans ₅	1	Ans ₄	6
3	Ans ₆	1	Ans ₃	8
3	Ans ₆	1	Ans ₄	7
4	Ans ₅	1	Ans ₃	9
4	Ans ₅	1	Ans ₄	10
4	Ans ₆	1	Ans ₃	12
4	Ans ₆	1	Ans ₄	11

Рис. 3.16. Реляционная база данных, описывающая сетевую модель *DiAM*, изображенную на рис. 3.15

Эффективным инструментом, позволяющим избежать большого количества шагов в трудно реализуемых диалоговых методах и упростить реляционную базу данных является использование процессов демонов *AnsDemon*, обрабатывающих ответы и определяющих индексы последующих узлов. Недостатком использования этого инструмента является необходимость участия профессионала программиста в создании и редактировании диалогового метода, что исключает его создание силами автора (психолога, преподавателя, проектировщика). Если при конструировании диалогового метода не используются процессы-демоны, то создание и редактирование как диалогового метода, так и всего сценария диалога может выполнять автор диалогового метода при помощи полноэкранного редактора базы данных.

3.4.2. Диалогическое представление вопроса

Диалоговые методы решения проблем часто предполагают использование комбинации вербальных и невербальных элементов при конструировании вопросов. Примерами могут быть диалоговые методы, разрабатываемые с целью обучения или психодиагностики. С точки зрения когнитивной психологии такие комбинированные вопросы могут рассматриваться как *интеррогативные стимулы*. В ходе реализации диалогового метода на каждом его шаге активный агент диалога передает реактивному агенту интеррогативный стимул и анализирует полученную реакцию реактивного агента. В зависимости от полученной реакции выбирается последующий шаг и реактивному агенту передается последующий интеррогативный стимул. Поэтому понятие «вопрос» целесообразно трактовать в широком смысле как интеррогативный стимул, состоящий из вербальной и невербальной информации. Полагаем, что *интеррогативный стимул представлен набором видео- и аудио-объектов*.

Когнитивный цикл вопросно-ответного диалога (рис. 3.2) позволяет также предположить, что интеррогативный стимул

включает все ответы/реакции реактивного агента, ожидаемые на этот стимул-вопрос.

Даталогическое представление интеррогативного стимула и ассоциированных с ним реакций будем называть сценой. Реактивный агент воспринимает сцену при помощи зрительно-го и слухового анализаторов. Понятие «сцена» может рассматриваться как обобщение понятия «шаг». На рис. 3.5 выделен шаг диалога, включающий вопрос и ожидаемые ответы.

В данном подразделе рассмотрены основные идеи даталогической модели сцены, ориентированной на использование в сценариях искусственного активного агента вопросно-ответного диалога. Сцена моделируется цепью видео- и аудио-объектов, интерпретируемых через видео- и аудиоканалы компьютера. Таким образом, любой логический компонент сцены (интеррогативный стимул и ожидаемые реакции) моделируются участком цепи объектов сцены. На рис. 3.17 графически представлено принятое соглашение о логической структуре сцены.

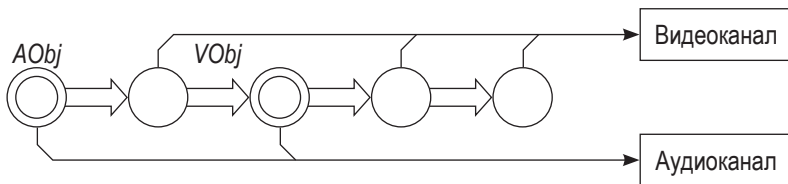


Рис. 3.17. Сцена, представленная цепью видео- ($VObj$) и аудио-объектов ($AObj$)

На рис. 3.17 сцена изображена в виде цепи, состоящей из пяти объектов, последовательно интерпретируемых через видео- и аудиоканалы компьютера. Объекты сцены проиндексированы в пределах каждой сцены, их интерпретация осуществляется последовательно, начиная с первого. Поскольку для формирования одного и того же логического компонента сцены могут использоваться и видео-, и аудио-объекты (напри-

мер, синхронный вывод звука и изображения при интерпретации интеррогативного стимула), то работа видеоканала должна быть согласована с работой аудиоканала. Согласование каналов предполагает наличие набора правил, позволяющих разработчику сцены правильно располагать объекты в цепи, составляющей сцену.

Любой объект сцены будем описывать совокупностью атрибутов, используемых при его интерпретации через видео- и аудиоканалы компьютера. Однако представление объекта сцены исключительно в виде набора атрибутов возможно лишь для простейших объектов. В случае сложных объектов атрибутивное описание необходимо дополнять данными, которые невозможно или нерационально разлагать на атрибуты. Такое описание будем называть неатрибутивным. Таким образом, любой объект сцены может быть представлен двумя типами описаний: (1) *атрибутивное описание* и (2) *неатрибутивное описание*.

Неатрибутивное описание объекта сцены соответствует информации, поступающей непосредственно от его *прототипа* [57]. Под прототипом будем понимать некий внешний объект (например, фотографическое изображение), используемый при конструировании видео-объекта сцены.

Атрибутивное описание формируется в результате *интеллектуальной деятельности автора сцены* при синтезе объекта сцены.

Исходная информация о прототипе видео-объекта представлена в виде светового потока. Для формирования неатрибутивного описания видео-объекта необходимо устройство, преобразующее световой поток в набор данных, характеризующих цвет и яркость каждой точки. На рис. 3.18 приведена диаграмма, иллюстрирующая процесс формирования неатрибутивного описания и его последующую интерпретацию в видео-объект сцены.

Для того чтобы неатрибутивное описание видео-объекта имело конечный объём, необходимо, чтобы преобразователь выполнял дискретизацию светового потока по пространству, уровню и цвету. В этом случае преобразователь представляет

собой известные устройства типа сканер либо цифровая фото- или видеокамера. Ряд эффективных и практически пригодных методов формирования неатрибутивного описания объекта сцены предложено в [58]. Не рассматривая методы формирования неатрибутивного описания, будем считать, что в даталогическом описании видео-объекта сцены этот тип описания представлен в виде цифрового файла соответствующего формата.

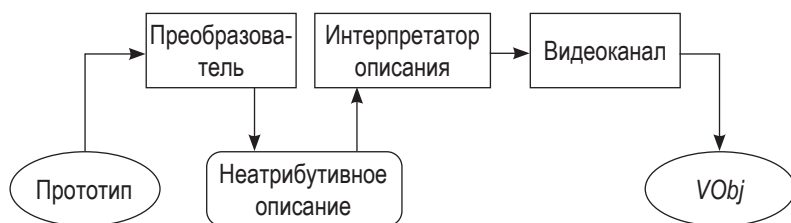


Рис. 3.18. Формирование и интерпретация неатрибутивного описания видео-объекта сцены

Атрибутивное описание объекта сцены формируется автором сцены. Атрибутивное описание видео-объекта может формироваться в результате восприятия существующего прототипа, его категоризации и последующей интеллектуальной обработки воспринятого ментального образа. Атрибутивное описание любого объекта сцены может формироваться автором сцены в результате творческой деятельности и на основании экспертных знаний предметной области диалогового процесса. Смысл интеллектуальной обработки воспринятого прототипа видео-объекта заключается в упрощении ментального образа представлением его в виде набора атрибутов. На рис. 3.19 приведена диаграмма, иллюстрирующая процесс формирования атрибутивного описания видео-объекта и его последующей интерпретации.

Ясно, что задача формирования атрибутивного описания объекта сцены решается неоднозначно и определяется автором сцены. Можно получить сколь угодно много атрибутивных

описаний одного и того же прототипа в зависимости от того, например, с какой целью составляется описание.

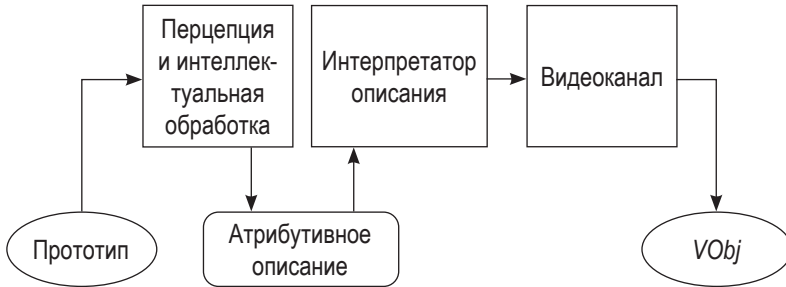


Рис. 3.19. Формирование и интерпретация атрибутивного описания видео-объекта сцены

Систему атрибутивного описания объекта сцены можно рассматривать как некую *систему классификации*, соответствующую архитектуре системы, которая интерпретирует атрибутивное описание видео-объекта и воспроизводит его на экране монитора компьютера. Как уже отмечалось, генерация вопроса-стимула осуществляется стандартным циклическим процессом *DiMC*. Будем считать, что стандартный циклический процесс *DiMC* строится, исходя из следующих соглашений.

1. Видео-объект представляет собой прямоугольник, который может иметь разные размеры (от одного пикселя до размера экрана) и отображаться в различных участках монитора.
2. После того, как процесс отображения видео-объекта на экране монитора завершён, он может подвергаться различным топологическим преобразованиям, например поступательному перемещению.
3. Текст и неатрибутивное описание отображаются внутри видео-объекта.
4. Видео-объект является «окном», через которое просматривается неатрибутивное описание.

5. Видео-объект предназначен как для визуальной презентации компонентов сцены, так и для связи процессов *AnsDemon* и *QueDemon* со сценой.
6. Внешним процессам *AnsDemon* и *QueDemon* доступна память истории ответов *PathMem* а также все информационные ресурсы компьютера.

Местоположение видео-объекта на экране задаётся соответствующими атрибутивными описателями. Из всех объектов, составляющих сцену, только один, интерпретируемый в данный момент, является активным, а остальные рассматриваются как фон.

Имеется только один буфер, обслуживающий активный объект во время его интерпретации. Та часть буфера, которая может «просматриваться» через каркас видео-объекта, задаётся соответствующими атрибутивными описателями.

Все ответы реактивного агента запоминаются в памяти истории ответов. В зависимости от того, как описан шаг в диалоговом методе *DiAM*, текущий ответ реактивного агента может либо непосредственно трансформироваться в идентификатор следующей сцены, либо опосредованно при помощи процесса *AnsDemon*.

Предложенное понимание атрибутивного описания видео-объекта сцены и принятые соглашения об организации стандартного циклического процесса *DiMC* позволяют сформировать множество классов атрибутивного описания видео-объектов. На рис. 3.20 приведена диаграмма классов, которая моделирует один из возможных вариантов структуры сцены в соответствии с принятыми соглашениям.

На рис. 3.20 использованы следующие обозначения:

Scene — класс сцен диалогового сценария.

VObj — класс видео-объектов сцены.

Attributed — класс атрибутивных описателей видео-объекта.

NonAttr — класс неатрибутивных описателей видео-объекта.

VisualAttr — класс атрибутивных описателей визуальных элементов видео-объекта.

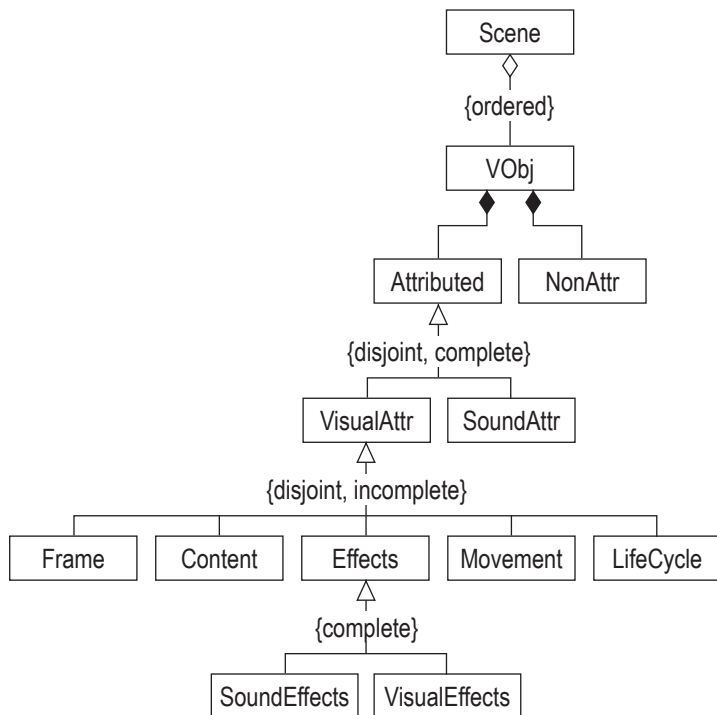


Рис. 3.20. Классы атрибутивных описателей видео-объекта сцены

SoundAttr — класс атрибутивных описателей звуковых элементов видео-объекта.

Frame — класс атрибутивных описателей каркаса видео-объекта.

Content — класс атрибутивных описателей содержимого видео-объекта.

Effects — класс атрибутивных описателей визуальных и звуковых эффектов, сопровождающих интерпретацию видео-объекта.

Movement — класс атрибутивных описателей поступательного перемещения видео-объекта.

LifeCycle — класс атрибутивных описателей «времени жизни» видео-объекта.

SoundEffects — класс атрибутивных описателей звуковых эффектов, сопровождающих интерпретацию видео-объекта.

VisualEffects — класс атрибутивных описателей визуальных эффектов, сопровождающих интерпретацию видео-объекта.

В общем случае умозрительно видео-объект сцены (экземпляр класса *VObj*) может представлять собой произвольную, плоскую фигуру, ограниченную замкнутой линией. Этого достаточно для формирования любого двухмерного или трёхмерного изображения. Однако сцены большинства прототипов прикладных вопросно-ответных диалоговых систем представлены композицией перекрывающихся прямоугольников [8, 59–62]. Поэтому полагаем, что видео-объект сцены представляет собой прямоугольник произвольного размера, размещаемый в произвольной точке экрана монитора. Видео-объекты сцены генерируются последовательно и на протяжении некоторого отрезка времени формируют сцену. Последующий видео-объект может располагаться в произвольной точке экрана монитора, в том числе перекрывать предыдущий объект. Рис. 3.21 иллюстрирует отмеченный способ взаимного расположения объектов сцены.

Как показано на диаграмме классов (рис. 3.20), класс видео-объектов сцены (класс *VObj*) состоит из двух классов, моделирующих структурно разнородные группы описателей: атрибутивные описатели видео-объекта сцены (класс *Attributed*) и неатрибутивные описатели видео-объекта сцены (класс *NonAttr*).

Любой видео-объект может быть описан совокупностью атрибутов, характеризующих:

- геометрию видео-объекта и его расположение на сцене (класс *Frame*);
- текстовое или графическое содержимое буфера, которое отображается в видео-объекте (класс *Content*);

- специфику видео- и аудио-элементов объекта (классы *VisualAttr* и *SoundAttr*);
- «поведение» объекта на сцене (классы *Effects* и *Movement*);
- длительность существования видео-объекта на сцене (класс *LifeCycle*).

Некоторые атрибуты являются обязательными. К таким атрибутам относятся, например, атрибуты классов *Frame*, *Content* и *LifeCycle*. Атрибуты, описывающие «поведение» объекта на сцене, являются опцией. Они нужны для конструирования динамической сцены.

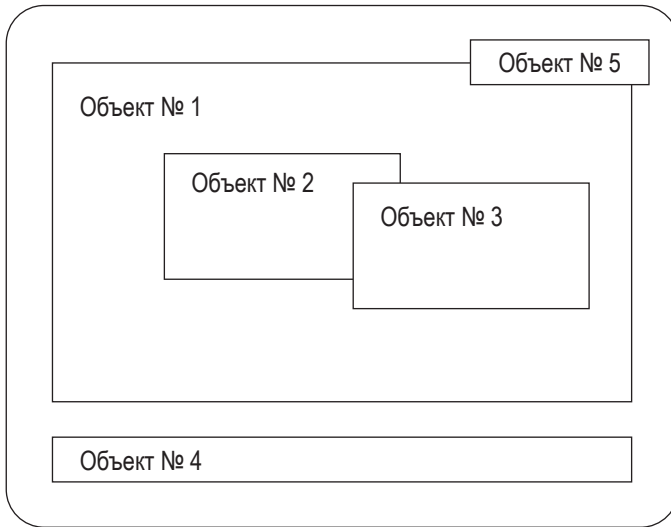
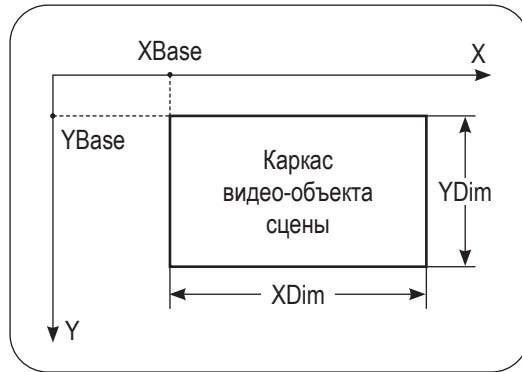


Рис. 3.21. Пример взаимного расположения видео-объектов на сцене

Формирование подробного списка атрибутов классов на рис 3.20 является предметом проектирования программы искусственного активного агента вопросно-ответного диалога. В качестве примера приведем только примерный набор атрибутов для некоторых классов нижнего уровня.

Атрибуты класса *Frame*

Каркасом видео-объекта сцены будем называть прямоугольную рамку с указанием ширины, высоты и координаты базовой точки. В качестве базовой точки выбран левый верхний угол прямоугольника. Рис. 3.22 иллюстрирует идею каркаса и атрибуты класса *Frame*, описывающие каркасы видео-объектов.



№ п.п.	Наименование	Интерпретация
1	XBase	Горизонтальная координата базовой точки
2	YBase	Вертикальная координата базовой точки
3	XDim	Размер по горизонтали
4	YDim	Размер по вертикали

Рис. 3.22. Каркас видео-объекта сцены и атрибуты класса *Frame*

Атрибуты класса *VisualEffects*

Презентация информации внутри каркаса видео-объекта может быть одномоментной и статической. Однако часто

требуется динамическая презентация в течение некоторого времени, необходимого для её восприятия. Например, последовательный вывод текстовой информации в каркас видео-объекта может упростить её восприятие реактивным агентом. Атрибуты класса *VisualEffects* (атрибуты визуальных эффектов) описывают различные способы презентации информации внутри каркаса. Вариацией этих атрибутов можно получить различные динамические эффекты при интерпретации видео-объекта сцены, которые, как показывает опыт, при их правильном использовании могут упростить и ускорить восприятие как самой сцены, так и её отдельных объектов.

Как уже отмечалось, каркас видео-объекта является «окном», позволяющим просматривать буфер, содержащий неатрибутивное описание видео-объекта, которое может быть, в частности, текстом. В таблице на рис. 3.23 приведены возможные атрибуты визуальных эффектов, а рис. 3.24 иллюстрирует их действие при интерпретации сцены.

Атрибуты класса *Movement*

После завершения интерпретации содержимого каркаса в соответствии с атрибутами, приведенными в таблице на рис. 3.23, видео-объект сцены может совершать поступательное перемещение по экрану. Такое его перемещение позволяет создавать динамические сцены.

Динамические сцены могут использоваться при создании сценариев диалоговых методов, направленных на решении проблем обучения и передачи знаний, когда, например, для облегчения восприятия сцены, содержащей обучающий интеррогативный стимул, перемещение очередного видео-объекта используется для организации паузы, необходимой для восприятия обучающего материала [57]. Рис. 3.25 иллюстрирует перемещение, задаваемое атрибутами класса *Movement*, и описание атрибутов этого класса.

№ п.п.	Группа	Наименование	Интерпретация
5	Начальное позиционирование каркаса	XPos	Начальное позиционирование каркаса на буфер по горизонтали.
6		YPos	Начальное позиционирование каркаса на буфер по вертикали.
7		XStep	Шаг позиционирования по горизонтали. Каркас перемещается по буферу в горизонтальном направлении пошагово на расстояние кратное Xstep.
8		YStep	Шаг позиционирования по вертикали (см. объяснения для XStep).
9	Моды начального заполнения каркаса	InitMod1	«Посимвольно» через заданный промежуток времени. «Символ» имеет размеры Xstep, Ystep.
10		InitMod2	«Построчно» через заданный промежуток времени. «Строка» состоит из символов. Длина строки равна длине каркаса.
11		InitSpeed	Промежуток времени, используемый в InitMod1 или InitMod2.
12		InitPause	Выдержка времени после первоначального заполнения каркаса. [В некоторых случаях нужна для заполнения буфера.]
13	Моды перемещения содержимого каркаса после его начального заполнения	MoveMod1	Вверх через каркас до «дна» буфера с шагом в 1 пиксель.
14		MoveMod2	Вниз через каркас до «верха» буфера с шагом в 1 пиксель.
15		MoveMod3	Непрерывное перемещение — «барабан» вверх с шагом в 1 пиксель.
16		MoveMod4	Непрерывное перемещение — «барабан» вниз с шагом в 1 пиксель.
17		MoveMod5	Не перемещать содержимое.
18		MoveCycle	Количество циклов для MoveMod3 и MoveMod4.

Рис. 3.23. Атрибуты класса *VisualEffects*

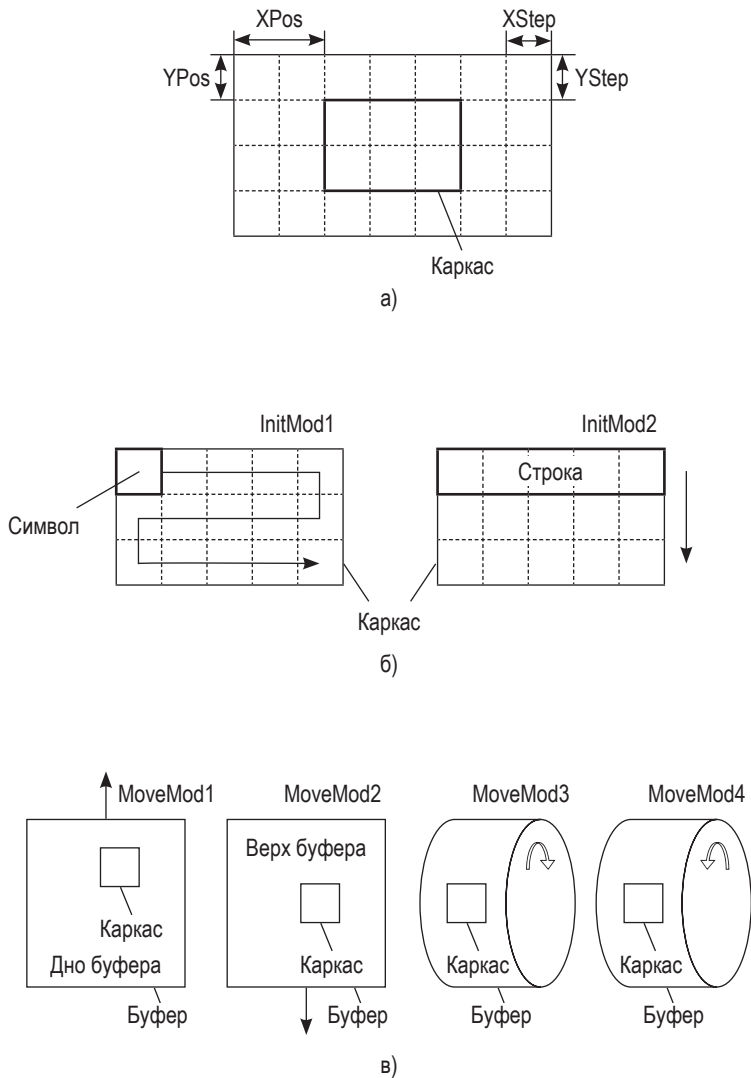
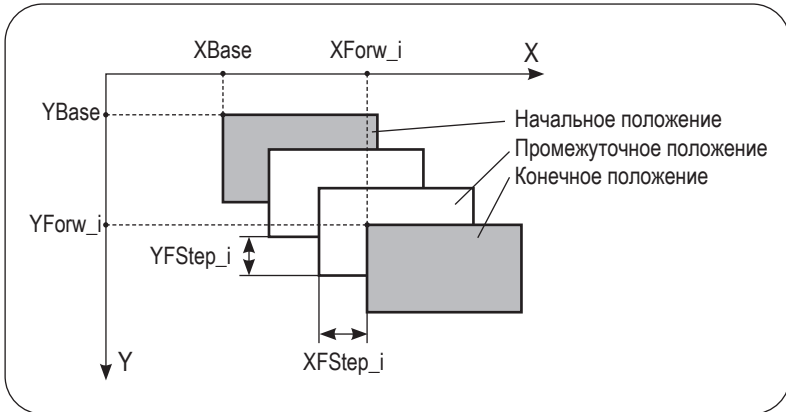


Рис. 3.24. Действие атрибутов визуальных эффектов:
 а – начальное позиционирование каркаса,
 б – начальное заполнение каркаса,
 в – перемещение содержимого каркаса после его начального
 заполнения



№ п.п.	Наименование	Интерпретация
19	XForw_i	Горизонтальная координата конечного положения VObj для i-го отрезка траектории.
20	YForw_i	Вертикальная координата конечного положения VObj для i-го отрезка траектории.
21	XStep_i	Шаг перемещения по горизонтали для i-го отрезка траектории.
22	YStep_i	Шаг перемещения по вертикали для i-го отрезка траектории.
23	ForwVis_i	Признак «визуализации траектории перемещения». Возможны два значения: – сохранять начальное и все промежуточные положения; – не сохранять начальное и все промежуточные положения.
24	ForwPause_i	Выдержка времени после перемещения в конечное положение. [Эта выдержка может использоваться исходя из «логики презентации» сцены, а также специально для заполнения входного буфера].

Рис. 3.25. Перемещение видео-объекта и атрибуты класса *Movement*

Траектория поступательного перемещения видео-объекта сцены может состоять или из одного отрезка, соединяющего начальное и конечное положения базовой точки, или из нескольких последовательно соединённых отрезков. Если траектория поступательного перемещения видео-объекта состоит из нескольких отрезков, то каждый отрезок траектории должен описываться отдельно.

Атрибуты класса *LifeCycle*

После завершения интерпретации видео-объекта на сцене (включая возможные визуальные эффекты и поступательное перемещение) он может либо сохраняться в пределах сцены неопределённо долго, либо исчезнуть через заданный промежуток времени. Атрибуты «времени жизни» класса *LifeCycle* предназначены для управления длительностью нахождения объекта на сцене. Эти атрибуты приведены в таблице на рис. 3.26.

№ п.п.	Наименование	Интерпретация
25	ObjPause	Выдержка времени после полной интерпретации видео-объекта на сцене.
26	DelAfterPause	Удаление объекта со сцены после завершения интерпретации и выдержки времени ObjPause. Принимает два значения: – удалить после паузы; – не удалить после паузы.
27	DelAfterStep	Удаление объекта со сцены после завершения интерпретации всех её объектов. Принимает два значения: – удалить после генерации сцены; – не удалить после генерации сцены.

Рис. 3.26. Атрибуты класса *LifeCycle*

Понятно, что при конструировании видео-объекта сцены не обязательно использовать все его атрибуты. Очевидно, что

обязательными являются, например, атрибуты каркаса и атрибуты «времени жизни», однако не обязательно, чтобы каждый видео-объект использовал атрибуты визуальных эффектов и участвовал в поступательном перемещении.

3.4.3. Искусственный активный агент диалога

Слово «диалог» можно понимать в узком и широком смысле. В узком смысле слово «диалог» обозначает логически связанную последовательность вербальных вопросно-ответных пар (вопросно-ответных транзакций), генерируемых активным (спрашивающим) и реактивным (отвечающим) агентами диалога. Классическими примерами диалога в узком смысле являются диалоги Платона [1]. Первый и второй разделы нашей книги посвящены, главным образом, исследованию диалога в узком смысле этого слова.

Диалог в широком смысле следует рассматривать с бихевиористической точки зрения как логически связанную последовательность пар стимул-реакция. Стимул генерируется активным агентом диалога, воздействует на зрительный и слуховой анализаторы реактивного агента диалога и воспринимается последним как интеррогатив. Реакция реактивного агента имеет статус ответа. Стимулы и реакции, в общем случае, являются комбинацией образной и вербальной формами представления информации. Подраздел 3.4, посвященный сетевой даталогической модели диалога, рассматривает диалог в широком смысле этого слова.

Как неоднократно отмечалось, диалог, как в узком, так и в широком смысле, является средством решения некоторых проблем. Диалоговый метод, названный методом Сократа, иллюстрирует использование вопросно-ответного диалога (диалога в узком смысле) как средства для решения проблем [8]. Психодиагностические тесты, рассматриваемые как диалоги в широком смысле, часто используют невербальные диагностические стимулы и предполагают вербальную реакцию. Примером мо-

жет служить метод прогрессивных матриц Равена, используемый для тестирования интеллектуальных способностей человека [54].

Проблемы, которые могут быть решены в ходе диалогового процесса, относятся к классу плохо формализуемых. Впервые на необходимость разработки методов решения плохо формализуемых проблем обратил внимание Ньюэлл [64]. Ньюэлл считал, что к плохо формализуемым относятся проблемы, обладающие одной или несколькими из следующих характеристик:

- проблемы не могут быть заданы в числовой форме;
- цели не могут быть выражены в терминах точно определённых целевых функции;
- не существует алгоритмического решения проблемы;
- алгоритмическое решение существует, но его нельзя использовать в связи с ограниченностью компьютерных ресурсов.

В прикладных системах искусственного интеллекта рассматривается ряд систем, функционирование которых направлено на решение плохо формализуемых проблем. Одним из наиболее обширных классов таких систем являются интеллектуальные системы обучения [8]. Проблема обучения и передачи знаний является плохо формализуемой и может решаться диалоговыми методами.

В области теории и практики диалоговых процессов важным является вопрос о том, каким образом можно *создавать, хранить и многократно воспроизводить диалоговые методы*. В целенаправленном диалоговом процессе, в ходе которого осуществляется решение плохо формализуемой проблемы, ключевую роль играет активный агент, поскольку именно он — носитель метода решения задачи.

На рис. 3.27 приведены компоненты обобщённой структуры искусственного активного агента вопросно-ответного диалога, которая синтезирована на основе результатов, полученных в подразделе 3.4.

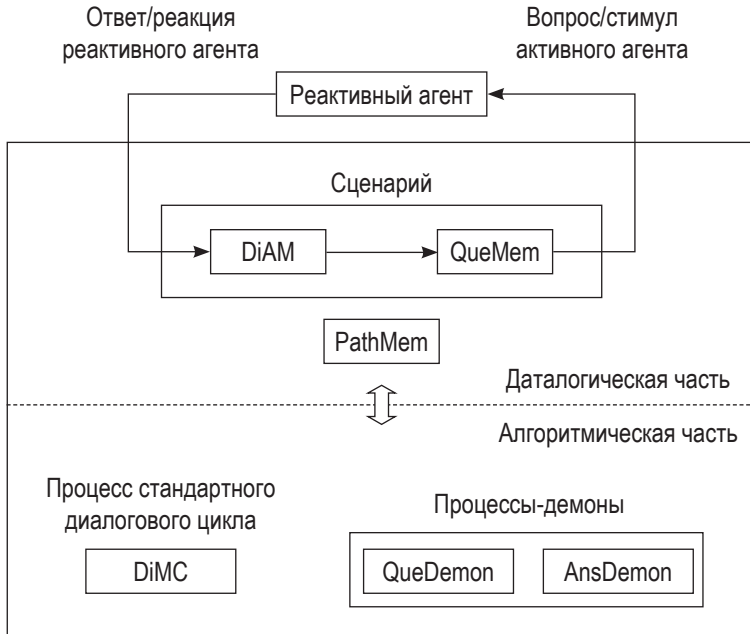


Рис. 3.27. Структура системы, моделирующей искусственного активного агента диалога

В системе на рис. 3.27 выделены даталогическая и алгоритмическая части.

Метод решения плохо формализуемой проблемы хранится в сценарии вопросно-ответного диалога и описан не алгоритмически, а даталогически. Данный сценарий можно рассматривать как трансформатор, преобразующий текущий ответ/реакцию реактивного агента в последующий вопрос/стимул активного агента. Сценарий включает память вопросов-сцен *QueMem* и диалоговый метод доступа к памяти вопросов *DiAM*. Память вопросов-сцен содержит даталогические описатели всех интеррогативных сцен, необходимых для реализации диалогового метода. Логика диалогового метода решения проблемы воплощена в *DiAM*.

Последовательность ответов реактивного агента запоминается в памяти ответов *PathMem*.

Память вопросов-сцен, диалоговый метод доступа к памяти вопросов-сцен, а также память ответов отображаются в данные и представляют собой диалогическую часть системы.

Реализация диалогового метода решения проблемы осуществляется пошагово (один шаг соответствует одной вопросно-ответной транзакции) при помощи «машины диалога». Ключевым компонентом «машины диалога» является стандартный циклический процесс *DiMC*, интерпретирующий сценарий диалога. Алгоритм *DiMC* на каждой итерации воспроизводит когнитивный цикл вопросно-ответного диалога, описанный в подразделе 3.1 (см. рис 3.2).

Вторым компонентом «машины диалога» является набор процессов-демонов. Итерации, выполняемые *DiMC*, предусматривают возможность вызова процесса-демона, если интерпретация сценария требует действий, выходящих за рамки стандарта, задаваемого структурой дескрипторов вопросов-сцен и стандартным диалоговым циклом *DiMC*. Предполагается использование двух типов процессов-демонов, отличающихся назначением.

Процесс *QueDemon* прерывает стандартный цикл *DiMC* на этапе интерпретации сцены и реализует нестандартный способ формирования сцены. Процесс *AnsDemon* прерывает стандартный цикл *DiMC* на этапе формирования индекса последующей сцены и реализует нестандартный способ формирования этого индекса. Процессы-демоны не обязательно участвуют в процессе диалогового решения проблемы, но если они используются, то к началу реализации диалогового метода должны присутствовать коды всех процессов-демонов, необходимых для интерпретации сценария. Стандартный процесс *DiMC* и набор процессов-демонов представляют собой алгоритмическую часть искусственного активного агента вопросно-ответной диалоговой системы.

4. КОГНИТИВНЫЕ ТЬЮТОРЫ

Проблемно-независимые модели вопросно-ответного диалога, описанные в подразделе 3.4, и модель искусственного активного агента вопросно-ответного диалога, изображённая на рис. 3.27, могут использоваться в качестве «оболочки» для разработки, хранения и многократного воспроизведения диалоговых методов решения плохо формализуемых проблем в области обучения или передачи знаний. Метод обучения или способ передачи знаний в процессе диалогового общения искусственного активного агента диалога с учеником является *проблемно-зависимым* и отображается в сценарии, а также в алгоритмах процессов-демонов.

Будем называть диалоговую систему, реализующую какой-либо метод обучения или передачи знаний *тьюторской системой*, поскольку диалоговое поведение активного агента диалога такой системы имитирует поведение тьютора, осуществляющего индивидуальную и персонализированную работу с учеником.

Диалоговое поведение *идеальной тьюторской системы* опишем рядом характеристик. Идеальная тьюторская система в процессе диалогового взаимодействия с учеником выполняет следующие действия:

1. Непрерывно адаптируется к текущим знаниям ученика и степени понимания им учебного материала.

2. Выбирает учебный материал, для которого имеет место несоответствие между текущими и требуемыми знаниями и/или степени понимания ученика.
3. Автоматически или по запросу ученика дополняет учебный материал тематически связанной информацией (например, информацией справочного характера); уделяет больше внимания «трудным местам» и пропускает уже известный материал.
4. Тестирует ученика и определяет его когнитивные характеристики и когнитивный профиль.
5. Проводит процесс обучения методом, соответствующим текущим когнитивным способностям и когнитивному профилю ученика.

Поскольку идеальная тьюторская система адаптируется к когнитивным характеристикам обучаемого, будем называть её *когнитивным тьютором*, следуя терминологии, введенной Джоном Андерсоном [65].

4.1. Системы коммуникации знаний

Очевидно, что поведение когнитивного тьютора, обеспечивающее его специфические характеристики, возможно только при наличии постоянного и целенаправленного диалогового взаимодействия между тьютором и учеником. В ряде публикаций, подчеркивая коммуникационный характер поведения тьюторских систем, такие системы называют *системами коммуникации знаний* [8]. Коммуникация знаний определяется, как способность инициировать и поддерживать целенаправленный обмен знаниями между двумя носителями знаний (тьютором и учеником) при помощи ограниченного набора коммуникационных ресурсов. Акцент на коммуникативные свойства тьюторских систем означает фокусировку внимания на структурном понимании диалогового процесса обмена знаниями.

В данном подразделе приведен обзор ряда классических систем коммуникации знаний, которые в разной степени реализу-

ют поведение идеальной тьюторской системы. Рассмотренные ниже примеры систем коммуникации знаний распределены между тремя классами: (1) интеллектуальные тьюторские системы; (2) обучающие среды с интеллектуальной поддержкой; (3) интеллектуальные репетиторы.

4.1.1. Интеллектуальные тьюторы

Одна из первых, ныне классических, систем коммуникации знаний, относящихся к классу интеллектуальных тьюторских систем, — это система SCHOLAR, предложенная Джеймом Карбонеллом [66].

Существенной с точки зрения вопросно-ответного диалога характеристикой системы SCHOLAR является её способность поддерживать диалог, в котором *роли диалоговых агентов не зафиксированы*. Это означает, что в процессе обучения вопросы может задавать не только система, но и ученик. Система SCHOLAR поддерживает вербальный диалог и использует ограниченный английский язык. Генерируемые вопросы и фразы не произвольные, они создаются при помощи шаблонов, дополняемых информацией, получаемой при помощи внутренней модели. Процедура генерации не предполагает «понимания» системой SCHOLAR неверных ответов ученика и, следовательно, не может извлекать диагностическую информацию из них. На рис. 4.1 приведен оригинальный фрагмент протокола диалога ученика и SCHOLAR, который иллюстрирует как свободное распределение ролей агентов диалога, так и языковую ограниченность системы. Фразы, генерируемые SCHOLAR, записаны прописными буквами.

Следующей существенной характеристикой SCHOLAR, является способ представления знаний о предметной области (география Южной Америки) и экспертных знаний тьютора. В этой системе впервые в качестве среды для построения модели предметной области и экспертных знаний тьютора используется концепция семантической сети. Карбонелл возлагал большие надежды на семантические сети в качестве средства

представления знаний и ввёл понятие *модель идеальных знаний ученика* как синоним понятия модели предметной области в виде полной семантической сети. Пример фрагмента модели предметной области SCHOLAR приведен на рис. 4.2. [8].

APPROX WHAT IS THE AREA OF BRAZIL?

2,500,000 square miles.

WRONG. PLEASE INDICATE IF THE FOLLOWING STATEMENT IS CORRECT: THE AREA OF PARAGUAY IS APPROX 47,432 SQUARE MILES.

What is the area of Brazil?

THE AREA OF BRAZIL IS APPROX 3,600,000 SQUARE MILES. NOW ANSWER THE QUESTION YOU DIDN'T ANSWER BEFOR.

False.

REPHRASE YOUR STATEMENT. I CAN'T UNDERSTAND IT.

Incorrect.

VERY GOOD. THE AREA OF PARAGUAY IS APPROX 157,000 SQUARE MILES. THE LANGUAGE IN BRAZIL IS FRENCH. RIGHT OR WRONG?

Wrong.

YOU ARE CORRECT. THE LANGUAGE IN BRAZIL IS PORTUGUESE. IS IT TRUE THAT THE CAPITAL IN BRAZIL IS BRAZILIA?

Yes.

THAT'S FINE. APPROX WHAT IS THE POPULATION IN BRAZIL?

60,000,000.

APPROXIMATELY TRUE, BUT A MORE EXACT VALUE IS 70,000,000. WHAT IS BRAZIL?

Tell me about Peru.

. . .

Рис. 4.1. Фрагмент протокола диалога ученика и системы SCHOLAR

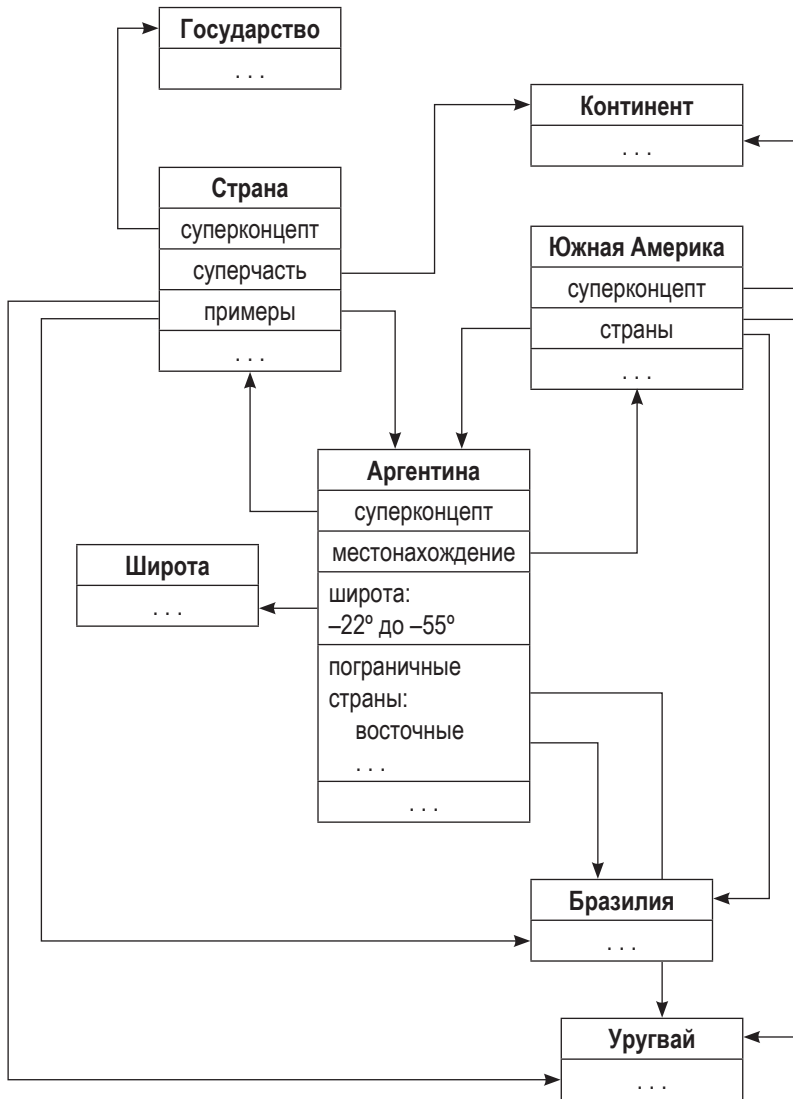


Рис. 4.2. Фрагмент семантической сети, моделирующей предметную область системы SCHOLAR

Для формирования ответов на вопросы ученика SCHOLAR используется механизм логических умозаключений. В первоначальной версии логические умозаключения строились путем навигации по узлам семантической сети. Например, система могла легко прийти к заключению, что Сантьяго находится в Южной Америке, поскольку узел «Сантьяго» связан с узлом «Чили» отношением «часть-целое», который, в свою очередь, связан тем же отношением с узлом «Южная Америка».

Метод обучения, используемый в первоначальной версии SCHOLAR, весьма прост и базируется на случайном выборе локальных тем, которые учитывают вопросы ученика.

Идеи первоначальной версии SCHOLAR получили дальнейшего развития в работах Коллинса и его коллег в направлении дальнейшего развития механизма вывода и диалогового метода обучения [67,68].

Развитие механизма вывода осуществлялось в направлении моделирования механизмов вывода, выполняемых человеком в условиях неопределённости или *правдоподобного вывода (plausible inference)*. Например, человек может сделать правдоподобное заключение о том, что на континенте не производится сахар из знаний о том, что ни в одной стране континента не производится сахар и что он располагает полным списком стран. Ясно, что получить такой вывод простой навигацией по узлам сети невозможно, так как невозможно включить в семантическую сеть все сведения о том, что не производится в Южной Америке. Среди прочих предложений по дальнейшему развитию механизма вывода SCHOLAR отметим технику вывода, названную авторами функциональным анализом, под которым понимался анализ фактов, от которых зависела некоторая гипотеза. Например, SCHOLAR получила возможность делать вывод о климате в регионе, анализируя его местоположение.

Развитие метода обучения осуществлялось в направлении замены процедуры случайного выбора локальных тем процедурой, действующей на основе набора эвристических правил, полученных из наблюдений за работой опытных тьюторов.

Тема, выбранная либо на основе отмеченных эвристических правил, либо по запросу ученика, динамически включалась в *план занятия (agenda)*, и ей выделялось время на обсуждение, пропорциональное её важности. Таким образом, *динамически редактируемый план занятий* становится главной управляющей структурой тьюторского диалога SCHOLAR.

Новая версия SCHOLAR, с точки зрения совершенствования метода обучения, получила также ограниченные возможности по разбору и обсуждению ошибок ученика.

Как следует из приведенного описания, метод обучения SCHOLAR обладает двумя существенными недостатками. Во-первых, он не опирается на достижения педагогической науки и игнорирует множество методов и методик обучения, предложенных педагогами. Во-вторых, метод не адаптируется к уже имеющимся знаниям и когнитивным характеристикам ученика.

Интеллектуальная тьюторская система WHY, предложенная Стивенем и Коллинсом, развивает основные идеи системы SCHOLAR в части моделирования предметной области и метода обучения, воплощённого в диалоговом процессе. В системе WHY впервые была сделана попытка компьютерной реализации диалогового метода обучения, известного как метод Сократа [61].

Метод Сократа является сугубо диалоговым и предполагает такой способ ведения вопросно-ответного диалога, когда ученик (реактивный агент), «ведомый» вопросами тьютора (активный агент), ставится в положение, когда он должен при помощи дедуктивных умозаключений сам выводить новые, неизвестные ему ранее, казуальные отношения между известными фактами предметной области.

Коллинс и его коллеги предположили, что метод Сократа может быть описан набором универсальных правил, определяющих диалоговое поведение тьютора. Коллинс и его коллеги предложили первоначально несколько десятков, а затем более полусотни правил, сформулированных в виде продукционных правил (условных предложений), которые определяют поведение тьютора, использующего метод Сократа, применительно к

изучению метеорологии (предметная область системы WHY) [69,70]. Каждое продукционное правило связывает один из ожидаемых текущих ответов ученика с последующим вопросом (обучающим стимулом) тьютора.

На рис. 4.3 приведено одно из продукционных правил, управляющих диалоговым поведением тьютора в системе WHY. Это правило иллюстрирует способ, при помощи которого тьютор стимулирует ученика обнаруживать ошибку в собственных рассуждениях.

При помощи этого правила тьютор и пытается вовлечь ученика в логическую «ловушку», которая заключается в том, что среди последовательности сделанных ранее заключений ученика имеются противоречивые.

ПРАВИЛО 6. Общее правило для несущественного фактора

ЕСЛИ ученик даёт объяснение одному
или нескольким несущественным факторам
ТО сформулируйте общее правило, утверждающее, что
данный фактор является существенным
И спросите ученика является ли это правило истинным

ОСНОВАНИЕ ДЛЯ ИСПОЛЬЗОВАНИЯ

Побудить ученика обращать внимание на остальные
казуальные факторы

ПРИМЕР ИСПОЛЬЗОВАНИЯ

Если ученик называет воду причиной выращивания риса
в Китае, спросите его «Считаете ли Вы, что в любом месте
с достаточным количеством воды может произрастать рис?»

Рис. 4.3. Пример продукционного правила, которое используется в системе WHY

На рис. 4.4 приведен фрагмент протокола диалога ученика и системы WHY, который иллюстрирует использование правила, приведенного на рис. 4.3. (четвёртый вопрос WHY).

WHY: DO YOU THINK THE AMAZON JUNGLE HAS HEAVY RAINFALL OR LIGHT RAINFALL?

(Rule 1: Ask for a prediction about a particular case.)

Student: Heavy rainfall

WHY: WHY DOES THE AMAZON HAVE HEAVY RAINFALL?

(Rule 2: Ask for prior cause.)

Student: Because it's near the Atlantic

WHY: YES, THE AMAZON JUNGLE IS NEAR THE EQUATORIAL ATLANTIC. HOW DOES THAT AFFECT RAINFALL IN THE AMAZON JUNGLE?

(Rule3: Ask for intermediate cause.)

Student: The water in the current is warm

. . .

WHY: DO YOU THINK THAT ANY PLACE WITH MOUNTAINS HAS HEAVY RAINFALL?

(Rule 6: Form a general rule for an insufficient cause.)

Student: Yes

WHY: SOUTHERN CALIFORNIA HAS MOUNTAINS. WHY DOESN'T SOUTHERN CALIFORNIA HAVE HEAVY RAINFALL?

(Rule 7: Pick a counter-example for an insufficient cause, and Rule 2. Ask for prior causes.)

. . .

Рис. 4.4. Фрагмент протокола диалога ученика и системы WHY

Как видно из этого фрагмента система WHY поддерживает диалог, в котором роли диалоговых агентов зафиксированы. Система WHY (интеллектуальный тьютор) выступает в роли активного агента, а ученик — в роли реактивного.

Знания о предметной области в WHY хранятся в виде *иерархии сценариев*. Общее определение сценария, предложен-

ное Коллинсом и его коллегами, определяет его как набор казуальных и темпоральных отношений между событиями, под которыми естественно понимать вопросы тьютора и ответы ученика. Таким образом, процессы генерации вопросов и оценки ответов в WHY задаются множеством сценариев, упорядоченных иерархически. Сценарии наиболее высокого уровня хранят знания в наиболее общей форме, а их детализация осуществляется сценариями более низкого уровня.

Предполагается, что обучение начинается со сценария, находящегося на наиболее высоком уровне. Сценарии WHY хранят знания о предметной области и стратегию обучения, но не используются для хранения метода диагностики ученика. Таким образом, WHY обладает ограниченными диагностическими возможностями, сводимыми к независимому анализу текущих ответов ученика.

Последующие исследования, направленные на дальнейшее приближение WHY к идеальной тьюторской системе, проводились по двум направлениям: (1) моделирование целенаправленного глобального обучающего поведения; и (2) расширение концепции сценарной модели предметной области.

Моделирование целенаправленного глобального обучающего поведения WHY базировалось на изучении поведения опытного человека тьютора, применяющего метод Сократа. Анализ полученного фактического материала позволили Коллинсу и Стивену выявить две глобальные цели, определяющие два типа диалоговых эпизодов в процессе обучения: *диагностика и коррекция* [70,71].

Диагностика определяется как процесс, позволяющий проследить связь между множеством «поверхностных ошибок» и глубоким непониманием материала. Возможность коррекции базируется на способности вначале классифицировать ошибки, а затем выбирать релевантный корректирующий метод. Коллинс и Стивен выявили пять типов «концептуальных ошибок», используемых тьютором человеком для корректировки обучающего метода: от простой ошибки в факторе, которая может быть скорректирована простым утверждением, до ошибочного

способа ведения дедуктивных умозаключений, который может быть скорректирован дополнительным обучением. Отмеченные свойства целенаправленного обучающего поведения человека тьютора были представлены Коллинсом и Стивенсом в виде теории, но не были воплощены в очередную версию WHY.

Расширение концепции сценарной модели предметной области базировалось на критике сценарно-ориентированного способа моделирования предметной области. По мнению авторов, сценарий в большей степени пригоден для управления диалоговым процессом и в меньшей степени для описания *ментальной модели* изучаемой предметной области. Ментальная модель определяется как внутренняя, доступная сознанию, модель реальной физической системы, используемая индивидом для построения логических умозаключений и предсказания поведения реальной физической системы. Ментальная модель одной и той же физической системы имеет вертикальную структуру. Стивен и Коллинс обнаружили, например, четыре типа ментальных моделей метеорологических процессов [72]:

- На планетарном уровне обычно используется *климатическая модель*, которая описывает глобальные водные и воздушные течения и их влияние на климат.
- С климатической моделью связана *модель циркуляции воды*, которая описывает абсорбцию воды воздушными массами, её транспортировку и возвращение на поверхность в виде осадков.
- Модель циркуляции воды связана с *макрофункциональной моделью парообразования*, с такими факторами как температура и влажность.
- Наконец существует *микромодель парообразования*, которая учитывает столкновение молекул воды и их испарение с поверхности.

Множественность ментальных моделей предметной области изменяет точку зрения на понимание сложной системы и обучения сложной системе. Понимание сложной системы в большой степени зависит от способности использовать различные ментальные модели и переходить от одной модели

к другой. Для каждой из моделей Стивен и Коллинз предложили шкалу, отражающую степень её усвоения и *определили обучение, как процесс последовательных приближений ментальной модели обучаемого к реальной физической системе*. Из такого понимания процесса обучения следует стратегия обучения, основанная на постоянном сопоставлении модельных предсказаний и фактов реальной физической системы, что является основной техникой метода Сократа. Несоответствие, обнаруженное в процессе диагностики, должно служить причиной корректировки ментальной модели ученика, а поскольку между ментальными моделями существует иерархическая упорядоченность, то обучение может рассматриваться также как *процесс диагностики и коррекции ментальных моделей ученика на всё более глубоких уровнях*.

4.1.2. Обучающие среды с интеллектуальной поддержкой

Система SOPHIE, описанная в работах Брауна и его коллег, открывает принципиально новый подход к разработке систем коммуникации знаний [73–77]. Разработчики системы SOPHIE не ставили перед собой цель разработать «искусственный тьютор», который «проводит» ученика через материал предметной области с учётом его когнитивной индивидуальности. Система SOPHIE представляет собой в большей степени «интеллектуальную среду» для обучения путём экспериментирования. Метод обучения SOPHIE относится к философии упрощенного конструктивизма и заключается в том, что обучение осуществляется в процессе исследования некоторой, специальным образом подготовленной обучающей среды. Ученик, экспериментируя со средой, проверяет свои идеи и гипотезы и получает в ответ критические замечания и советы. Предметной областью SOPHIE является поиск дефектов в электронных схемах.

В работах Брауна и его коллег описано три версии SOPHIE. Отличительной особенностью всех трёх версий является совершенный естественно-языковый интерфейс, базирующийся

на концепции семантического разбора предложений, предложенного Бартоном [78]. Фрагмент протокола диалога системы SOPHIE иллюстрирует возможности интерпретировать «некорректные» естественно-языковые вопросы (рис. 4.5).

Student: Now what is the outpt?
 SOPHIE: (OUTPT в OUTPUT) THE OUTPUT VOLTAGE IS 1.2 VOLTS. IF YOU ARE NOT SURE WHETHER A VALUE IS CORRECT, YOU CAN ASK "IS THAT VALUE CORRECT?"
 (После того, как ошибка в слове «outpt» распознана и разрешена как «output voltage» SOPHIE информирует ученика о своих возможностях)

Student: Is that right?
 SOPHIE: IN A WORKING CIRCUIT THE OUTPUT VOLTAGE IS 30.1 VOLTS. THE VALUE IS INCORRECT.
 (Заключение об анафоре: слово «that» относится к величине предыдущего измерения)

Student: What is the voltage between the current limiting transistor and the constant current source?
 SOPHIE: THE VOLTAGE AT №4 IS 2.48 VOLTS.
 (Слово «between» интерпретируется как ссылка на определённую точку, в которой измеряется напряжение по отношению к земле)

Student: What is it between there and the base of Q6?
 SOPHIE: THE VOLTAGE BETWEEN №4 AND №5 IS 1.28 VOLTS.
 (Слово «it» относится к напряжению, а слово «there» — к №4. Слово «between» в этом случае интерпретируется не так как в предыдущем вопросе ученика)

. . .

Рис. 4.5. Фрагмент протокола диалога ученика и системы SOPHIE

Версия SOPHIE I, следуя философии упрощенного конструкторизма, отличается тем, что не проявляет никакой педагогической инициативы и поэтому не содержит ни модель ученика, ни какую-либо стратегию обучения. Единственное, что делает система в смысле метода обучения, это фиксирует дефект и предлагает студенту его обнаружить. Следствием такого подхода является принятое распределение ролей партнеров диалога. Как видно из фрагмента протокола диалога, приведенного на рис. 4.5, распределение ролей зафиксировано и представляет собой полную противоположность WHY: ученик задаёт вопросы, а система отвечает на них. Таким образом, цель встроенного в обучающую среду тьютора — *формировать осмысленные ответы на текущие вопросы ученика, рассматривая их как независимые*.

Версия SOPHIE I отличается от ранее рассмотренных систем тем, что использует не одну, а несколько моделей предметной области: (1) *симулятор* в виде математической модели электронной схемы; (2) *процедурную модель* в виде набора «интеллектуальных специалистов», использующих модель в ходе логического вывода; и (3) *декларативную модель* в виде семантической сети фактов. Процедурная модель предназначена для ответов на вопросы ученика проведением соответствующих экспериментов с симулятором. «Специалисты» процедурной модели решают четыре типа задач логического вывода.

Задача первого типа заключается в формировании ответов на гипотетические вопросы ученика. Такие, например, как: «Что произойдет с таким-то компонентом, если будет произведено такое-то действие?»

Задача второго типа заключается в оценке гипотезы, которая предлагается учеником.

Задача третьего типа заключается в перечислении всех возможных гипотез, соответствующих проведенным измерениям, по соответствующему запросу ученика.

Четвёртая задача (считающаяся наиболее сложной) заключается в оценке новых измерений, предлагаемых учеником.

На рис. 4.6 приведен фрагмент протокола диалога SOPHIE I, иллюстрирующий способность системы решать перечисленные выше четыре типа задач логического вывода.

Student: What could be wrong?
SOPHIE: HERE ARE SOME FAULTS WHICH EXPLAIN THE MEASUREMENTS YOU OBSERVED: THE BETA OF Q3 IS LOW THE BETA OF Q4 IS LOW OR R22 HAS THE WRONG RESISTANCE

Student: What would the output voltage be if beta of Q3 were 8?
SOPHIE: IF THE BETA OF Q3 WERE 8 UNDER THE PRESENT SETTING, THE OUTPUT VOLTAGE WOULD BE 10.5 VOLTS. IN A WORKING CIRCUIT IT IS 19.9 VOLTS.

Student: Is the emitter current of Q4 correct?
SOPHIE: THE EMITTER CURRENT OF Q4 IS -61 MILLIAMPS. IN A WORKING CIRCUIT THE EMITTER OF Q4 IS -49 MILLIAMPS. THAT VALUE SEEMS TO BE INCORRECT.

Student: Could R22 be low?
SOPHIE: IS IT OK IF I USE 25 OHMS FOR THE RESISTANCE OF R22?

Student: Yes
SOPHIE: THAT FAULT AGREES WITH ALL OF MEASUREMENTS YOU HAVE MADE.

. . . .

Рис. 4.6. Фрагмент протокола диалога обучаемого и системы SOPHIE I, иллюстрирующий способность системы решать задачи логического вывода

Версия SOPHIE II отличается от версии SOPHIE I тем, что кроме «интеллектуальной среды» для обучения путём экспери-

ментирования *содержит также «эксперта по поиску неисправностей»*, который может продемонстрировать ученику свою (предположительно эталонную) стратегию рассуждений. Прежде чем привести цепь рассуждений, эксперт может задать ученику уточняющие вопросы. Таким образом, диалог SOPHIE II похож на диалог SCHOLAR в том смысле, что роли партнеров не зафиксированы, следовательно, в процессе обучения вопросы может задавать как система, так и ученик. Эксперт демонстрирует ученику свой вариант логических умозаключений вне зависимости от того, как реально развивается процесс поиска неисправности. Следовательно, эксперт, встроенный в SOPHIE II, не может «отслеживать логику ученика» в индивидуальном процессе поиска неисправностей. Задача преодоления этого недостатка была поставлена при разработке версии SOPHIE III.

4.1.3. Интеллектуальные репетиторы

Серия программ под общим наименованием WUSOR — представитель отдельного класса систем коммуникации знаний, получившим наименование *интеллектуальные репетиторы*.

Главная цель интеллектуального репетитора заключается в поддержке деятельности ученика при выполнении им некоторого задания путём трансформации ошибок, совершаемых учеником в процедурный опыт. Репетитор не вмешивается в процесс выполнения задания до тех пор, пока не обнаруживает ошибочные действия. Таким образом, репетитор осуществляет непрерывное и незаметное отслеживание активности ученика и прерывает его работу только тогда, когда действия ученика не совпадают с действиями, ожидаемыми репетитором. Искусство репетиторства заключается в нахождении оптимального момента прерывания активности ученика и в корректности формулировки подсказки. Опубликован ряд педагогических принципов компьютерного репетиторства, предложенных Бартоном и Брауном [79].

Программы WUSOR являются репетиторами компьютерной игры WUMPUS, известной тем, что она часто используется как предметная область при исследовании и изучении процесса логических умозаключений в прикладных системах искусственного интеллекта [21,80]. Эта компьютерная игра предполагает перемещение игрока через последовательность пещер в лабиринте. Игрока поджидают различные опасности: чудовище WUMPUS, бездонные ямы и летучие мыши, которые подхватывают игрока и роняют его в пещеру случайным образом. Как только игрок попадает в новую пещеру, он получает список близлежащих пещер и набор признаков, позволяющих сделать вывод о возможной опасности: сквозняк или писк, которые предупреждают о яме или летучей мыши в одной из смежных пещер; плохой запах, говорящий о том, что на расстоянии не более чем в две пещеры находится WUMPUS. Игрок перемещается, выбирая одну из смежных пещер. Для победы в игре необходимо выстрелить одной из пяти стрел в логово WUMPUSa. Игра проиграна, если игрок проваливается в яму или попадает в логово с WUMPUSом, или безрезультатно израсходует все стрелы. Решение о том, какая из окружающих пещер является безопасной и, следовательно, может быть выбрана следующей, требует от игрока проведения *логических умозаключений*.

Первая версия репетитора WUSOR-I включает два модуля: «эксперт» и «советчик» [81]. Модуль «эксперт» содержит описание предметной области в виде набора правил продукции, связывающих признаки и опасности и позволяющих получить прогноз об опасности или безопасности смежных пещер. Модуль «советчик» не хранит метод обучения. Он вмешивается в игру всякий раз, когда ученик делает неоптимальный выбор следующей пещеры, и предоставляет ему набор соответствующих объяснений. WUSOR-I не диагностирует состояние знаний ученика и не адаптирует своё вмешательство соответственно.

Главное отличие WUSOR-II от WUSOR-I заключается в наличии модели ученика, базирующейся на теории оверлея (overlay theory) Кара и Гольдштейна [82]. Оверлей Кара и Гольд-

штейна является классической парадигмой и часто используется при моделировании ученика в интеллектуальных тьюторских системах. Теория оверлея предполагает оценку знаний ученика сравнением его поведения с поведением эксперта. При этом принимается, что знания ученика являются частью знаний эксперта. Существует ряд трудностей при применении оверлейной модели. Например, она неприменима, если предметная область допускает множество парадигм решения задачи и, следовательно, может быть представлена множеством экспертов. На рис. 4.7 приведен фрагмент протокола диалога ученика с WUSOR-II.

WUSOR: HI MARY. THERE ARE 15 CAVES, 2 BATS, 2 PITS, AND 1 WUMPUS. YOU ARE NOW AT CAVE 15 WITH NEIGHBORS 4, 14, AND 0. BRRR! THERE IS A DRAFT. YOU ARE NEAR A PIT. WHAT A STENCH! THE WAMPUS IS NEAR! WHAT NOW?

Student: 4

WUSOR: YOU ARE NOW AT CAVE 4 WITH NEIGHBORS 15, 14, AND 2. BRRR! THERE IS A DRAFT. SQUEAK! A BAT IS NEAR. WHAT NOW?

Student: 14

WUSOR: MARY. IT ISN'T NECESSARY TO TAKE SUCH LARGE RISKS WITH PITS. ONE OF CAVES 2 AND 14 CONTAINS A PIT. LIKewise, ONE OF CAVES 0 AND 14 CONTAINS A PIT. THIS IS MULTIPLE EVIDENCE OF A PIT IN CAVE 14 WHICH MAKES IT LIKELY THAT CAVE 14 CONTAINS A PIT. IT IS LESS LIKELY THAT CAVE 0 CONTAINS A PIT. HENCE, MARY, WE MIGHT WANT TO EXPLORE CAVE 0 INSTEAD. DO YOU WANT TO TAKE BACK YOUR MOVE?

. . . .

Рис. 4.7. Фрагмент протокола диалога ученика и системы WUSOR-II

Программа WUSOR-III отличается от WUSOR-II ещё большим акцентом на представление знаний, ориентированное на конкретного ученика. При проектировании WUSOR-III была предложена концепция генетического графа [83], который является графическим представлением эволюционных отношений между фрагментами знаний и базируется на эволюционной теории интеллекта Пиаже [22]. Узлы генетического графа соответствуют элементарным навыкам, представленным индивидуальным правилом, а ветви — их эволюционным отношениям, таким как обобщение или аналогия. В WUSOR-III модель пользователя представлена оверлеем узлов генетического графа, включающих правильные и неправильные правила, а индивидуальная история пользователя представлена оверлеем связей генетического графа.

С точки зрения проблематики диалоговых методов решения проблем генетический граф является одним из немногих подходов, который может рассматриваться как унифицированная сетевая (сценарная) модель, включающая метод обучения и стратегию диагностики и предоставляющая средства персонализированного подхода к обучению.

Серия программ под общим наименованием «когнитивный тьютор», разработанных в университете Карнеги Мелон (США), является одним из наиболее известных примеров интеллектуальных репетиторов. Главный идеолог когнитивных тьюторов — Джон Андерсон, разработавший несколько версий унифицированной теории познания, известной под наименованием АСТ (Adaptive Control of Thought) [31,84,85]. Когнитивные тьюторы первоначально разрабатывались с целью проверки ключевых идей АСТ, главной из которых является убеждение в том, что процедурные знания человек хранит в виде системы продукционных правил, упорядоченных иерархией целей.

Когнитивные тьюторы разработаны для обучения в двух предметных областях: программирование на алгоритмических языках (LISP, Pascal) и элементарная математика (геометрия, алгебра) [65, 86–90].

Метод репетиторства, используемый в когнитивных тьюторах, называется «трассировка модели». Когнитивный тьютор хранит эталон процедурных знаний, которые должен освоить ученик, в виде последовательности шагов-действий. Если фактическая последовательность шагов-действий ученика совпадает с эталонной, то когнитивный тьютор никак не вмешивается в действия ученика. Если действия ученика отклоняются от эталонных, то тьютор реагирует выводом сообщения, содержащего расшифровку ошибки и рекомендации о дальнейших действиях. Как при описании эталона, так и при описании возможных отклонений используется система продукционных правил, упорядоченных иерархией целей. На рис. 4.8 приведен пример продукционного правила, используемого в LISP тьюторе.

```
ЕСЛИ цель-кодирование функции с табличным входным параметром
    и функция должна иметь доступ к каждому элементу
    таблицы, и структура таблицы произвольно сложная
ТО кодируйте cond для обеспечения car-cdr рекурсии и
    устанавливайте подцелью кодирование случаев останова и
    кодируйте случаи рекурсии
```

Рис. 4.8. Пример продукционного правила, используемого в LISP тьюторе

Одной из особенностей когнитивных тьюторов Андерсона, как систем коммуникации знаний, является отсутствие явного вербального диалога с учеником. Интерфейс программы разделяет экран на две части, одна из которых — поле действий ученика, а другая — поле для вывода сообщений тьютора. В последующих работах идея когнитивного тьютора развивалась, главным образом, в направлении использования новой версии унифицированной теории познания ACT-R для моделирования более тонких, направленных на передачу процедурных знаний элементов поведения ученика при решении задачи [91, 92].

4.1.4. Обобщённая структура системы коммуникации знаний

В завершении опишем обобщённую структуру рассмотренных систем коммуникации знаний. Все рассмотренные системы коммуникации знаний могут быть представлены четырьмя базовыми компонентами:

- модель предметной области;
- модель ученика;
- метод обучения;
- интерфейс системы с учеником.

Модель предметной области (domain knowledge, domain expertise) выполняет двойную функцию. С одной стороны, она представляет собой источник предметных знаний, «порционно» представляемых ученику, а с другой — желаемый стандарт знаний ученика (или цель обучения), который может использоваться для оценки его текущих знаний.

Модель ученика (student model) предназначена в основном для того, чтобы предсказывать диалоговое поведение ученика с целью его интерпретации. Идеальная модель ученика должна делать точные предсказания поведения любого ученика в любом контексте предметной области. Например, в виде множества возможных реакций/ответов на текущую обучающую сцену/вопрос. С моделью ученика часто связывается понятие *диагностического процесса (diagnostic process)*, под которым обычно понимается процесс адаптации модели ученика на основании данных, полученных в процессе диалога.

Метод обучения (pedagogical expertise) определяет какая обучающая сцена и на каком шаге диалога должна быть передана ученику. Метод обучения не кодируется жестко, а выводится из текущей ситуации в соответствии с набором принципов и/или продукционных правил. Дидактические решения принимаются на основе предсказания реакции ученика, генерируемой моделью ученика, его реальной реакцией и отношений между обучающими сценами, хранимыми в модели предметной области. Методы обучения варьируются по отношению к степени «свободы», предоставляемой ученику, и могут быть упо-

рядоченно расположены между двумя полюсами: от полного контроля всех действий ученика и выполнения «лоцманских» функций при проведении ученика через обучающие сцены до полной свободы навигационной активности ученика и выдачи рекомендаций по его запросу.

Интерфейс системы коммуникации знаний с учеником в приведенных выше системах, как правило, варьировался вокруг вариантов ограниченного естественного языка. Таким образом, авторы отмеченных систем рассматривали диалог как сугубо вербальный процесс. Графические элементы предположительно включались в интерфейс, как правило, только для презентации обучающих сцен в интеллектуальных обучающих средах.

Рис. 4.9 иллюстрирует отношения между отмеченными четырьмя базовыми компонентами систем коммуникации знаний.

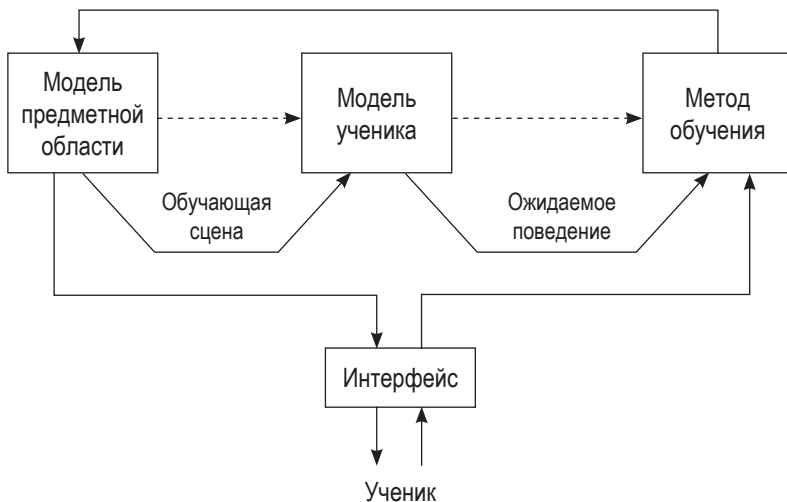


Рис. 4.9. Основные компоненты системы коммуникации знаний

В третьем разделе (на рис. 3.27) приведена обобщённая структурная модель искусственного активного агента вопро-

сно-ответного диалога. Модель проблемно-независима и применима для реализации любого диалогового метода решения задачи, в том числе и для реализации любого диалогового метода обучения, используемого в системах коммуникации знаний. Поэтому следует ожидать, что эта модель является более «широкой», чем структурная систем коммуникации знаний, приведенная на рис. 4.9. Рассмотрим, каким образом соотносятся функции компонентов искусственного активного агента вопросно-ответного диалога и функции компонентов системы коммуникации знаний.

Функции системы коммуникации знаний сосредоточены в её базовых компонентах: (1) метод обучения; (2) модель ученика; (3) модель предметной области; (4) интерфейс с учеником. В этом смысле система коммуникации знаний является *системой с сосредоточенными функциями*. В искусственном активном агенте вопросно-ответного диалога функции системы распределены, главным образом, между узловыми элементами сценария диалога. В этом смысле искусственный активный агент вопросно-ответного диалога является *системой с распределенными функциями*.

Метод обучения. В системе коммуникации знаний метод обучения представляет собой компонент, главной функциональной обязанностью которого является определение того, какая обучающая сцена и на каком шаге диалога должна быть передана ученику. Метод обучения может варьироваться во время одной диалоговой сессии. В искусственном активном агенте вопросно-ответного диалога метод обучения реализуется сетевой структурой, названной «диалоговый метод доступа» *DiAM*. Понятно, что топология сети диалогового метода доступа может быть различной на разных этапах диалоговой сессии и отражает вариации метода обучения.

Модель ученика. В системе коммуникации знаний функции модели ученика заключаются в предсказании диалогового поведения ученика. По сути, такое предсказание должно отвечать на вопрос: «Какую реакцию следует ожидать от ученика на данную обучающую сцену?» В искусственном активном агенте вопросно-ответ-

ного диалога функции модели ученика реализуются в структуре шагов диалога, из которых строится сеть *DiAM*. При проектировании шага учитывается только *множество ожидаемых реакций ученика*. Эти реакции распознаются диалоговым агентом. Все остальные реакции рассматриваются как нераспознаваемые.

Модель предметной области. В системе коммуникации знаний функции модели предметной области заключаются, главным образом, в детерминировании знаний о предметной области, которые «порционно» предоставляются ученику в виде обучающих сцен. В искусственном активном агенте вопросно-ответного диалога эти функции реализуются памятью вопросов *QueMet*, хранящей описание интеррогативных стимулов. Даталогическое представление интеррогативных стимулов (смотри подраздел 3.4.2) позволяет конструировать обучающие сцены для любой предметной области.

Интерфейс с учеником. В системе коммуникации знаний интерфейс реализует функции вербальной коммуникации и его структура не меняется на протяжении всей диалоговой сессии. В искусственном активном агенте вопросно-ответного диалога функции интерфейса реализуются сценой, которая включает интеррогативный стимул и ожидаемые реакции реактивного агента. Поэтому интерфейс искусственного диалогового агента является *динамически изменяющимся* и *индивидуальным* для каждого шага диалога. В этом смысле интерфейс искусственного активного агента вопросно-ответного диалога *адаптивен*.

В подразделе 4.2 рассмотрены основные положения теории обучения и классические методы лингводидактики, которые могут быть использованы при конструировании сценариев для искусственных активных агентов вопросно-ответного диалога, ориентированных на обучение иностранным языкам.

4.2. Теория обучения

Несмотря на то, что метод обучения является одним из четырёх базовых концептов систем коммуникации знаний,

авторы описанных в литературе систем коммуникации знаний, как правило, не опираются на результаты, полученные в исследованиях в области *теории обучающих процессов*. Теория обучения является методологической основой проектирования методов обучения и должна обязательно учитываться при их проектировании. В настоящее время теория обучения не существует в виде завершенной формальной системы, использующей единый концептуальный базис, а представлена в виде следующих относительно обособленных разделов:

- философия обучения;
- сущность обучения;
- отношение между учеником и обучающим материалом;
- обучение и деятельность.

4.2.1. Философия обучения

Философия обучения представлена следующими основными направлениями [93]:

- структурализм;
- конструктивизм;
- деятельностно-ориентированное обучение;
- обучение с множественными связями.

В самом общем случае *структурализм* (structuralism) может быть определён при помощи своей центральной идеи, заключающейся в том, что знания существуют в виде множества «кусков» или фрагментов, а не в виде гомогенной структуры. Другой отличительной чертой структурализма является противопоставление структуры знаний их использованию или представлению. Структура в этом противопоставлении является более важной, чем способ её представления. Ясно, что, если автор метода обучения стоит на позициях структурализма, то это отражается на предлагаемом им методе. Например, если полагать, что знания существуют в виде фрагментов, то следствием этого является

концепция «обучающего эпизода». Более того, от концепции фрагментации знаний можно легко перейти к концепции фрагментации самого процесса обучения [94]. В то время как не структуралист ожидает, что все части нового материала изучать одинаково легко (или одинаково трудно), структуралист предполагает, что нужны различные ментальные усилия, зависящие от того, необходимо ли ученику создать новую ментальную модель, модифицировать существующую модель или дополнить её.

Термин *конструктивизм* (constructivism) в последнее время активно используется при обсуждении проблем образования вообще и проектирования обучающих методов, в частности. Конструктивизм обозначает позицию, заключающуюся в том, что знания человека есть результат взаимодействия двух факторов: внешнего (обучающего) стимула и предыдущего ментального опыта. Предположение, что обучение обусловлено предыдущим ментальным опытом, является одним из ключевых в теории обучения и изучения. Конструктивизм можно рассматривать как теорию, заполняющую пространство между двумя экстремумами: случаями, когда определяющими факторами обучения являются внешние стимулы и хорошо работает механизм трансмиссии знаний, и случаями, когда доминирует предыдущий ментальный опыт и модель трансмиссии не работает. *Философия конструктивизма утверждает, что обучение — не единственная причина изучения, а только одна из причин, действующей в виде внешних, обучающих стимулов. Упрощенный конструктивизм отрицает необходимость и возможность обучения путём трансмиссии знаний и предполагает, что единственным методом обучения является экспериментирование с предметом изучения на основе накопленного ментального опыта.*

Обучение с ориентацией на деятельность (activity-orientedness learning), направленную на изучаемый предмет, исследует проблему, которую можно сформулировать в виде следующего вопроса: «Должен ли ученик активно воздействовать на изучаемый материал или достаточно пассивного вос-

приятия?» Некоторые конструктивисты считают, что положительный ответ на этот вопрос следует непосредственно из идеи конструктивизма, утверждая, что успех изучения определяется степенью деятельностной активности ученика с изучаемым материалом [95]. Ясно, однако, что такая точка зрения соответствует упрощенному конструктивизму и является дискуссионной. Не отрицая роли деятельностной активности ученика над изучаемым материалом, отметим, что в ряде случаев трансмиссия знаний успешно осуществляется без отмеченной активности. Можно предположить, что существуют несколько причин, вызывающих необходимость деятельностной активности: (1) недостаточность пассивного восприятия для понимания и (2) необходимость завершить изучение в ограниченные временные рамки.

Фундаментальным свойством человеческой памяти есть забываемость. Наиболее загадочным фактом, относящимся к феномену забываемости, является отсутствие каких бы то ни было доказательств того, что забываемость биологически необходима и обеспечивает человеку эволюционные преимущества, например, освобождает ресурс памяти для размещения новых знаний. Но существует ряд фактов, подтверждающих, что человек хранит в своей памяти всю (или почти всю) информацию, запомненную ранее. Если какая-либо информация утрачена в результате забывания, то она может быть частично или полностью реконструирована. Бартлет показал, что реконструирование является одним из фундаментальных и всеобъемлющих свойств памяти человека и определяется мощностью *множественных связей*, ассоциированных с реконструируемым концептом [96]. Отмеченный факт порождает *принцип обучения с множественными связями*, заключающийся в том, что чем больше связей у изучаемого материала с ранее накопленными знаниями, тем больше шансов у этого материала быть изученным, запомненным и в последствии реконструируемым. Ясно, что если бы человеческая память не обладала свойством забываемости, множественные связи не были бы так важны.

4.2.2. Сущность обучения

Сущность обучения рассматривается в следующих контекстах:

- соотношение декларативных и процедурных компонентов в обучении;
- социально распределённая природа знаний;
- групповое (школьное) обучение.

Деятельностно-ориентированное обучение не означает, что предпочтение должно быть отдано изучению процедур, а не концептов. Как показано Пиаже, в его теории развития интеллекта концепты или декларативные знания являются результатом обобщения, которое получено в ходе деятельностной активности, а их изучение существенно увеличивает мощь процедурных знаний обучаемого [22]. Шахматный эксперт, например, демонстрирует высокий уровень процедурных знаний, выражающихся в последовательном перемещении фигур только потому, что опирается на опыт, хранимый в виде декларативных знаний. Главной сутью всех академических предметов во всех дисциплинах является описание, а не сенсомоторный личный опыт. *Академические знания, таким образом, в первую очередь являются декларативными знаниями.* В высшем образовании по крайней мере деятельностно-ориентированное обучение должно интерпретироваться, в первую очередь, как деятельность, ориентированная на описание объекта изучения, а не на физические действия с ним.

Знания человека и, следовательно, то чему он собирается учить других, являются социально распределёнными в том смысле, что существует единая семантическая интерпретация этих знаний для всех людей. Путнам, исследуя феномен социальной распределённости знаний, отмечает роль общепринятого языка общения и приводит в качестве примера такие слова, как «вода» и «золото», которые имеют и должны иметь одинаковый смысл для большинства людей [97].

В системе школьного обучения широко распространено представление о знаниях как о чём-то, что может быть протек-

стировано на экзамене, и вера в то, что методы, подготавливающие учащегося к успешной сдаче экзамена, являются универсальными механизмами обучения. Исследования, проведенные с целью изучения школьного обучения, выявляют его отличительные характеристики [98]. Главный вывод заключается в том, что школьное обучение деконтекстуализировано и характеризуется безконтекстностью, обобщенностью и абстрактностью. Школьное обучение формирует навыки, специфические для школьной культуры (например, вспоминание материала во время тестирования).

4.2.3. Отношение между учеником и обучающим материалом

Отношение между учеником и обучающим материалом рассмотрим в контексте двух проблем:

- мотивация к обучению;
- выбор очередной задачи учеником.

Множество психологических теорий и моделей, применимых в теории обучения (бихевиоризм, модели познания SOAR [99] и АСТ-R [31]) не учитывают мотивацию и, следовательно неявно, предполагают, что мотивация не влияет на обучение. Согласно этим теориям обучение осуществляется унифицировано, неизбежно и независимо от желания ученика изучать материал. Однако если мотивация мало влияет на обучении в детском возрасте, то для взрослого человека мотивация является главным фактором обучения. Взрослый человек мало интересуется новым материалом, если он не релевантен его интересам, однако демонстрирует высокую эффективность при изучении материала, в котором он заинтересован. *Мотивацию можно рассматривать как готовность ученика выполнять ментальное усилие.* Мотивация, таким образом, позволяет компенсировать ожидаемые неприятные ощущения, связанные с ментальным усилием, необходимым для изучения нового материала.

Философия конструктивизма для случая, когда доминирует ментальный опыт, предполагает, что ученик самостоятель-

но осуществляет навигацию в обучающей среде и, следовательно, должен решать проблему выбора очередной задачи. Важность рассматриваемой проблемы выбора для теории обучения вообще и для конструктивистского подхода, в частности, подтверждается тем, что в ряде случаев она выделяется в отдельный принцип обучения [93]. Если выбор очередной задачи осуществляется не случайным образом, а целенаправленно, то это означает, что *ментальный опыт ученика включает метод изучения в виде метода навигации по изучаемому материалу*.

4.2.4. Обучение и деятельность

Обучение и деятельность является отдельным разделом теории обучения и рассматривает следующие основные вопросы:

- роль обратной связи в процессе обучения;
- обучение с учётом необходимой степени деятельностной активности.

Несмотря на то, что успешное изучение нового материала не всегда зависит от обратной связи (например, презентация задачи не обязательно должна сопровождаться исследованием того, насколько успешной она была), обратная связь часто является ключевым моментом процесса обучения. Наличие управляющей обратной связи от задачи к ученику превращает простое решение задачи в задаче-ориентированное обучение. Как философия конструктивизма, так и деятельностно-ориентированное обучение предполагают использование обратной связи. Роль обратной связи в процессе изучения может быть ранжирована следующим образом:

- полное отсутствие обратной связи, исключающее обучение (например, невозможно бросить камень в цель, которую вы не видите);
- внутренняя обратная связь, позволяющая самому ученику судить об успешности или не успешности обучения (например, бросание камня в цель, которая видна);

- простая оценочная обратная связь от окружающих (например, реакция окружающих на рассказанный анекдот);
- диагностика, или дескриптивная обратная связь, которая сообщает не только об успешности или не успешности обучения, но и о том, что необходимо изменить в методе обучения и, возможно, почему это нужно сделать.

Чем больше знаний накоплено учеником, тем меньше он нуждается во внешней обратной связи, тем больше функций он возлагает на внутреннюю обратную связь и тем более автономным становится его поведение. Однако поскольку невозможно гарантировать достаточности внутренней обратной связи, *конструктивистская стратегия необходимо требует наличия диагностической обратной связи*, как необходимого компонента обучения.

Ясно, что выполнение действий над предметом изучения не всегда является необходимым условием обучения. Большинство представителей академического мира уверены, что они обучаются при помощи книг, а астрономы и историки, например, изучают такие предметы, с которыми они никак не могут выполнять какие-либо действия — удаленные звёзды и прошлые цивилизации. Но, существует множество аргументов в пользу точки зрения, согласно которой обучение невозможно без деятельности над предметом изучения. В литературе, посвященной теории обучения и изучения, можно встретить выражение: «я действую и я понимаю», отражающее наиболее радикальный взгляд на роль деятельности в изучении [100]. Рациональная точка зрения на необходимую степень деятельностной активности в процессе изучения может быть сформулирована следующим образом. Деятельность над предметом изучения важна во всех случаях, однако степень её влияния на эффективность и конечный результат изучения зависят от ряда следующих факторов.

- Обучение без деятельностной активности успешно в том случае, когда ученик располагает ментальной моделью, хорошо адаптированной к новой информации. Чем более новым является изучаемый материал, тем более важна обратная связь и тем важнее деятельностная активность.

- Необходимость получения информации по цепи внешней обратной связи (что необходимо влечет деятельностную активность) часто важно в случае, когда внутренняя обратная связь ещё не сформирована или когда часть решения полностью утеряна (забыта) и ученик не может продвинуться дальше без внешней помощи.
- Выполнение действий над предметом изучения может рассматриваться как элемент тестирования процесса обучения.

4.3. Классические методы лингводидактики

Компьютерная лингводидактика занимается моделированием и разработкой компьютерных программ обучения иностранным языкам. Компьютерная лингводидактика является идеальной предметной областью для развития идей диалоговых когнитивных тьюторов и их проверки при помощи проектирования реальных систем. Обучение иностранным языкам является относительно стабильной предметной областью и не подвержено резким методологическим изменениям. Здесь применимы практически все идеи как систем коммуникации знаний, так и теории обучения. Привлекательной является и попытка встроить известные, но «некомпьютерные» методы обучения иностранным языкам в архитектуру системы коммуникации знаний.

К классическим методам обучения иностранным языкам, рассмотренным в настоящем подразделе, относятся:

- ситуативный метод;
- аудио-лингвистический метод;
- метод общего физического отклика.

4.3.1. Ситуативный метод обучения

В теории обучения иностранным языкам *ситуативный метод обучения* (situational language teaching) часто классифи-

цируется, как Британский структурализм [101]. Британские теоретики предложили свою версию структурализма, опирающуюся на понятие «ситуация». Они подчеркивали близкую связь структуры языка, контекста и ситуаций, в которых применяется язык. Таким образом, предположение о том, что *знание языковых структур должно быть увязано с ситуациями, в которых они употребляются*, — одно из отличительных черт ситуативного метода обучения. Основным видом аудиторной работы при ситуативном методе является *устное представление моделей предложений* с последующей тренировкой представленных структур. Следовательно, приобретение навыков автоматического контроля базисных структур языка при ситуативном методе достигается, в основном, через устную речь. Холидэй и другие исследователи отмечали, что язык, как средство общения, вместе с участниками общения может рассматриваться как средство целенаправленной деятельности, а речевое высказывание индивидуума всегда осуществляется с какой-либо целью [102].

Метод ситуативного обучения предполагает индуктивный подход к изучению грамматики. Это означает, что значение слов и грамматических структур воспринимается не переводом на родной язык, а непосредственно из ситуации. Следовательно, соответствующим образом подобранная ситуация является обязательным окружением отмеченного индуктивного процесса.

Основу при обучении ситуативным методом составляют так называемые *ситуативные программы*. Ситуативная программа представляет собой список ситуаций, предназначенных для изучения базисных грамматических структур и моделей предложений, расположенных в порядке их изучения. Таким образом, ситуация используется для первоначальной презентации и последующей тренировки моделей предложений, которые извлекаются из ситуации [103]. При презентации языкового материала ситуацией необходимо управлять так, чтобы максимизировать степень понимания материала путём увеличения количества связей с контекстом. Приёмы аудиторной работы,

в основном, сводятся к целенаправленным повторениям, включая повторение хором, диктанты, тренировки и упражнения по целенаправленному чтению вслух и описанию, работа в парах и другие виды групповой работы.

Роль ученика на начальных стадиях изучения языка заключается в том, чтобы слушать и повторять за учителем обучающий материал и отвечать на вопросы и команды. Ученик ещё не обладает навыками управления содержанием материала и, следовательно, ему необходимо подчиняться командам учителя. Позже, при появлении отмеченных навыков соучастие ученика в процессе обучения в виде проявления инициативы в ответах, постановки вопросов, обмену вопросами следует поощрять. Однако последовательностью введения и тренировкой нового языкового материала управляет учитель.

Роль учителя может быть описана следующим образом [104]. На начальной стадии учитель служит моделью, целенаправленно генерирует ситуации и побуждает ученика давать правильные ответы. На последующих стадиях учащимся предоставляется возможность пользоваться языком в менее контролируемых ситуациях, а учитель следит за выявлением и устранением грамматических и структурных ошибок. Ситуативный метод абсолютно нетерпимо относится к разного рода ошибкам и предполагает их немедленное исправление.

Учебник должен содержать тщательно спланированные, организованные вокруг грамматических структур уроки. Наглядные пособия совместно с грамматическим материалом — важный компонент ситуативного метода. Однако учебник используется только как руководство к процессу изучения языка. Учитель всегда должен оставаться «хозяином положения».

Ниже описаны варианты обучающего поведения учителя, дающие представление о ситуативном методе обучения [103].

1. Аудио-тренировка. Учитель привлекает внимание учащихся и повторяет образец грамматической структуры или слова несколько раз. В первый раз — медленно и раздельно (Where ... is ... the ... pen?).

2. Проговаривание хором. Все учащиеся повторяют за учителем предлагаемые им образцы (лучше, когда учитель даёт чёткую инструкцию «повторите» и жестом руки задаёт такт и ударение).
3. Индивидуальная имитация. Учитель просит отдельных учащихся повторить модель, с целью контроля произношения.
4. Вычленение. Учитель выделяет звуки, слова или группы слов, вызвавшие затруднение и «отрабатывает» их несколько раз прежде, чем поместить в контекст.
5. Конструирование новой модели. Учитель предлагает учащимся задавать вопросы и отвечать на вопросы, пользуясь уже освоенными штампами, анализирует ответы с целью определения последующей модели.
6. Воспроизводство. Учитель, используя мимику, наводящие слова и жесты, побуждает учащихся задавать вопросы, делать утверждения и приводить примеры обсуждаемых моделей.
7. Тренировка на вопросах и ответах. Учитель побуждает одного из учащихся задавать вопрос, а другого — отвечать на него до тех пор, пока это упражнение не выполнит большинство учащихся.
8. Корректировка. Учитель отмечает факт допущенной ошибки и предлагает учащимся самостоятельно её отыскать.

4.3.2. Аудио-лингвистический метод обучения

Аудио-лингвистический метод (the audiolingual method) возник как оппозиция традиционному грамматическому подходу, который жестко привязывает изучение языка к грамматике.

Методологической основой аудио-лингвистического метода являются: структурная лингвистика, контрастный анализ (со-

поставление различий между родным и изучаемым языками), аудирование и бихевиористическая психология.

Важная характеристика структурной лингвистики — утверждение о том, что первичной языковой средой является устная речь. Во многих языках отсутствует письменность. Люди учатся говорить прежде, чем читать и писать. Брукс, например, утверждает: «... первично то, что проговаривается устно, а то, что представлено в письменной форме — вторично» [105]. Это утверждение противоречит точке зрения на взаимоотношение устной и письменной форм языка, согласно которой язык существует в виде символов, записанных на бумаге, а речь — это несовершенная реализация версии чистой письменной формы языка.

С бихевиористической точки зрения человеческая психика хранит и воспроизводит множество поведений. Проявление конкретного поведения определяется тремя факторами, важными, с точки зрения обучения: (1) *стимул*, «извлекающий» поведение; (2) *реакция*, являющаяся откликом на стимул; (3) *подкрепление*, указывающее, является ли реакция успешной (или неуспешной) и следует (или не следует) её запомнить для дальнейшего повторения [106,107]. Рис. 4.10 иллюстрирует отношение между отмеченными тремя факторами.

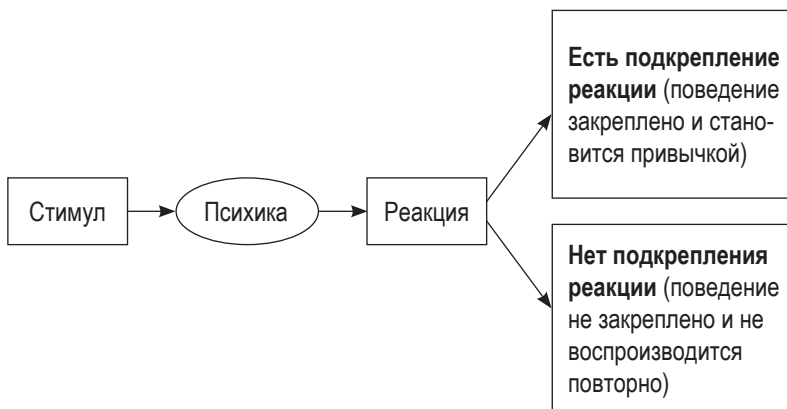


Рис. 4.10. Бихевиористическое понимание процесса обучения

Подкрепление (reinforcement) — важный элемент процесса обучения, так как увеличивает шанс повторного использования «удачного» поведения и в итоге превращает его в автоматический навык.

Если применить бихевиористическую модель к обучению иностранному языку, то необходимо рассматривать: психику, как субъект обучения (ученик); поведение, как вербальное поведение; стимул, как обучающий эпизод; реакцию, как реакцию ученика на обучающий эпизод; подкрепление, как внешнее одобрение, или внутреннее удовлетворение от применения изучаемого языка. Овладение иностранным языком, таким образом, может рассматриваться как формирование у ученика множества языковых цепочек, связывающих обучающие эпизоды с вербальными поведенческими реакциями. В основе аудио-лингвистического метода лежат следующие основные постулаты [108].

- Процесс изучения иностранного языка представляет собой процесс механического формирования навыков. Навыки формируются при правильных, а не ошибочных реакциях-ответах. Запоминание эталонов сводит к минимуму возникновение ошибок.
- Языковые навыки формируются более эффективно, если материал вначале представлен в устной форме, а затем — в письменной.
- Аналогия создает лучшую основу для изучения языка, чем анализ. Поэтому объяснять правила следует не раньше, чем учащиеся изучат образцы в разнообразных контекстах и научатся воспринимать материал по аналогии. Следовательно, подход к обучению грамматики является индуктивным, а не дедуктивным.
- Изучение семантики слов должно осуществляться не в изоляции, а в лингвистическом и культурном контексте. Следовательно, в обучение языку непременно вовлекается изучение культурной системы тех людей, которые говорят на этом языке.

Стратегии обучения, использующие аудио-лингвистический метод, часто рассматриваются в контексте краткосрочных и дол-

госрочных целей. Краткосрочные цели включают: восприятие текста на слух и его понимание; отработку произношения; распознавание речевых символов как графических знаков печатного текста и способность воспроизвести эти символы в письменной форме. Долгосрочные цели предусматривают знание второго языка в той мере, в какой им владеет двуязычный индивидуум.

Роль ученика определена моделью, приведенной на рис. 4.10. Ученики, ведомые последовательностью обучающих стимулов, генерируемых учителем, не могут явно управлять содержанием, темпом и стилем процесса обучения. Их не поощряют к инициативе, так как это может привести к ошибкам. Тот факт, что на начальной стадии ученики могут не понимать смысла и значения слов или фраз, которые они повторяют, не считается недостатком, поскольку считается, что слушая, имитируя преподавателя и выполняя его задания они изучают новую форму вербального поведения.

Роль учителя определяется как центральная. Учитель, являясь активным партнером общения, моделирует изучаемый язык, контролирует направление и темп обучения, управляет и направляет деятельность учеников, варьирует задания и т.п. Предполагается, что изучение языка — результат обучающего вербального взаимодействия учителя и ученика.

Учебник должен помогать учителю и на начальной стадии изучения языка не предоставляется ученикам. Считается, что на начальной стадии обучения предъявление печатного материала нежелательно, поскольку это отвлекает учащихся от устного восприятия. Магнитофонные записи и другие аудиовизуальные средства играют ключевую роль в аудио-лингвистическом методе. Если учитель не является носителем изучаемого языка, то магнитофон воспроизводит точные модели диалогов и тренировочных упражнений. Поэтому весьма существенный фактор, определяющий успех аудио-лингвистического метода, — это наличие лингафонной лаборатории.

Ниже изложены варианты типичного обучающего поведения учителя, дающие представление об аудио-лингвистическом методе.

1. Учащиеся прослушивают модель диалога, содержащего ключевые объекты всей аудиторной работы во время урока. Диалог повторяется построчно индивидуально или хором. Ошибки в произношении и грамматике исправляются учителем незамедлительно. Запоминание диалога происходит постепенно, строчка за строчкой. В случае необходимости строчка может быть разбита на несколько фраз. Диалог читают вслух хором, причем одна половина группы читает от лица говорящего, а другая — от лица оппонента. На этой фазе учащиеся не сверяются с учебником.
2. Диалог адаптируется к интересам учащихся самими учащимися путем замены слов или фраз.
3. Из диалога выбираются определенные ключевые структуры, которые служат основой для последующей тренировки, осуществляемой хором или индивидуально. На этом этапе допускается минимальные грамматические объяснения.
4. Учащимся разрешается пользоваться учебником для сопровождения чтения и работы со списком слов, на основе которых строится диалог. На начальном этапе, письмо является имитационным и сводится к переписыванию тренируемых предложений. По мере наращивания умений учащиеся могут выписывать варианты структурных единиц или писать короткие сочинения на заданные темы.
5. В лингафонной лаборатории выполняются упражнения на повторение за диктором и осуществляется дальнейшая работа с диалогом.

4.3.3. Метод общего физического отклика

Метод общего физического отклика (total physical response) — это метод преподавания иностранного языка, опи-

рающийся на координацию речи и моторную активность ученика [109].

Метод общего физического отклика связан с психологической теорией «трассировки», в которой утверждается, что чем чаще и интенсивнее трассируются связи в памяти, тем прочнее будут ассоциации и тем легче их вспомнить [110]. Трассировку связей можно осуществлять как вербально, так и в сочетании с моторной активностью, которая увеличивает вероятность успешного воспоминания. Так, речь, обращенная к детям, состоит, в основном, из команд, на которые дети реагируют физически раньше, чем научатся давать словесные ответы. Сторонники метода утверждают, что глагол в повелительном наклонении является главным лингвистическим средством, вокруг которого организуется изучение языка. По их мнению ученик может воспринять грамматическую структуру языка, не прибегая к абстракциям. Метод предполагает наличие игровой атмосферы, при которой ученики освобождаются от стресса и которая способствует возникновению положительных эмоций.

Основные методологические принципы, лежащие в основе метода общего физического отклика, могут быть сформулированы следующим образом [111]:

- Понимание предшествует языковым навыкам.
- Обучение говорению должно быть задержано до тех пор, пока не выработано умение понимать.
- Навыки, полученные при слушании, трансформируются в другие навыки.
- Обучение должно делать акцент на значение, а не на форму.
- Обучение должно минимизировать стрессовое состояние ученика.

Роль ученика в основном заключается во внимательном слушании команд, предлагаемых учителем с последующей физической реакцией на эти команды. Ученик ограниченно влияет на содержание обучаемого материала. От него ожидается умение распознавать и физически реагировать на новые обучающие стимулы, являющиеся комбинацией ранее изученных

обучающих стимулов. Ученик сам оценивает свой прогресс в обучении и начинает говорить, когда чувствует потребность и готовность к говорению.

Роль учителя. Учитель играет активную и непосредственную роль. Учитель должен располагать детально разработанным сценарием урока. Ашер отмечает: «Полезно записать точно фразы, которые вы собираетесь использовать, особенно новые команды, потому что ответная реакция может быть столь быстрой, что не останется времени создавать что-либо спонтанно» [112]. Учитель инициирует взаимодействия учеников друг с другом.

Учебник. В учебниках, ориентированных на метод общего физического отклика, как правило, отсутствуют базисные тексты. На начальных стадиях обучения достаточно голоса, действий и жестов учителя, чтобы вызвать ответные реакции учеников. Далее, учитель использует наличествующие в аудитории предметы: книгу, ручку, мебель и т.д. Затем используются подготовленные наглядные пособия: картинки, слайды, фильмы, таблицы и т.д. Специальные наборы предметов могут способствовать обыгрыванию тематических ситуаций и конструированию предложений.

Ниже описаны элементы типичного обучающего поведения учителя, дающие представление о методе общего физического отклика [112].

1. Игровые действия для закрепления ранее изученного материала. Каждому из группы учащихся даётся индивидуальное задание для совершения какого-либо действия.
2. Команды на прежде введенные повелительные глаголы с вариацией лексических единиц. Например: вымой (руки, лицо, чашку); найди (полотенце, мыло) и т.д.
3. Команды, которые используют одно и то же слово в нескольких предложениях. Например, для слова «треугольник»: (а) начерти треугольник на доске; (б) выбери из фигур на столе треугольник и дай его мне; (в) положи треугольник рядом с квадратом.

4. Простые вопросы, на которые учащийся отвечает жестом. Например, «Где книга?» или «Где стол учителя?»
5. Смена ролей учащегося и учителя. Один из учащихся отдаёт команды, управляющие поведением учителя или других учащихся.
6. Чтение и писание. Учитель пишет на классной доске новое слово и предложение, которое его иллюстрирует. Затем произносит слово и производит действие, соответствующее содержанию предложения. Учащиеся, выслушав учителя, записывают информацию в тетради.

4.4. Сцены лингводидактических тьюторов

Как отмечено в начальной части раздела, идеальная тьюторская система в процессе диалогового взаимодействия с учеником адаптируется к текущему уровню знаний ученика и степени понимания учебного материала, а также осуществляет процесс обучения методом, соответствующим текущим когнитивным способностям и стилю ученика. Из этого следует, что диалоговая сессия когнитивного тьютора и ученика должна включать два типа эпизодов: (1) эпизоды, осуществляющие тестирование ученика с целью диагностики текущего уровня знаний и определения его когнитивных способностей и когнитивного стиля; (2) эпизоды, осуществляющие непосредственно обучение. Сцены, используемые для проектирования обучающих эпизодов будем называть *обучающими сценами*, а сцены, используемые для проектирования эпизодов, осуществляющих тестирование — *диагностическими сценами*.

Прежде чем рассматривать диагностические и обучающие сцены, используемые в лингводидактике, разработаем универсальную классификацию структур лингводидактических сцен, которая инвариантна к вопросу, является ли сцена обучающей или диагностической. Поскольку лингводидактическая сцена представляет собой ассоциацию лингводидактического сти-

мула и ожидаемых реакций, классификация лингводидактических сцен может быть построена путём ассоциации типов лингводидактических стимулов и ожидаемых реакций. Понятно, что для реализации такого подхода к классификации лингводидактических сцен необходимо вначале сформировать множество типов лингводидактических стимулов и множество типов ожидаемых реакций. При формировании лингводидактических стимулов учтём описание обучающих стимулов, приведенное в подразделе 4.3, а также описание диагностических стимулов, используемых в психодиагностике.

Психодиагностика характеризуется широким спектром используемых методик [113–116]. Многообразие таких методик порождает множество систем классификации как методик, так и используемых в них диагностических стимулов в зависимости от того, какие классификационные признаки рассматриваются как значимые. Для наших целей значимым классификационным признаком может служить *применимость психодиагностической методики для проектирования сценария диалога и возможность реализации диагностического стимула средствами даталогической сцены* (смотри подраздел 3.4.2).

Лингводидактические стимулы классифицируются по способу презентации интеррогатива на следующие классы:

- вербальные (словесные) стимулы;
- невербальные стимулы.

В общем случае вербальные и невербальные стимулы могут воздействовать как на зрительный, так и на слуховой анализатор ученика, однако чаще, (особенно в психодиагностике) вербальные стимулы представлены текстом и воспринимаются органом зрения.

Вербальные стимулы классифицируются по возможности их адаптации к индивидуальности ученика на следующие классы:

- стандартизованные стимулы;
- индивидуально-ориентированные стимулы.

Стандартизованные стимулы одинаковы для всех учеников, в то время как индивидуально-ориентированные сти-

мулы подбираются индивидуально для каждого конкретного ученика.

Невербальные стимулы классифицируются по возможности изменять свои параметры во время их восприятия учеником на следующие классы:

- статические стимулы;
- динамические стимулы.

Параметры статических невербальных стимулов (такие, например, как форма, цвет, высота тона звука и т. д.) неизменны во время их восприятия учеником, а параметры динамических невербальных стимулов могут изменяться.

На рис. 4.11 приведена диаграмма, иллюстрирующая множество типов лингводидактических стимулов.

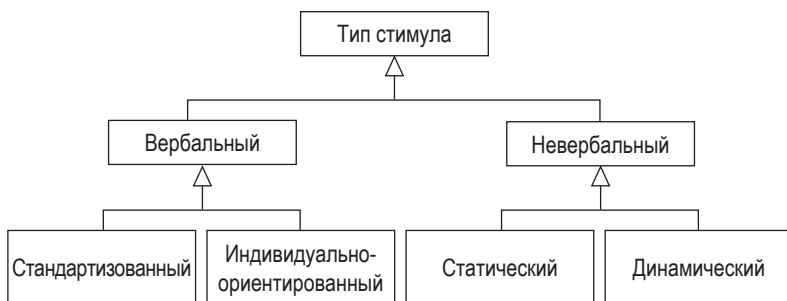


Рис. 4.11. Типы лингводидактических стимулов

При формировании классов ожидаемых реакций, так же как и при формировании классов лингводидактических стимулов, будем учитывать описание лингводидактических методов, приведенное в подразделе 4.3, и описание методов, используемых в психодиагностике. Ожидаемые реакции классифицируются по возможности, которая предоставляется ученику для формирования реакции на классы:

- закрытый набор ожидаемых реакций;
- открытый набор ожидаемых реакций;
- динамическое реагирование.

Закрытый набор формирования ожидаемой реакции предполагает, что все ожидаемые реакции образуют группу конечного количества заранее детерминированных реакций. Открытый набор формирования ожидаемых реакций предполагает, что ученик формирует реакции в свободной форме. При этом могут быть регламентированы лишь самые общие требования к форме реакции (вербальная или невербальная, максимальный объём или время реакции). При динамическом реагировании ученик формирует релевантную реакцию при помощи клавиатуры и/или манипулятора «мышь», используя свои перцептивные и моторные навыки.

Закрытый набор ожидаемых реакций классифицируется по способу формирования релевантной реакции на следующие классы:

- реакции, выбираемые из «меню»;
- реакции, формируемые оценением признака;
- реакции, формируемые восстановлением целого из частей;
- реакции, формируемые переструктурированием.

Реакция, выбираемая из «меню», формируется учеником выбором одной из альтернативных реакций, предлагаемых на сцене. Реакция, формируемая оценением признака, предполагает, что ученик осуществляет оценку некоторого признака стимула, а *шкала оценивания предлагается ученику*. Реакция, формируемая восстановлением целого из частей, предполагает, что ученик «собирает» предложение или фигуру из частей, которые презентуются стимулом. Реакция, формируемая переструктурированием предполагает, что ученику предлагается изменить структуру стимула в соответствии с заданной инструкцией.

Открытый набор ожидаемых реакций классифицируется по способу формирования релевантной реакции на следующие классы:

- реакции, формируемые дополнением;
- реакции, формируемые свободным конструированием.

Реакция, формируемая дополнением, предполагает, что ученик дополняет стимул некоторыми данными, которые ни-

как не регламентируются, а реакция, формируемая свободным конструированием, предоставляет ученику полную свободу в формировании реакции. На рис. 4.12 приведена диаграмма, иллюстрирующая множество типов ожидаемых реакций.



Рис. 4.12. Типы ожидаемых реакций

4.4.1. Диагностические сцены

Когнитивный тьютор осуществляет *персонализированное обучение*, которое мы понимаем как адаптацию обучающего материала к текущему уровню знаний ученика и степени понимания им учебного материала, а также адаптацию обучающего материала к текущим когнитивным способностям ученика. Такое понимание персонализированного обучения предполагает рассмотрение методов, используемых в психодиагностике, поскольку именно эта наука накопила наибольший опыт в области диагностики как общепсихологических, так и когнитивных характеристик человека.

В общем случае тип лингводидактической сцены детерминируется типами лингводидактических стимулов и типами ожидаемых реакций. Типология диагностических сцен должна учитывать ещё один классификационный признак — *способ анализа набора данных*, полученного во время диалоговой сессии, осуществляющей психодиагностический эксперимент.

Рассмотрим варианты основных типов диагностических сцен, которые сформируем путём комбинации типов диагностических стимулов, приведенных на рис. 4.11, и типов ожидаемых реакций ученика, приведенных на рис. 4.12. Каждый вариант диагностической сцены дополним описанием способа анализа наборов данных, полученных в результате психодиагностического эксперимента с использованием данного варианта сцены.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 1

Диагностический стимул:

вербальный, стандартизованный.

Реакция ученика:

закрытый набор, выбор из "меню".

Сочетание вербальных стандартизованных стимулов с ожидаемыми реакциями, формируемыми из закрытого набора ожидаемых реакций выбором из «меню», определяет самый обширный класс диагностических сцен, используемых в психодиагностических тестах-опросниках. Стимулы в диагностических сценах типа 1 представлены в виде вербальных вопросов. В «меню» ожидаемых реакций чаще всего используется двух- или трёх-альтернативный набор. Например, «Верно», «Неверно», «Не могу сказать». Тесты-опросники условно разделяют на: опросники-анкеты, биографические опросники и личностные опросники [114,116–118].

Алгоритмы обработки наборов данных, полученных в результате психодиагностического эксперимента, для тестов-о-

просников, как правило, представляют собой несложную процедуру подсчета количества совпадения ответов ученика с так называемым *диагностическим ключом* и последующее приведение полученного результата к нормированному виду.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 2

Диагностический стимул:
невербальный, статический.

Реакция ученика:
закрытый набор, выбор из «меню».

Диагностические сцены типа 2 часто применяются в психодиагностических тестах для исследования пространственного воображения, комбинаторных способностей и способностей соотнесения стимула к определённом классу из заданного множества классов.

Обработка наборов данных, полученных в результате психодиагностического эксперимента на основе диагностической сцены 2, сводится к оценке количества правильно произведенных выборов. В качестве диагностического параметра нередко используется время решения заданий теста.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 3

Диагностический стимул:
вербальный, стандартизованный
невербальный, статический.

Реакция ученика:
закрытый набор, восстановление целого из частей.

Психодиагностические тесты, построенные на основе диагностической сцены типа 3, являются вариантами тестов с множественным выбором. Отличительной особенностью этих

тестов является то, что ученику предлагается самому сконструировать реакцию-ответ, не прибегая к «меню» возможных реакций. С одной стороны, ученик не ограничен какими-либо рамками при конструировании ответа, с другой — предполагается, что существует правильный ответ-шаблон.

Примером психодиагностического теста, построенного на основе диагностической сцены типа 3, может служить тест Векслера нахождения недостающих деталей [119]. При помощи теста Векслера изучаются особенности зрительного восприятия обучаемого, наблюдательность и способность различить существенные детали.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 4

Диагностический стимул:

вербальный, стандартизованный
невербальный, статический.

Реакция ученика:

закрытый набор, переконструирование.

Диагностические сцены типа 4, включающие реакции типа переконструирование, предполагают, что ученик составляет целую сущность путём комбинации заданного набора вербальных или невербальных элементов. Как правило, существует заранее известная правильная комбинация элементов, но это не является обязательным условием, поскольку при оценке правильности могут учитываться такие параметры, как, например, оригинальность созданной комбинации. В качестве типичного примера психодиагностического теста, построенного на основе диагностической сцены типа 4, можно привести тест последовательности картинок в шкале измерения интеллекта Векслера. С помощью этого теста исследуется способность ученика к организации фрагментов в логическом поле, понимание и предвосхищение событий [116].

Обработка наборов данных, полученных в результате психодиагностического эксперимента на основе диагностических сцен 3 и 4, содержит как количественный, так и качественный компоненты. Но в целом ограниченный набор элементов, подвергающихся переструктурированию, и, соответственно, ограниченный и заранее известный набор правильных комбинаций даёт основание отнести соответствующие тесты к хорошо структурированным и допускающим высокую степень формализуемости процедуры обработки.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 5

Диагностический стимул:

вербальный, стандартизованный
невербальный.

Реакция ученика:

закрытый набор, оценивание признака по заданной шкале.

Психодиагностические тесты, построенные на основе диагностической сцены типа 5, связаны с оценкой учеником в элементах стимула (словесных утверждениях, изображениях предметов или лиц и т. п.) качества, заданного шкалой (например, «теплый — холодный», «сильный — слабый» и т. д.). Примером может служить тест семантического дифференциала, используемый для измерения различий в интерпретации понятий [114].

При использовании техники семантического дифференциала для оценки множества объектов одним учеником или одного объекта несколькими учениками формируются числовые таблицы, содержащие данные в виде пар «объект — признак», которые в дальнейшем могут быть подвергнуты анализу многомерных группировок как объектов, так и признаков. Этот анализ позволяет выявлять особенности психосемантики и отношений к объектам и явлениям. Метод семантического дифференциала позволяет получить информацию о «понятийной

структуре» ученика, направленности и особенностях личностного общения, идентификации себя с окружающими.

Обработка наборов данных, полученных в результате применения метода семантического дифференциала трудно формализуема и, как считается, основную работу по обработке полученных данных должен выполнять психодиагност, который помимо своих профессиональных знаний должен хорошо представлять особенности алгоритмов выявления структуры многомерных данных.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 6

Диагностический стимул:

вербальный, индивидуально-ориентированный
невербальный.

Реакция ученика:

закрытый набор, оценивание признака по заданной шкале.

Примером психодиагностического теста, построенного на основе диагностической сцены типа б, является техника репертуарных решёток. Такая техника, предложена Келли в 1955 году и направлена на изучение индивидуально-личностных конструктов, опосредующих восприятие и самовосприятие при анализе личностного смысла понятий [116]. Несмотря на внешнее сходство техники репертуарных решёток и метода семантического дифференциала, между ними существует радикальное отличие. В технике репертуарных решёток используются не заданные и стандартизованные элементы стимула, а предлагаются индивидуальные элементы, создаваемые непосредственно в ходе тестирования ученика. Поэтому техника репертуарных решёток является гораздо более гибким и тонким инструментом. Однако эти качества обуславливают значительную вовлеченность психодиагноста в процесс тестирования. Хотя для эффективного использования техники репертуарных решёток совершенно необходима компьютерная поддержка

методики, в самой процедуре обследования и интерпретации результатов акцент ещё более смещается в сторону использования трудно формализуемых знаний и опыта эксперта психодиагноста.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 7

Диагностический стимул:
вербальный, стандартизованный.

Реакция ученика:
открытый набор, дополнение.

Примером психодиагностического теста, построенного на основе диагностической сцены типа 7, может служить методика завершения предложений. Ученику даётся серия незаконченных предложений, состоящих из одного или нескольких слов, и предлагается завершить их по своему усмотрению. Предложения формулируются таким образом, чтобы стимулировать ученика формировать реакции, относящиеся к изучаемым свойствам его личности. Другим примером психодиагностического теста, построенного на основе диагностической сцены типа 7, является методика завершения историй. Здесь в отличие от предыдущей методики стимулом служит небольшой рассказ.

Стандартизация обеих методик осуществляется отнесением произвольных ответов ученика к некоторому опорному множеству психологических категорий.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 8

Диагностический стимул:
невербальный, статический.

Реакция ученика:
открытый набор, дополнение.

Примером психодиагностического теста, построенного на основе диагностической сцены типа 8, является рисуночный тест Вартегга [120]. При проведении такого теста ученику предлагается выполнить рисунок. Рисунок необходимо выполнить в заданном и ограниченном пространстве, в котором уже изображены некоторые элементы. В процессе тестирования фиксируется последовательность выполнения рисунков учеником. После завершения выполнения задания ученику предлагается ответить на ряд вопросов. Например, «Какой рисунок более удался?», «Каково содержание рисунков?» и т. д. При интерпретации результатов тестирования исходят из определённых *свойств, приписываемых графическим знакам*. Например, точке приписывают свойства органичности и неопределённости; квадрату — свойства угловатости, темноты, тяжести и т. д.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 9

Диагностический стимул:

вербальный, стандартизованный
невербальный.

Реакция ученика:

открытый набор, свободное конструирование.

Для психодиагностических тестов, предполагающих реакции обучаемого, формируемые из открытого набора типа «свободное конструирование» стимулы могут быть самыми разнообразными — как вербальными, так и невербальными. То же самое можно сказать и о типах реакций ученика — это могут быть рисунки на созвучную тему, рассказы, интерпретации изображений и т. д. Значительная часть психодиагностических тестов, построенных на основе диагностической сцены типа 9, приходится на проективные тесты.

Психодиагностика с помощью проективных тестов осуществляется на основе анализа взаимодействия ученика с внешне нейтральным материалом, который, в силу недостаточ-

ной структурированности, становится стимулом, порождающим фантазию и воображение ученика [116]. К проективным тестам относятся и рассмотренные выше методики с реакциями в виде дополнения стимулов. Однако целесообразно выделить методики, использующие реакции в виде открытого набора со свободным конструированием в отдельную группу, поскольку они ещё менее формализованы.

В качестве примеров проективных тестов приведём два популярных теста — тест Роршаха, в котором ученик должен придать смысл симметричным аморфным чёрно-белым и цветным изображениям [121], и тест тематической аперцепции, основанный на толковании учеником сюжетов специально подобранных картинок [122].

При обработке наборов данных, полученных в результате применения проективного теста, количественные диагностические оценки могут быть получены на основании измерения объёма реакции ученика, подсчёта частоты обращения к отдельным темам и т. п. На практике оценка результатов применения проективных методик часто опирается на интуицию и навыки психодиагноста.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 10

Диагностический стимул:

невербальный, динамический.

Реакция ученика:

динамическое реагирование.

Комбинация типов диагностического стимула и реакции ученика в диагностической сцене типа 10 соответствует классу психодиагностических тестов, которые обычно называют аппаратными тестами. Эти тесты используются в исследованиях параметров времени реакции и типологических особенностей высшей нервной деятельности. Наличие обратной связи между реакциями ученика и стимулами свойственно большому коли-

честву критериально-ориентированных аппаратурных тестов, в которых моделируются условия какой-либо критериальной деятельности. В традиционной психодиагностике реализация этих методик требовала специального и часто дорогостоящего оборудования. Современные средства программирования графического интерфейса прикладных программ открывают новые перспективы развития аппаратурных методик. На экране монитора могут моделироваться разнообразные виды деятельности, имитироваться объекты слежения, управления и т. д. Параллельно с помощью специальных датчиков может производиться съём и ввод в компьютер психофизиологической информации. Отдельно можно выделить также подкласс аппаратурных тестов, который в настоящее время стал активно развиваться и в котором моделирование опосредуется при помощи компьютерных игр.

ДИАГНОСТИЧЕСКАЯ СЦЕНА 11

Диагностический стимул:

вербальный, индивидуально-ориентированный.

Реакция ученика:

открытый набор.

Психодиагностические тесты, построенные на основе сцены типа 11, можно определить как сугубо диалоговые, в которых предполагается непосредственный контакт психодиагностом с учеником и учитываются специфические особенности конкретного психодиагностического эксперимента.

4.4.2. Обучающие стимулы и сцены

При помощи обучающей сцены тьютор (активный агент) передаёт ученику (реактивный агент) порцию знаний, которую должен понять и запомнить ученик. Обучающая сцена — это

всегда интеррогативный стимул, являющийся либо задачей, подлежащей решению, либо вопросом, на который необходимо ответить. В традиционной лингводидактической практике, в зависимости от выбранного метода обучения, отдаётся предпочтение различным обучающим стимулам, которые в лингводидактическом тьюторе отображаются в обучающие сцены.

Обучающие сцены могут классифицироваться так же, как и диагностические сцены — комбинацией типа лингводидактического стимула (рис. 4.11) и типа множества ожидаемых реакций (рис. 4.12).

Проведём анализ несколько примеров типичных обучающих стимулов, применяемых в различных классических методах обучения иностранным языкам, и рассмотрим возможность их имитации в виде обучающих сцен, даталогическая структура которых приведена на рис. 3.20.

В ситуативном методе обучения часто используются следующие обучающие стимулы [103,123].

Обучающий стимул 1. Учитель держит в руках наручные часы и произносит вслух: «This is a watch». Потом он указывает на настенные часы и говорит «That is a clock». Цель такого обучающего стимула — научить учащихся понимать и различать выражения: «This is a...» и «That is a ...». Стимул воздействует одновременно на зрительный и слуховой анализаторы.

Такой обучающий стимул может быть реализован с помощью обучающей сцены прямой рефлексией, когда описанный выше обучающий стимул имитируется в виде цепи видео- и аудио-объектов, включающих изображения наручных и настенных часов и синхронное воспроизведение звука.

Обучающий стимул 2. Учащиеся хором повторяют слова и фразы, ранее произнесённые учителем. Цель стимула такого рода — обучить учащихся правильному произношению. Стимул воздействует на слуховой анализатор и не требует вовлечения моторных реакций для формирования ответа.

Такой обучающий стимул крайне сложно реализовать с помощью обучающей сцены прямой рефлексией, но можно предложить другую сцену, *эквивалентную рассматриваемо-*

му обучающему стимулу с точки зрения достижения цели. Например, сцена после соответствующих инструкций, выводимых при помощи видео-объектов, произносит фразу (аудио-объект), которую затем повторяет ученик. Контроль произношения осуществляется с помощью микрофона и программы, реализуемой при помощи процесса-демона *AnsDemon* (смотри рис. 3.27).

Обучающий стимул 3. Учитель последовательно демонстрирует и называет учащимся небольшие предметы, извлекая их из коробки. Учащиеся повторяют вслух наименования предметов. Цель такого стимула двоякая: обучить учащихся правильному произношению и установить ассоциации произносимых наименований предметов с их зрительными образами. Стимул воздействует одновременно на зрительный и слуховой анализаторы.

Такой обучающий стимул может быть имитирован с помощью обучающей сцены следующим образом. На сцене появляется видео-объект, содержащий изображение коробки с предметами. Учащийся, последовательно нажимая одну из клавиши, «извлекает» предметы из коробки. При каждом извлечении предмета из коробки аудио-объект сцены произносит наименование предмета, а на сцене появляется видео-объект с просьбой повторить вслух это наименование. Для проверки правильности произношения используется микрофон и процесс-демон.

Для аудио-лингвистического метода характерны следующие обучающие стимулы [105].

Обучающий стимул 4. Учащемуся предлагается предложение, из которого удалён ряд слов. Учащийся должен восстановить предложение с минимальным искажением исходного предложения. Цель такого стимула — обучение учащегося применению конкретной грамматической конструкции. Стимул воздействует на зрительный анализатор и при формировании реакции требует применения моторных навыков в виде письма.

Такой обучающий стимул может быть реализован с помощью обучающей сцены прямой рефлексией. На сцену выводит-

ся видео-объект, содержащий предложение, из которого удалён ряд слов. Ученику предлагается вписать удалённые слова при помощи клавиатуры. Моторная реакция ученика реализуется работой с клавиатурой.

Обучающий стимул 5. Учитель предлагает паре учащихся прочитывать вслух диалог. Цель такого обучающего стимула — научить учащихся воспринимать речь и правильно произносить фразы. Стимул воздействует на слуховой анализатор.

Ясно, что такой стимул невозможно реализовать в обучающей сцене прямой рефлексией, но можно разработать сцену, при помощи которой может быть достигнута сформулированная цель обучения. Например, вначале при помощи аудио-объекта сцены произносится фраза диалога, затем при помощи видео-объекта на сцену выводится текст последующей фразы диалога, а также просьба произнести эту фразу вслух. Контроль произношения ученика осуществляется при помощи микрофона и процесса-демона.

Для метода общего физического отклика характерны следующие обучающие стимулы [112].

Обучающий стимул 6. Учитель предлагает учащимся устные команды, которые они должны понять и выполнить. Стимул воздействует на слуховой анализатор и предполагает моторную реакцию. Цель стимула — тренировка понимания устной речи.

Для реализации обучающей сцены, имитирующей подобный обучающий стимул, можно использовать следующий подход. Ученик работает со сценой, на которой при помощи видео-объекта изображен, например, интерьер некоторого помещения. При помощи аудио-объектов ученик получает аудио-команды о необходимости выполнять некоторые действия. Например, указать на конкретный предмет интерьера или передвинуть его. Ученик выполняет аудио-команды, используя манипулятор типа «мышь».

Обучающий стимул 7. Учитель вслух задаёт простые вопросы, на которые учащийся должен отвечать жестом, например указывая на предмет рукой. Цель этого стимула такая же, как

и в предыдущем случае — получение навыков понимания устной речи. Стимул воздействует на слуховой и зрительный анализаторы, а также предполагает моторную реакцию.

Реализация обучающей сцены для этого стимула может быть осуществлена способом, описанным для случая обучающего стимула 6. При этом единственным действием, которое выполняет обучаемый, является выбор предмета при помощи манипулятора «мышь».

В [124] приведен ряд упражнений, из которых выделим следующие обучающие стимулы 8 и 9.

Обучающий стимул 8. Учитель предлагает учащимся небольшой фрагмент текста, в котором в неявной форме описано некоторое понятие. Задача учащегося — определить, о каком понятии идёт речь в тексте.

Обучающий стимул 9. Учитель предлагает учащимся текст, из которого удалено несколько слов. Учащиеся должны прочесть текст и попробовать заполнить пустые места словами, которые соответствуют тексту по смыслу.

Целью обучающих стимулов 8 и 9 является изучение, понимание и запоминание новых слов с использованием контекста. Этот вид стимулов помогает учащемуся понять, что благодаря контексту можно легче выяснить значение трудных или незнакомых слов. Стимул воздействует преимущественно на зрительный анализатор и требует вовлечения моторных реакций.

Поскольку обучающие стимулы 8 и 9 имеют одинаковую цель, они могут быть реализованы при помощи следующих однотипных обучающих сцен. На сцену при помощи видео-объекта выводится некоторый текст, который находится на экране в течение фиксированного промежутка времени. Затем, при помощи последующего видео-объекта учащемуся предлагается список возможных заголовков этого текста. Учащийся должен выбрать подходящий заголовок.

В [125] рассмотрены примеры обучающих сценариев, основанных на коммуникационном методе обучения (communicative methodology), для которого характерны следующие обучающие стимулы.

Обучающий стимул 10. Учитель предлагает учащемуся текст. Учащийся, прочитав текст, должен выразить его содержание различными средствами (с помощью другого текста, рисунка и т.д.). Цель стимула — развитие навыков чтения и понимания текстового материала, а также умения выразить полученные знания.

Можно предложить несколько вариантов реализации обучающей сцены для обучающего стимула 10. Описанные варианты сцен достигают ту же цель, что и приведенный обучающий стимул. Вариант 1. Сцена предъявляет ученику при помощи нескольких видео-объектов текст и список вопросов, релевантных тексту. Ученик должен выбрать в тексте предложения, которые являются ответами на предъявленные вопросы. Вариант 2. Сцена предъявляет ученику при помощи нескольких видео-объектов фрагменты текста в произвольном порядке. Ученик должен расположить фрагменты текста в правильном порядке.

Проведенный анализ типовых обучающих стимулов классических методов обучения иностранным языкам позволяет сделать несколько заключений о возможности их реализации при помощи обучающих сцен, даталогическая структура которых приведена на рис. 3.20.

Реализация реального обучающего стимула в виде обучающей сцены всегда предполагает трансформацию реального обучающего стимула. В некоторых случаях проблема трансформации не вызывает затруднений. Стимул трансформируется методом, названным выше непосредственной или прямой рефлексией. Прямая рефлексия часто предполагает разработку анимационного фрагмента, копирующего реальный обучающий стимул. В ряде случаев (например, случай хорового пения) непосредственная рефлексия неприменима. В этом случае предлагается конструировать обучающие сцены в соответствии с *принципом эквивалентности целей*. Предполагается, что в каждом обучающем стимуле можно выделить некую цель, заложенную в обучающий стимул учителем. Предполагается также, что достижение цели обучающего стимула важнее способа её достижения, заложенного в исходный обучающий стимул.

Принцип эквивалентности целей означает, что реальный обучающий стимул, который невозможно отобразить в обучающую сцену непосредственной рефлексией, заменяется другим, эквивалентным реальному в смысле поставленной цели, но легко отображаемым в обучающую сцену, которая реализуется средствами даталогической модели.

Модель искусственного активного агента вопросно-ответного диалога (рис. 3.27) и даталогическая модель сцены (рис. 3.20) были апробированы при конструировании ряда лингводидактических тьюторов.

Ниже приведено краткое описание лингводидактического тьютора «English Grammar in Use», с акцентом на реализацию различных обучающих сцен. Одна из причин, по которой описание этого тьютора включено в данный раздел, является желание проиллюстрировать эффективность даталогического подхода при проектировании сценария. При создании тьютора «English Grammar in Use» не было написано ни одной строчки программного кода. Все особенности, как обучающих сцен, так и методики обучения, реализуются даталогически.

Лингводидактический тьютор «English Grammar in Use» создан на основе материалов известного британского учебника «English Grammar in Use. A self-study reference and practice book for intermediate students», разработанного Раймондом Мэрфи и изданного в Cambridge University Press в 1988 году. Выбор отмеченного учебника в качестве прототипа сценария обусловлен тем, что его материал хорошо структурирован, поэтому относительно легко разбивается на последовательности обучающих сцен. Основное внимание при разработке тьютора «English Grammar in Use» уделялось проектированию обучающих сцен. Логика связи обучающих сцен в диалоговый обучающий сценарий относительно проста и исчерпывается следующими возможностями: (1) альтернативный выбор учеником одного из ожидаемых ответов, присутствующих на сцене; (2) заполнение одного или нескольких полей ввода последовательностью символов.

Тьютор «English Grammar in Use» хорошо структурирован в том смысле, что состоит из большого количества модулей, каждый из которых содержит материал примерно одинакового уровня сложности. Модули слабо связаны между собой и могут изучаться независимо. На рис. 4.13 и 4.14 приведены примеры начальных обучающих сцен двух различных модулей.

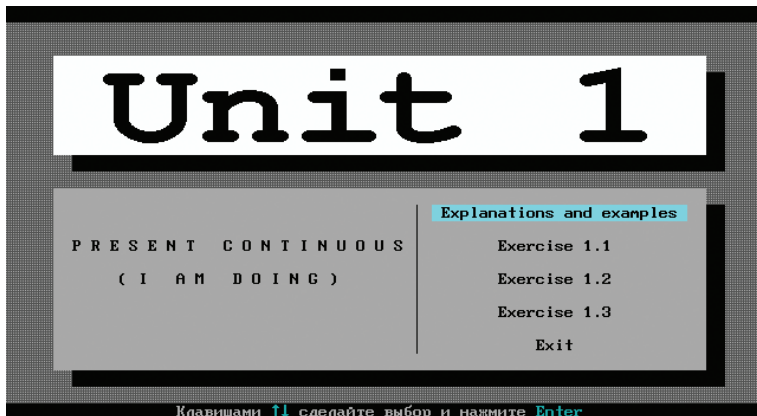


Рис. 4.13. Первая сцена модуля «Unit 1»

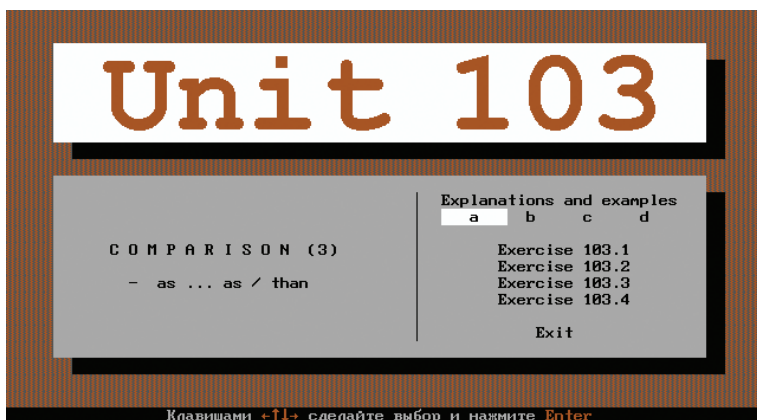


Рис. 4.14. Первая сцена модуля «Unit 103». Раздел «explanations and examples» включает четыре подраздела

Как видно из примеров обучающих сцен на рис. 4.13 и 4.14, тьютор «English Grammar in Use» не имеет стандартизованного интерфейса. На каждом шаге диалогового взаимодействия он предоставляет ученику интуитивно понятную сцену, которая включает интеррогативный обучающий стимул и список ожидаемых ответов. Такой подход к разработке интерфейса может рассматриваться как вариант интерфейса, адаптирующегося к текущей точке диалогового процесса. Вместо стандартизованного интерфейса в тьюторе используется стандартизованный стиль сцены, заключающийся в том, что обучающий стимул всегда располагается в верхней части экрана, а список ожидаемых реакций обучаемого — в нижней части экрана. Отметим, что использование такого стиля при проектировании обучающих сцен исключает появление тупиковых ситуаций, когда ученик не знает как продолжить диалоговый процесс.

Одной из особенностей тьютора «English Grammar in Use» является использование *динамических обучающих сцен*. В динамической обучающей сцене как объекты обучающего стимула, так и объекты ожидаемых реакций появляются на экране не одновременно, а последовательно. Таким образом, динамическая сцена постепенно разворачивается на экране. Это позволяет ученику легко воспринимать даже сложную сцену.

Рис. 4.15–4.18 иллюстрируют последовательную интерпретацию динамической сцены. С целью уменьшения количества рисунков ряд промежуточных фаз интерпретации динамической сцены пропущен.

Динамические обучающие сцены проектируются так, чтобы темп появления видео-объектов сцены и темп заполнения их информацией был психологически комфортен для обучаемого. Для этой цели используются соответствующие значения атрибутов поступательного перемещения, деформации, времени жизни при описании объектов сцены. На начальной фазе интерпретации динамической сцены, приведенной на рис. 4.15, в левую часть сцены выводится видео-объект, содержащий графическое изображение, которое в дальнейшем будет дополнено видео-объектами с текстом.

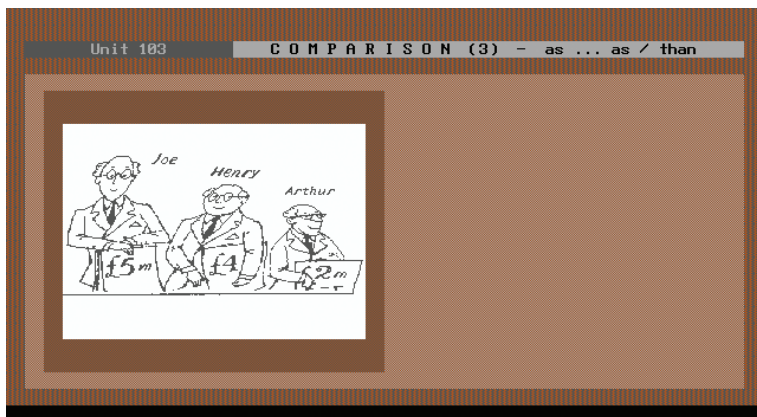


Рис. 4.15. Начальная фаза интерпретации динамической сцены

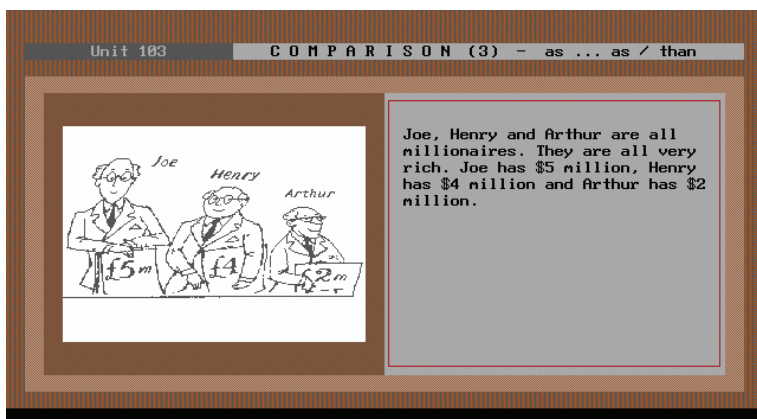
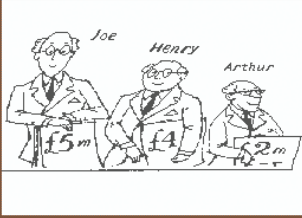


Рис. 4.16. Промежуточная фаза интерпретации динамической сцены

Рис. 4.17 иллюстрирует момент, когда на сцене проинтерпретированы все видео-объекты, составляющие обучающий стимул сцены, однако отсутствует список ожидаемых реакций. Текст в правом видео-объекте появляется порционно в темпе, достаточном для его комфортного чтения и восприятия. Темп интерпретации выбран в расчёте на способности «слабого» ученика.

Unit 103 C O M P A R I S O N (3) - as ... as / than



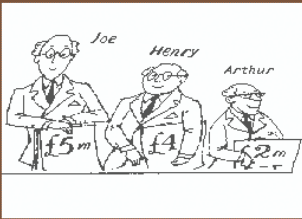
Joe, Henry and Arthur are all millionaires. They are all very rich. Joe has \$5 million, Henry has \$4 million and Arthur has \$2 million.

So:

- + Henry is rich.
- + He is richer than Arthur.
- + But he isn't as rich as Joe.
(= Joe is richer than he is)

Рис. 4.17. Ещё одна промежуточная фаза интерпретации динамической сцены

Unit 103 C O M P A R I S O N (3) - as ... as / than



Joe, Henry and Arthur are all millionaires. They are all very rich. Joe has \$5 million, Henry has \$4 million and Arthur has \$2 million.

So:

- + Henry is rich.
- + He is richer than Arthur.
- + But he isn't as rich as Joe.
(= Joe is richer than he is)

Enter – продолжение работы, F2 – словарь, Esc – к началу блока.

Рис. 4.18. Заключительная фаза интерпретации динамической сцены. Сцена дополнена списком ожидаемых реакций

Темп интерпретации, ориентированный на «слабого» ученика может оказаться слишком медленным для более подготовленного ученика. Поэтому в тьюторе предусмотрена возможность управления темпом интерпретации сцены самим

учеником. Ученик при желании, может «мгновенно» завершить интерпретацию текущего видео-объекта. Для этого необходимо одновременно нажать клавиши: «Alt-W». Для мгновенного завершения интерпретации всей сцены необходимо одновременно нажать клавиши «Alt-S».

Динамические сцены часто используются при проектировании сцен-упражнений, предполагающих моторную реакцию ученика. В ряде упражнений ученику предлагается заполнить недостающие фрагменты предложений, используя конечный набор слов, выводимых в отдельном видео-объекте в виде списка. На рис. 4.19 приведена начальная фаза интерпретации такой динамической сцены-упражнения.

В случае, когда ученик сделал правильный выбор, слово перемещается по сцене и занимает нужное положение в предложении. На рис. 4.20 показана работа ученика с упражнением. Первое выбранное слово «age» уже заняло своё место в предложении, а только что выбранное слово «you» перемещается по экрану.

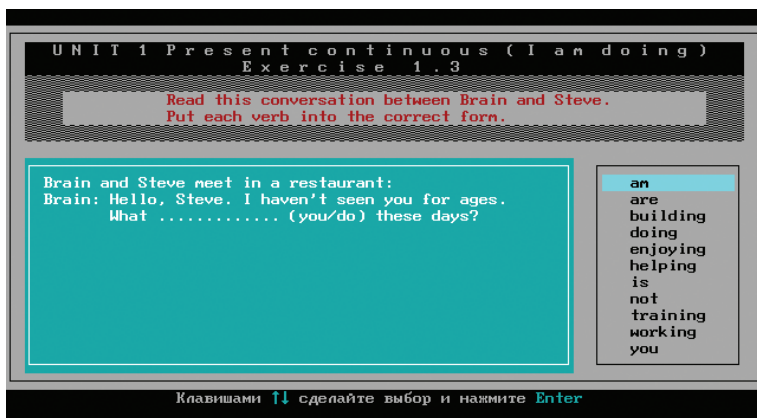


Рис. 4.19. Начальная фаза интерпретации динамической сцены-упражнения. Заполнение недостающего фрагмента предложения осуществляется выбором слов из предлагаемого списка

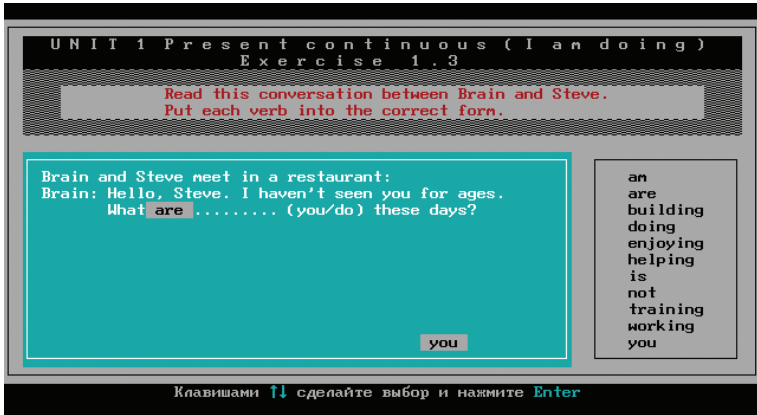


Рис. 4.20. Работа ученика с динамической сценой-упражнением

На рис. 4.21 показан момент завершения работы ученика с первым предложением. Словосочетание «are you» заняло своё место, а заключительное слово «doing» перемещается по экрану.



Рис. 4.21 Завершение работы ученика с первым предложением

Если во время выполнения упражнения рассматриваемого типа ученик сделал ошибку выбора, он получает контекстно

зависимое замечание. Рис. 4.22 и 4.23 иллюстрируют «обработку» сценарием отмеченной ошибки ученика.

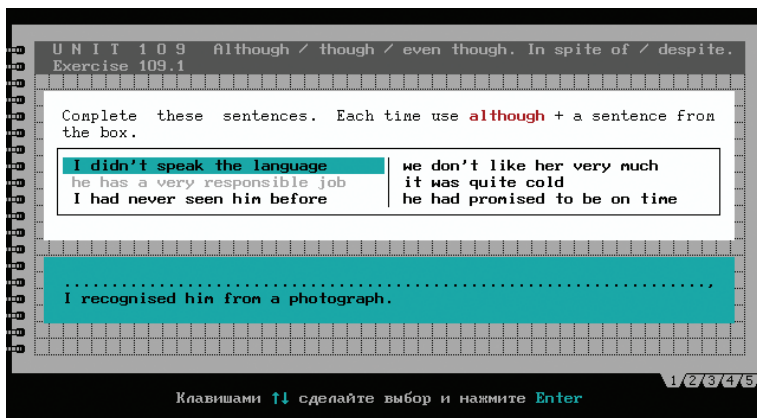


Рис. 4.22. Сцена-упражнение, в которой ученик совершает ошибку

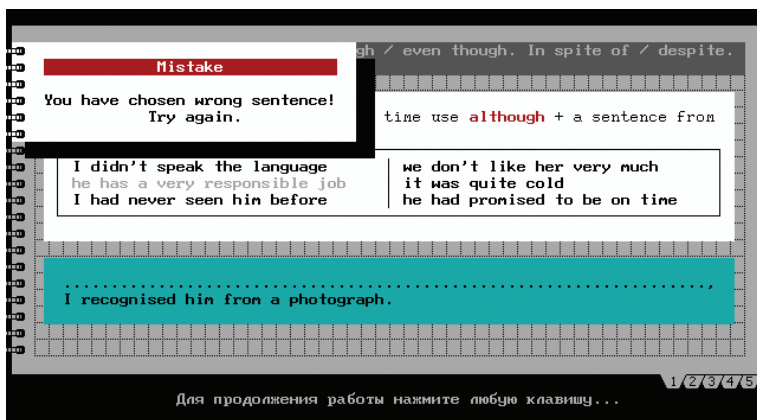


Рис. 4.23. Сцена-упражнение, приведенная на рис. 4.22, и дополненная сообщением об ошибке

При работе над тьютором «English Grammar in Use» соблюдались общие соглашения о логической структуре сцены, способах организации «теоретических» сцен и сцен-упражнений.

Однако дизайн сцен от раздела к разделу изменялся, оставаясь однотипным в пределах одного раздела. Такой подход исключает «уставание» ученика от однообразного дизайна обучающих сцен. Рис. 4.24 и 4.25 иллюстрируют иной вариант дизайна сцены-упражнения и «обработки» ошибки.



Рис. 4.24. Сцена, в которой ученик совершает ошибку



Рис. 4.25. Сцена-упражнение, приведенная на рис. 4.24, и дополненная сообщением об ошибке

4.5. Диагностика и логический вывод

Как обучающие, так и диагностические сцены, описанные в подразделе 4.4, могут быть реализованы даталогическим средствами сцены (смотри рис. 3.20) и алгоритмами процессов-демонов, входящих в алгоритмическую часть модели активного агента диалога (смотри рис. 3.27).

Диагностические сцены используются в сценариях когнитивных тьюторов для определения таких когнитивных характеристик ученика, как текущий уровень знаний о предмете обучения и когнитивный профиль. Как правило, психодиагностические методики предполагают, что в процессе диалоговой диагностической сессии формируется набор данных, анализ которого позволяет сделать вывод о когнитивных характеристиках ученика. Анализ набора данных для ряда психодиагностических методик (например, тесты-опросники и тест Векслера) может осуществляться при помощи процессов-демонов, алгоритмы которых базируются на сравнении полученного набора данных с эталонными шаблонами. Отметим, что в основе такого способа обработки результатов психодиагностического теста лежит использование заключений, полученных экспертом психодиагностом в результате *ранее проведенных дедуктивных умозаключений*.

Анализ наборов данных, полученных в результате применения психодиагностических методик, построенных с использованием сложных диагностических сцен, таких как сцены типов 8–11, осуществляется непосредственно экспертом психодиагностом. Есть все основания предполагать, что такой анализ представляет собой логический вывод, использующий процедурные знания эксперта и что психодиагност приходит к заключению о когнитивных характеристиках ученика в результате дедуктивных умозаключений.

Сетевые модели диалогового метода доступа *DiAM*, описанные в подразделах 3.3 и 3.4, в контексте логического вывода могут рассматриваться как *сети вывода*, а вопросно-ответный диалоговый процесс может использоваться для моделирования

дедуктивных умозаключений. Таким образом, для реализации психодиагностических методик, построенных на основе сложных диагностических сцен типов 8–11, может использоваться вопросно-ответный диалоговый процесс. Это открывает перспективы для конструирования лингводидактических сценариев, совмещающих диагностику и обучение.

Основной гипотезой, составляющей основу исследований данного подраздела, является предположение о том, что вопросно-ответный диалоговый процесс обладает свойствами процесса логического вывода на основе дедуктивных умозаключений и может рассматриваться как средство *непрерывной диагностики ученика*.

4.5.1. Диалог и логический вывод в экспертных системах

Экспертные системы являются классом прикладных программ, ориентированных на решение плохо формализованных задач методами, которые моделируют логические умозаключения эксперта. Таким образом, сфера применимости экспертных систем пересекается со сферой применимости вопросно-ответных диалоговых систем. Два процесса, важные для функционирования диалоговых систем, являются также неотъемлемой частью функционирования экспертных систем: (1) *диалог* и (2) *дедукция*.

В этом подразделе рассмотрим свойства диалогового процесса, сопровождающего работу экспертной системы, а также методы, применяемые в экспертных системах при построении дедуктивных умозаключений. Рассматривается класс экспертных систем, в которых экспертные знания, хранящиеся в базе знаний, моделируются набором продукционных правил, поскольку в таких системах интенсивно используется диалоговый процесс и хорошо имитируется способность человека проводить дедуктивные умозаключения.

Диалог с пользователем является неотъемлемой частью функционирования экспертной системы. В процессе диалога

экспертная система может потребовать от пользователя посылки, которые необходимы ей для продолжения процесса вывода. Это происходит в тех случаях, когда эти посылки отсутствуют в базе знаний.

Категоризация экспертной системы как агента диалогового процесса с пользователем зависит от того, используется ли при функционировании экспертной системы её объяснительная компонента. В том случае, когда объяснительная компонента не используется, экспертная система выступает в роли активного агента, а пользователь — в роли реактивного агента. Экспертная система формирует вопрос таким образом, чтобы ответ пользователя содержал запрашиваемые факты-посылки, отсутствующие в базе знаний. Если используется объяснительная компонента, то диалог экспертной системы с пользователем относится к категории диалогов со свободным распределением ролей, а в процессе объяснений экспертная система выступает в роли реактивного агента. Ниже приведен фрагмент диалогового взаимодействия пользователя и экспертной системы MYCIN [63]. Экспертная система MYCIN была спроектирована для диагностирования бактерий. Вопросы и объяснения системы MYCIN записаны прописными буквами.

```
-----PATIENT-1-----
PATIENT'S NAME?
DAVID KING

AGE?
34

SEX?
WHY?
[I.E. WHY IS IT IMPORTANT TO DETERMINE THE SEX
OF PATIENT-1?]
[1.0] THE SEX OF A PATIENT IS ONE OF THE 4
STANDARD PARAMETERS THAT IDENTIFY A PATIENT.
THIS MAY BE USEFUL LATER IN THE CONSULTATION.
FOR EXAMPLE, IT IS RELEVANT TO DETERMINING
```

WHETHER AN ORGANISM IS NORMALLY FOUND AT THIS SITE AND IS RELEVANT IN DETERMINING PATIENT-1'S MOST RECENT CREATININE CLEARANCE (IN ML/MIN).

[BACK TO QUESTION 3]

MALE

RACE?

CAUCASIAN

ARE THERE ANY CULTURES FOR DAVID KING WHICH MAY BE RELATED TO THE PRESENT ILLNESS, AND FROM WHICH ORGANISMS HAVE BEEN GROWN SUCCESSFULLY IN THE MICROBIOLOGICAL LABORATORY?

WHY?

[I.E. WHY IS IT IMPORTANT TO DETERMINE ORGANISMS ISOLATED FROM PATIENT-1?]

[2.0] THIS AID IN DETERMINING THE THERAPEUTIC REGIMEN OF PATIENT-1. IF

[2.1] INFORMATION HAS BEEN GATHERED ABOUT ORGANISMS ISOLATED FROM THIS PATIENT, ORGANISMS NOTED ON SMEARS OF THIS PATIENT, NEGATIVE CULTURES OF THIS PATIENT, SUSPECTED INFECTIONS WITHOUT MICROBIOLOGICAL EVIDENCE, CURRENT DRUGS OF THIS PATIENT AND PRIOR DRUGS OF THIS PATIENT, AND

[2.2] AN ATTEMPT HAS BEEN MADE TO DEDUCE THE ORGANISMS WHICH REQUIRE THERAPY, AND

[2.3] YOU HAVE GIVEN CONSIDERATION TO ORGANISMS (OTHER THAN THOSE SEEN IN CULTURAL SMEARS) WHICH MIGHT BE PRESENT THEN DETERMINE THE BEST THERAPY RECOMMENDATION FROM AMONG THE DRUGS LIKELY TO BE EFFECTIVE AGAINST THE SIGNIFICANT ORGANISMS, OR INDICATE THAT NO THERAPY IS REQUIRED AT THIS TIME.

[RULE092]

[FIRST, OF COURSE, IT IS NECESSARY TO ASK ABOUT A POSITIVE CULTURE]

[BACK TO QUESTION 5]

WHY?

[I.E. WHY IS IT IMPORTANT TO DETERMINE THE THERAPEUTIC REGIMEN OF PATIENT-1?]
[3.0] DEDUCING THE THERAPEUTIC REGIMEN OF A PATIENT IS THE GOAL OF AN INFECTION DISEASE CONSULTATION

NO

ARE THERE ANY PENDING CULTURES FOR DAVID KING?
YES

-----CULTURE-1-----

FROM WHAT SITE WAS THE SPECIMEN FOR CULTURE-1 TAKEN?
CSF

PLEASE, GIVE THE DATE AND TIME WHEN THE PENDING CSF CULTURE (CULTURE-1) WAS OBTAINED (MO/DA/YR TIME)
JUNE/24/84 7:00

WAS A SMEAR EXAMINED AT THE TIME THAT THE SPECIMEN FOR THE PENDING CSF CULTURE (CULTURE-1) WAS OBTAINED?
YES

WERE ANY ORGANISMS SEEN ON THE SMEAR OF THE SPECIMEN SUBMITTED AS PENDING CSF CULTURE (CULTURE-1)?
NO

ARE THERE ANY OTHER PENDING CULTURE?
NO

HAVE THERE BEEN NEGATIVE (AS OPPOSITE TO PENDING) CULTURES OBTAINED FROM DAVID KING WHICH MAY BE RELEVANT TO THE PRESENT ILLNESS?
NO

DO YOU SUSPECT DAVID KING MAY HAVE AN INFECTION AT A SITE FROM WHICH YOU HAVE NOT OBTAINED CULTURE SPECIMENS?

NO

DURING THE PRESENT ILLNESS, HAS DAVID KING BEEN TREATED WITH AN ANTIMICROBIAL AGENT, WHICH IS NO LONGER BEING ADMINISTRATED?

NO

HAS DAVID KING RECENTLY HAD SYMPTOMS OF PERSISTENT HEADACHE OR OTHER ABNORMAL NEUROLOGICAL SYMPTOMS (DIZZINESS, LETHARGY, ETC.)?

YES

HAS DAVID KING RECENTLY HAD OBJECTIVE EVIDENCE OF ABNORMAL NEUROLOGICAL SIGNS (NUCHAL RIGIDITY, COMA, SEIZURES, ETC.) DOCUMENTED BY PHYSICIAN OBSERVATION OR EXAMINATION?

YES

THE CSF CULTURES WILL BE CONSIDERED TO BE ASSOCIATED WITH MENINGITIS.

Из приведенного фрагмента протокола диалога видно, что если в нём удалить все запросы Why?, вызывающие объяснительную компоненту, и следующие за ними объяснения, то оставшаяся часть является примером диалога с фиксированным распределением ролей, в котором MYCIN играет роль активного агента.

Диалог в экспертной системе MYCIN, как и в других экспертных системах, поддерживается отдельным интерфейсным блоком, в котором форма хранения данных отлична от формы хранения фактов и правил в базе знаний. Таким образом, для экспертной системы диалог является только формой общения с пользователем, а не методом решения проблемы диагностики.

Часто экспертные системы определяют, как класс программ, имитирующих способность человеческого интеллекта делать вывод построением дедуктивных цепей, используя имеющуюся совокупность фактов-посылок [126]. В экспертных системах, основанных на правилах (или продукционных экспертных

системах), знания, необходимые для проведения дедуктивных умозаключений, хранятся в виде фактов и связывающих их правил продукции. Факт, с точки зрения дедуктивных умозаключений — это либо посылка, либо заключение, либо заключение-посылка, если он располагается внутри дедуктивной цепи. Для того чтобы провести дедуктивный вывод, экспертной системе нет необходимости «понимать» значение фактов, достаточно оперировать лишь их именами [127].

Тип проблемы, решаемой при помощи экспертной системы, определяется двумя способами построения дедуктивной цепи, называемыми прямым выводом и обратным выводом. Цель прямого вывода можно определить, как переход от исходного набора фактов-посылок к результатному факту-заключению. При прямом выводе каждый из фактов исходного набора посылок связывается дедуктивными цепями с фактом-заключением. Целью обратного вывода является переход от исходного факта-заключения к результатному набору фактов-посылок. При обратном выводе исходный факт-заключение связывается дедуктивными цепями с каждым из фактов результатного набора фактов-посылок.

Таким образом, можно говорить о двух типах связей между фактами при проведении дедуктивных заключений.

1. Связи между фактами, детерминированные правилами продукции. Каждое правило связывает небольшое (обычно от 2-х до 4-х) количество фактов. Закона, позволяющего автоматически формировать правила продукции не существует. Совокупность правил продукции отображает процедурные знания эксперта (или группы экспертов) о предметной области.
2. Связи между фактами, детерминированные в процессе вывода при построении дедуктивных цепей. Дедуктивные цепи связывают между собой большое (десятки и сотни) количество фактов. Закон, управляющий формированием дедуктивных цепей, называется дедуктивным умозаключением.

Связи 1-го типа существуют в базе знаний экспертной системы *в явном виде* и выражаются в описании продукционных

правил, а связи 2-го типа — *в неявном виде*. Они формируются машиной вывода в процессе вывода. Однако в литературе, посвященной экспертным системам, основанным на правилах, можно встретить несколько способов явного графического представления связей 2-го типа. Связи 2-го типа можно представить явно графически при помощи следующих средств: (1) сеть вывода [128]; (2) дерево типа И/ИЛИ [129]; (3) дерево решений [130].

В основе того или иного способа явного представления дедуктивных цепей лежит способ графического изображения правила продукции. Существует множество нотаций и определений продукционных правил. Наиболее общее определение продукционных правил, используемых в системах искусственного интеллекта, приведено в [128].

$$\begin{array}{l} \text{ЕСЛИ } \text{условие}_1, \text{условие}_2, \dots \\ \text{ТО } \quad \text{действие}_1, \text{действие}_2, \dots \end{array} \quad (4.1)$$

Это определение отождествляет правило продукции с условным предложением *ЕСЛИ – ТО*. В левой части условного предложения, после слова *ЕСЛИ* записывается список условий. Если все эти условия имеют место, то должна быть выполнена последовательность действия, записанная в правой части условного предложения, после слова *ТО*.

В тех случаях, когда продукционное правило используется в логическом контексте, для построения дедуктивных заключений чаще используются понятия посылка и заключение, а правило продукции принимает вид:

$$\begin{array}{l} \text{ЕСЛИ } \text{факты-посылки} \\ \text{ТО } \quad \text{факт-заключение} \end{array} \quad (4.2)$$

В левой части правила продукции (4.2) записывается конъюнкция фактов-посылок. Если конъюнкция истинна, то имеет место факт-заключение. В процессе вывода факт-заключение, полученный в результате применения некоторого продукционного правила, может использоваться в качестве факта-посылки в последующем продукционном правиле.

На рис. 4.26 приведен графический символ продукционного правила, который используется для изображения сети вывода [128]. Используя этот графический символ правила продукции, можно графически представить все возможные цепи дедуктивных выводов для базы знаний, состоящей из заданного набора фактов и продукционных правил.

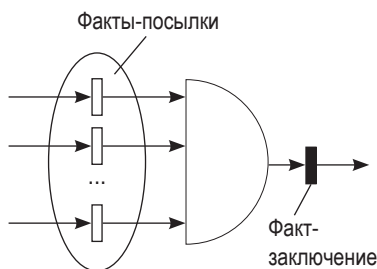


Рис. 4.26. Графический символ, используемый для изображения правил продукции в сети вывода

Рассмотрим пример, иллюстрирующий, каким образом знания, представленные в виде набора продукционных правил, могут быть связаны в сеть вывода с использованием графического символа, изображенного на рис. 4.26. Пусть имеется база знаний, состоящая из продукционных правил, связывающих характерные свойства млекопитающих с их видом. На рис. 4.27 приведен набор из шести продукционных правил ($R_1 - R_6$), которые можно использоваться для вывода заключения о принадлежности животного к конкретному виду при наличии соответствующих фактов предпосылок.

В зависимости от имеющихся посылок правила, приведенные на рис. 4.27, позволяют сделать пять прямых выводов (от фактов-посылок к фактам-заключениям) и получить пять заключений: (1) животное млекопитающее; (2) животное плотоядное; (3) животное копытное; (4) животное — это тигр; (5) животное — это зебра.

- R_1 : ЕСЛИ у животного есть волосы
ТО это млекопитающее
- R_2 : ЕСЛИ животное даёт молоко
ТО это млекопитающее
- R_3 : ЕСЛИ животное является млекопитающим
И у него заострённые зубы
И у него есть когти
И его глаза смотрят вперёд
ТО это плотоядное животное
- R_4 : ЕСЛИ животное является млекопитающим
И у него есть копыта
ТО это копытное животное
- R_5 : ЕСЛИ животное является плотоядным
И у него желтовато-коричневый цвет
И у него есть чёрные полосы
ТО это тигр
- R_6 : ЕСЛИ животное копытное
И у него белый цвет
И у него есть чёрные полосы
ТО это зебра

Рис. 4.27. Продукционные правила, моделирующие связи между характерными свойствами животных и их принадлежностью к виду

На рис. 4.28 изображена сеть вывода, построенная с использованием продукционных правил, приведенных на рис. 4.27. В сети вывода явно представлены связи 2-го типа, а сама сеть вывода может рассматриваться как *модель всех возможных выводов* дедуктивного типа для базы знаний, включающей продукционные правила, приведенные на рис. 4.27. На рис. 4.28 светлые прямоугольники соответствуют исходным фактам-по-

ссылкам, а тёмные прямоугольники — фактам-заключениям, полученным в процессе вывода.

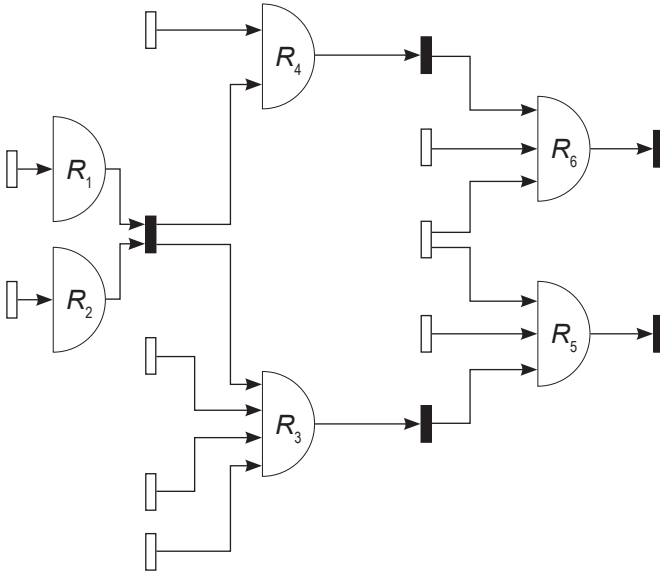


Рис. 4.28. Пример сети вывода для базы знаний, включающей правила, приведенные на рис. 4.27

Возможность явного представления связей 2-го типа, например при помощи сети вывода, позволяет утверждать, что для каждой базы знаний, состоящей из конечного количества фактов и продукционных правил, существует некоторое фиксированное число M , равное *максимальному количеству всех возможных выводов для этой базы знаний*. С этой точки зрения, «новые» факты, получаемые в ходе вывода, на самом деле должны находиться в базе знаний, а процесс вывода лишь осуществляет доступ к этим фактам.

4.5.2. Дедуктивные связи в сети диалогового метода

В экспертной системе база знаний, состоящая из фактов и правил продукции, формирование явных связей между фактами (связей 2-го типа) в процессе вывода может происходить либо при прямом, либо при обратном выводе. Как при прямом, так и при обратном выводе в рабочей памяти экспертной системы накапливаются факты. Отличие прямого вывода от обратного вывода заключается в интерпретации фактов, накопленных в рабочей памяти.

После завершения процесса вывода набор фактов, накопленных в рабочей памяти, можно разделить на два класса: промежуточные факты и терминальные факты. Пусть *FactP* обозначает факт-посылку продукционного правила, а *FactC* — факт-заключение продукционного правила. Тогда в процессе прямого вывода факты, которые накапливаются в рабочей памяти экспертной системы, интерпретируются как факты-посылки (*FactP*), а конечной целью прямого вывода является получение терминального факта-заключения (*FactC*). В процессе обратного вывода факты, которые накапливаются в рабочей памяти экспертной системы, интерпретируются как факты-заключения (*FactC*), а конечной целью обратного вывода является получение терминального списка фактов-посылок (*FactP*). Сеть вывода, приведенная на рис. 4.28, иллюстрирует приведенное понимание процессов прямого и обратного выводов.

Любой промежуточный факт является одновременно и фактом-посылкой (*FactP*) и фактом-заключением (*FactC*): фактом-заключением по отношению к предыдущему правилу и фактом-посылкой по отношению к последующему правилу. Промежуточные факты располагаются «внутри» сети вывода. Их количество может быть достаточно большим, однако они необходимы только для того, чтобы в конце концов получить терминальные факты.

Терминальные факты — это совокупность целевых фактов-посылок или фактов-заключений. Они располагаются

«по краям» сети вывода и определяют искомое решение проблемы. Ясно, что количество промежуточных фактов детерминирует длительность процесса вывода, поэтому желательно, чтобы в каждом конкретном выводе их количество было наименьшим.

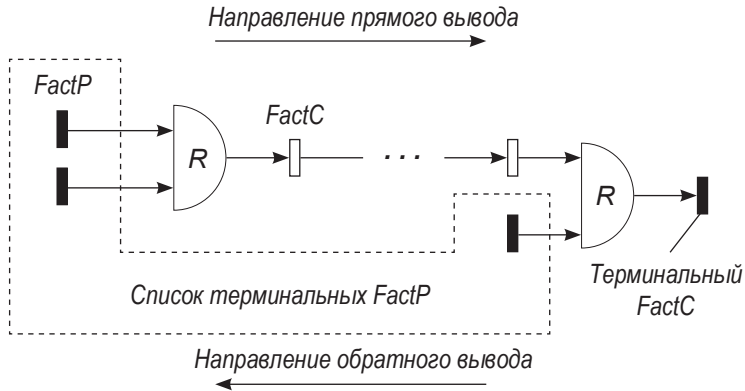


Рис. 4.28. Сеть вывода, иллюстрирующая понятия промежуточного и терминального фактов

После завершения процесса вывода и формирования сети вывода каждый конкретный вывод может быть представлен без промежуточных фактов в виде *отображения терминального списка фактов-посылок в терминальный факт-заключение* (для случая прямого вывода), или в виде *отображения терминального факта-заключения в терминальный список фактов-посылок* (для случая обратного вывода).

$$\{FactP_i\}, i = 1, \dots, n \rightarrow FactC, \text{ прямой вывод,} \quad (4.3)$$

$$FactC \rightarrow \{FactP_i\}, i = 1, \dots, n, \text{ обратный вывод,} \quad (4.4)$$

где:

$\{FactP_i\}$ — терминальный список фактов-посылок,
 $FactC$ — терминальный факт-заключение.

Таким образом, если количество фактов и правил в некоторой базе знаний остаётся неизменным и при помощи этой базы знаний был однажды проведен прямой вывод, то сеть этого вывода может быть «свёрнута» и представлена одним эквивалентным правилом продукции типа (4.5), которое *включает только терминальные факты*:

$$\begin{array}{l} \text{ЕСЛИ } FactP_1 \wedge FactP_2 \wedge \dots \wedge FactP_n \\ \text{ТО } FactC \end{array} \quad (4.5)$$

Если правило типа (4.5) помещено в базу знаний, то в случае, когда возникнет необходимость повторно провести точно такой же вывод, машина вывода экспертной системы сможет воспользоваться правилом типа (4.5) вместо повторного построения сети вывода.

Правило типа (4.5) позволяет провести прямой вывод за n шагов (где n — количество фактов-посылок в левой части правила (4.5)) в том случае, когда имеется метод, позволяющий на каждом шаге получить точно одну посылку, необходимую для левой части правила (4.5). Далее покажем каким образом вопросно-ответный диалоговый процесс может быть использован в качестве отмеченного выше метода.

Ранее (подраздел 4.5.1) мы пришли к выводу, что количество возможных выводов для базы знаний, состоящей из фиксированного количества фактов и правил, ограничено сверху и для такой базы знаний существует не более M возможных выводов. Следовательно, все возможные выводы для фиксированной базы знаний можно моделировать M правилами типа (4.5).

Конъюнкцию в левой части правила (4.5) можно представить цепью из n последовательно соединённых элементов, каждый из которых соответствует факту-посылке $FactP$ (по аналогии с известным формальным представлением операции «И» в переключательных схемах) [131]). Для этой цели используем возможности модели $DiAM$ в виде сети Петри (подраздел 3.3).

На рис. 4.29 правило типа (4.5) представлено в графической форме и изображено в виде фрагмента сети *DiAM*. На рис. 4.29 направление стрелок совпадает с направлением прямого вывода.

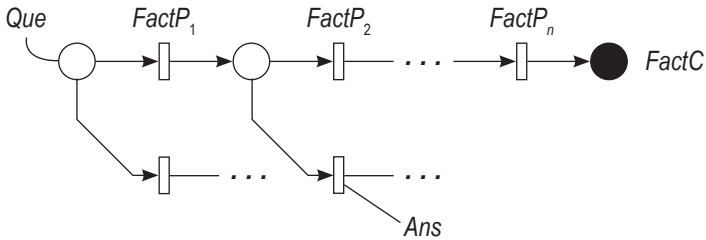


Рис. 4.29. Фрагмент сети *DiAM*, соответствующий правилу типа (4.5)

Поставим в соответствие фактам, которыми оперирует продукционная экспертная система, ответы на вопросы, которыми оперирует *DiAM* в Петри-модели вопросно-ответного диалога. Будем считать, что факту, которым оперирует экспертная система, соответствует один из ожидаемых ответов на вопрос. Поэтому на рис. 4.29 факты представлены переходами. Вопрос с точки зрения логического вывода играет второстепенную роль. Сцена, включающая вопрос, предоставляет реактивному агенту множество ожидаемых ответов, которые можно трактовать как множество родственных и альтернативных фактов. Отвечая на вопрос, реактивный агент выбирает один из альтернативных фактов, формируя, таким образом, очередной конъюнкт в левой части правила (4.5).

Очевидно, что если фрагмент сети *DiAM*, приведенный на рис. 4.29, эквивалентен одному выводу, то вся сеть *DiAM* эквивалентна всем M выводам, возможным для данной базы знаний. Движение фишки по сети *DiAM* моделирует динамику пошагового процесса вывода, а сам вывод завершается за n шагов диалога. На каждом шаге диалога в левую часть правила (4.5) добавляется точно один факт. На рис. 4.29 изображен

случай прямого вывода. При прямом выводе позиция, соответствующая факту-заключению *FactC*, завершает цепь вывода. Зачернённая позиция соответствует факту-заключению. Формирование заключения может завершить диалоговый процесс, если его целью является только диагностика. Однако это не обязательно. Заключение может использоваться для определения дальнейшего «направления» адаптивного диалогового процесса.

Одно из ключевых предположений данного подраздела, заключается в том, что после завершения процесса вывода и формирования сети вывода весь процесс вывода может быть «свёрнут» и заменён одним правилом продукции типа (4.5), которое включает только терминальные факты. Описанная процедура имеет психологические корни в *теории автоматических и осознанных ментальных процессов* [34,132].

Интроспекция и экспериментальные данные показывают, что многократное повторение и практика позволяют резко улучшить временные и качественные показатели, определяющие некоторую поведенческую активность человека. Объяснение этого феномена заключается в том, что многократное решение одной и той же проблемы переводит процесс её решения из разряда процессов, контролируемых сознанием, в автоматически выполняемые процессы. Автоматические процессы характеризуются следующими свойствами [132]:

- автоматические процессы выполняются быстрее, чем процессы, контролируемые сознанием;
- автоматические процессы не уменьшают ресурсы, необходимые для решения других проблем (поскольку не требуют привлечения механизма внимания);
- автоматические процессы не подвластны сознанию;
- автоматические процессы неотвратимы (автоматические процессы безусловно выполняются, если на организм воздействует соответствующий стимул, даже если на этом стимуле не сфокусировано внимание).

Согласно теории, предложенной Шнейдером и Шифриным, различие между автоматическими процессами и процесса-

ми, контролируемые сознанием, заключаются в следующем [133,134]:

- процессы, контролируемые сознанием, используют фиксированный и ограниченный ресурс сознания, базируются на механизме внимания и легко адаптируются к новым проблемам;
- автоматические процессы не зависят от имеющихся ресурсов сознания, не требуют привлечения механизма внимания, но их трудно модифицировать после того, как они были сформированы и запомнены.

Практика и многократное повторение *трансформируют процесс, контролируемый сознанием в автоматический процесс*, и такая трансформация сопровождается реструктуризацией ментальной схемы, соответствующей данному процессу (здесь термин «схема» используется в психологическом смысле этого слова) [135].

На практике, однако, не все процессы могут быть чётко разделены на два класса в соответствии с описанными выше признаками. Существует множество примеров процессов, являющихся автоматическими только частично. Согласно теории, предложенной Норман и Шалис, дихотомическая классификация ментальных процессов упрощает реальность, поскольку не учитывает процессы, занимающие промежуточное положение [136]. Вместо простого дихотомического деления процессов на автоматические и управляемые сознанием, Норман и Шалис предложили три уровня автоматизма ментальных процессов:

- полностью автоматический процесс, управляемый схемой;
- частично автоматический процесс, вовлекающий процедуру разрешения конфликтов нескольких схем, которая осуществляется без преднамеренного вмешательства сознания;
- процесс, находящийся под полным контролем механизма внимания.

Согласно Норман и Шалис, полностью автоматический процесс детерминируется одной схемой, хранящейся в долговре-

менной памяти, и реализуется действительно автоматически с очень малым его осознанием индивидуумом.

В ряде случаев существуют несколько схем, которые могут использоваться для управления процессом. Тогда необходим предварительный этап выбора одной из схем. Норман и Шалис ввели в свою модель специальную контролируемую систему, используемую для осуществления отмеченного выбора. Эта система осуществляет автоматическое разрешение конфликта схем и имеет наименование система конкурирующего планирования (contention scheduling). Частично автоматические процессы осознаются индивидом в большей степени, чем полностью автоматические.

Наконец, существует система, являющаяся, по сути механизмом внимания (supervisory attentional system). Эта система привлекается во всех случаях, когда решение проблемы или принятие решения осуществляется впервые. Она отличается наибольшей гибкостью и приспособляемостью к условиям решаемой задачи.

Таким образом, в отличие от модели Шнейдера и Шифрина модель Норман и Шалис предполагает наличие двух систем управления ментальными процессами: система конкурирующего планирования и механизм внимания.

Логан предложил теорию, объясняющую, каким образом многократное повторение ментальных процессов переводит их в категорию автоматических [137]. Теория Логана базируется на следующих принципах:

- Начальная версия схемы в виде отдельных «траекторий» формируется в долговременной памяти при первоначальном решении проблемы.
- Повторное решение той же проблемы ведёт к «разработке» схемы и увеличению информации как о самой проблеме, так и о методе её решения.
- Схема с увеличенным объёмом хранимой информации приобретает свойство быть быстро найденной в случае необходимости (повторного решения проблемы).
- «Автоматизм — это поиск в долговременной памяти: обработка является полностью автоматической, когда она

базируется на одношаговом и прямом доступе к прошлому решению в памяти» [137].

- При однократном решении проблемы она воспринимается как новая, и её решение формируется с привлечения механизма внимания. После того как проблема решалась многократно, решение сохраняется в долговременной памяти и может быть получено очень быстро.

В унифицированной когнитивной модели SOAR (State Operator and Result), предложенной Ньюэллом, используется механизм «сворачивания» знаний, полученных в результате решения проблемы. Когнитивная модель SOAR представляет собой систему, моделирующую процесс решения проблемы человеком [99]. SOAR моделирует механизмы центральной переработки и не учитывает перцепцию и моторное поведение. Архитектура SOAR включает долговременную и рабочую памяти. Знания, хранимые в долговременной памяти, моделируются набором продукционных правил. Поиск решения проблемы предполагает логический вывод. После формирования решения из фактов, накопленных в рабочей памяти, формируется продукционное правило, которое помещается в долговременную память.

Процессы решения проблем путём поиска в пространстве состояний, которые используются в интеллектуальных агентах, решающих проблемы, а также процессы логического вывода и принятия решений имеют много общих черт и могут моделироваться аналогичными системами или методами. Системы и методы, осуществляющие автоматическое решение отмеченных проблем, могут также рассматриваться как модели соответствующих ментальных функций. Поэтому такие системы должны «по-разному» использовать свои средства при первоначальном и повторном решении одной и той же проблемы. Внешние отличия могут заключаться в том, что первоначально решаемая проблема использует существенное количество ресурса системы (предположительно больший объём адресного пространства основной памяти), а её решение занимать существенное время. Повторно решаемая проблема решается быстрее и использует незначительный ресурс системы.

Если с описанной точки зрения проанализировать «сворачивание» сети вывода в правило типа (4.5), то такой подход можно рассматривать как отражение в компьютерной системе феномена дихотомии автоматических и осознанных ментальных процессов.

Проблема, решение которой осуществляется логическим выводом и решаемая впервые, порождает построение сети вывода. Это ресурсоёмкий и длительный процесс (аналогичный процессу, контролируемому сознанием), поскольку все используемые факты, как терминальные, так и промежуточные, должны быть использованы и «встроены» в сеть вывода. Ясно, что нерационально повторять этот процесс каждый раз, когда необходимо *повторно найти решение той же проблемы*. Если после каждого решения проблемы система генерирует правило типа (4.5) и запоминает его в *DiAM* во внешней памяти, то повторное решение может быть получено за минимальное время (n шагов диалогового процесса). Процесс нахождения решения при помощи правила типа (4.5) является аналогом автоматического процесса. Продолжая аналогию, можно рассматривать *DiAM* как модель долговременной памяти, в которой накоплен «прошлый опыт» решения задач.

4.5.3. Логический вывод в сети диалогового метода

С целью получения формализованного описания дедуктивных связей в сети диалогового метода рассмотрим пример трансформации базы знаний продукционной экспертной системы в сеть *DiAM*. Приведенный ниже пример описан в [126]. Ещё несколько примеров такой трансформации приведено в [138].

Пусть в базе знаний экспертной системы имеется набор фактов и продукционных правил, которые связывают эти факты как факты-посылки и факты-заключения. Набор фактов представим в виде следующих пропозиций:

$$\begin{array}{l}
 p_1 : \text{Намерение-отдых} \\
 p_2 : \text{Дорога ухабистая} \\
 p_3 : \text{Использовать внедорожник} \\
 p_4 : \text{Место отдыха-горы} \\
 p_5 : \text{Место отдыха-пляж} \\
 p_6 : \text{Нужна скорость}
 \end{array}
 \quad \left. \vphantom{\begin{array}{l} p_1 \\ p_2 \\ p_3 \\ p_4 \\ p_5 \\ p_6 \end{array}} \right\} \quad (4.6)$$

Пусть правила, связывающие факты (4.6), имеют вид

$$\begin{array}{l}
 r_1 : \text{ЕСЛИ } p_1 \wedge p_2 \text{ ТО } p_3 \\
 r_2 : \text{ЕСЛИ } p_4 \text{ ТО } p_2 \\
 r_3 : \text{ЕСЛИ } p_1 \text{ ТО } p_6 \\
 r_4 : \text{ЕСЛИ } p_5 \text{ ТО } p_2
 \end{array}
 \quad \left. \vphantom{\begin{array}{l} r_1 \\ r_2 \\ r_3 \\ r_4 \end{array}} \right\} \quad (4.7)$$

Отметим, что правила (4.7) можно рассматривать как некоторые правила вывода, связывающие факты-посылки с фактами-заключениями при помощи связей 1-го типа (подраздел 4.5.1). Например, правило r_1 позволяет вывести факт-заключение p_3 («Использовать внедорожник») при наличии фактов-посылок p_1 и p_2 («Намерение — отдых» и «Дорога ухабистая»). Правила (4.7) можно использовать для построения сети вывода, явно представляющей все возможные связи 2-го типа. На рис. 4.30 изображена сеть вывода, соответствующая описанной базе знаний.

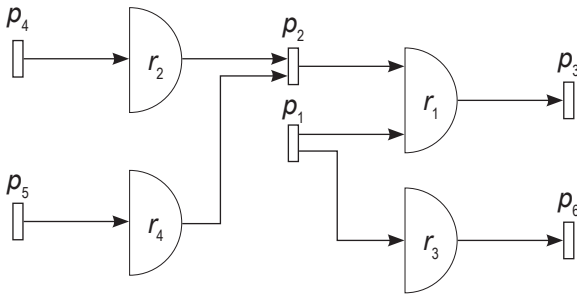


Рис. 4.30. Сеть вывода для базы знаний, состоящей из фактов (4.6) и правил (4.7)

Правило r_3 введено в базу знаний для иллюстрации потенциальной конфликтности набора правил, сформированных случайным образом. Ясно, что при наличии в рабочей памяти экспертной системы одновременно фактов p_1 и p_2 , возможны как вывод факта p_3 (правило r_1), так и вывод факта p_6 (правило p_3). Для разрешения конфликтных ситуаций, возникающих в процессе вывода в продукционных экспертных системах, используются специальные методы, в основе которых лежат знания внешнего (по отношению к базе знаний) интеллекта: стоимость перехода, коэффициент доверия и т.п. Оценки вводятся извне и определяются либо расчётами, либо внешними экспертами. Краткий обзор наиболее известных методов разрешения конфликтных ситуаций приведен в [139]. Отметим, что процесс снабжения правил продукции коэффициентами доверия можно рассматривать как неявный (скрытый) вопросно-ответный диалог с экспертом или как опрос эксперта, проведенный заранее при формировании базы знаний.

Факты-посылки, представленные базой знаний (4.6), (4.7), будем интерпретировать как *ответы на вопросы*. Можно сконструировать такие сцены для сценария диалога, в которых факт-посылка из базы знаний (4.6), (4.7) является одной из ожидаемых реакций реактивного агента. Поскольку факты в базе знаний (4.6), (4.7) представлены вербально, то в качестве интеррогативного стимула в такой сцене может фигурировать ли-вопрос, субъект которого представляет собой список возможных ответов реактивного агента, а требование предполагает альтернативность элементов списка (подраздел 1.2.1). Ясно, что *субъект ли-вопроса должен включать один из фактов-посылок*.

На рис. 4.31 приведена таблица, состоящая из трёх колонок. В первой колонке таблицы перечислены факты-посылки из базы знаний (4.6), (4.7). Во второй колонке — ли-вопросы, субъекты которых являются наборами альтернативных ожидаемых ответов и содержат соответствующие факты-посылки. В третьей колонке таблицы изображены элементы сети *DiAM*, соответствующие ли-вопросу и набору ожидаемых альтернативных ответов.

В таблице на рис. 4.31 предполагается, что факты-посылки p_4 и p_5 могут быть получены как альтернативные ответы на

один и тот же вопрос. Несмотря на то, что в базе знаний эти факты представлены как отдельные пропозиции, с точки зрения логики вопросно-ответных отношений они представляют собой элементы списка субъекта одного и того же вопроса.

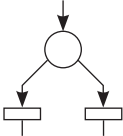
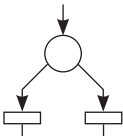
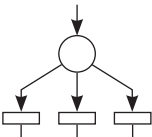
Факты-посылки базы знаний (4.6), (4.7)	Ли-вопросы, ответы на которые порождают факты-посылки	Графическое изображение элементов <i>DiAM</i>
p_1 Намерение — отдых	Что вы намерены делать: отдыхать или работать?	
p_2 Дорога ухабистая	Какая ожидается дорога: ухабистая или ровная?	
p_4 Место отдыха — горы p_5 Место отдыха — пляж	Где вы собираетесь отдыхать: в горах, на пляже или где-нибудь ещё?	

Рис. 4.31. Пример трансформации фактов-посылок продукционной экспертной системы в элементы сети *DiAM*

Таблица на рис. 4.31 позволяет построить сеть *DiAM*, которая моделирует все выводы, возможные для базы знаний (4.6), (4.7). Такая сеть приведена на рис. 4.32. Она соответствует прямому выводу, поскольку вывод завершается достижением позиции, соответствующей факту-заклчению *CFact*. На рис. 4.32 позиция, соответствующая факту-заклчению, зачернена.

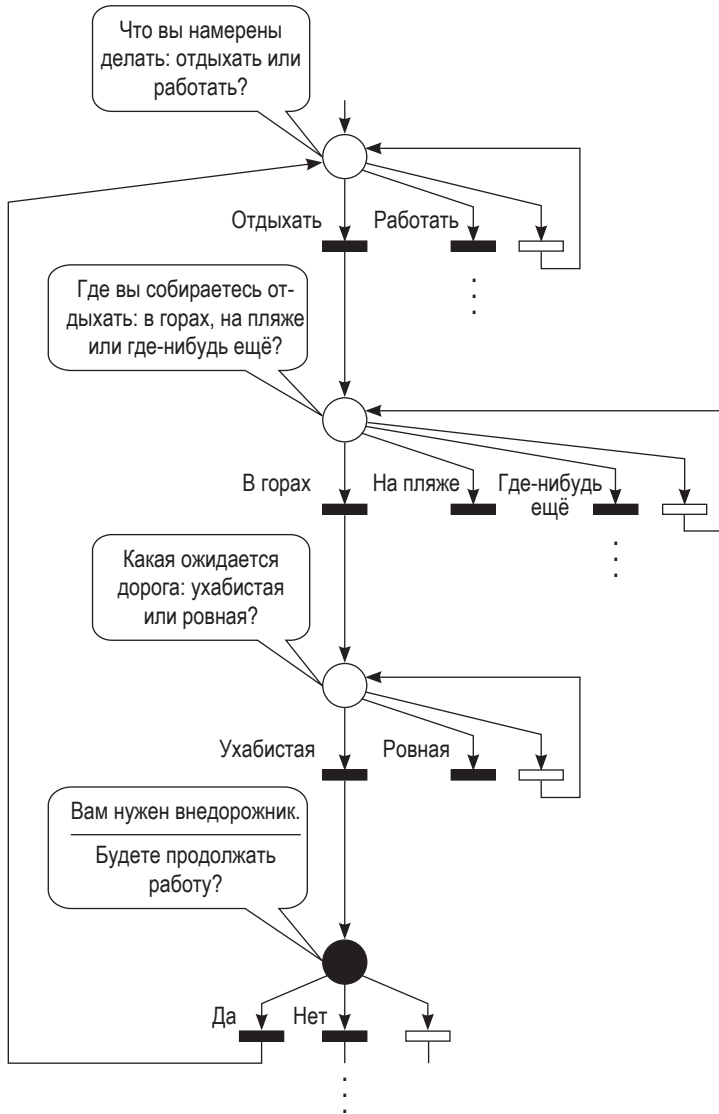


Рис. 4.32. Сеть *DiAM*, которая моделирует все выводы, возможные для базы знаний, (4.6), (4.7)

Для целевой позиции нарушается смысловая интерпретация позиций и переходов, принятая при моделировании прямого вывода при помощи сети Петри (напомним, что согласно договоренности, переход соответствует ответу, который интерпретируется как факт-посылка, а позиция — вопросу). Целевая позиция в случае прямого вывода может содержать либо факт-заключение, либо факт-заключение и очередной вопрос. При изображении сети на рис. 4.32 предполагалось, что после получения заключения и положительного ответа на вопрос: «Будете продолжать работу?» система возвращается в начальное состояние.

Сопоставляя способы представления знаний (факты и связи между ними) при помощи сети *DiAM* и в форме, принятой для продукционных экспертных систем, можно сделать следующие выводы и обобщения.

Сеть *DiAM* имеет один вход, соответствующий началу диалогового процесса, а сеть вывода имеет несколько входов. Наличие множества входов в сети вывода затрудняет или даже исключает совмещение диагностики и обучения в рамках классической продукционной экспертной системы.

В случае сети *DiAM* процесс получения заключения (процесс вывода) сводится к процессу навигации на сети *DiAM*, который управляется реактивным агентом. Таким образом, один из ключевых компонентов архитектуры экспертной системы — машина вывода трансформируется в простую «машину навигации», которая реализуется процессом стандартного диалогового цикла *DiMC* (подраздел 3.4.3).

Привлекательным свойством сети *DiAM* является принципиальное отсутствие конфликтных ситуаций в процессе вывода-навигации. Это следствие естественного ограничения, существующего в вопросно-ответном диалоговом процессе и отраженного в сети *DiAM*. Ограничение заключается в том, что в вопросно-ответном диалоге на любой ответ реактивного агента активный агент генерирует только один вопрос.

Ещё одно заметное свойство сети *DiAM* — возможность объединения нескольких фактов-посылок в одном шаге диало-

гового процесса в том случае, если они являются (или могут быть) элементами субъекта одного и того же вопроса. В рассмотренном примере два исходно различных факта: «Место отдыха — горы» и «Место отдыха — пляж» включены в субъект одного вопроса: «Где вы собираетесь отдыхать: в горах, на пляже или где-нибудь ещё?»

В процессе работы машины вывода экспертной системы количество данных, хранящихся в рабочей памяти, возрастает по мере развития процесса вывода и в связи с ограниченностью компьютерных ресурсов, должно быть регламентировано. В диалоговой системе, осуществляющей логический вывод и построенной на основе *DiAM*, необходимо на каждом шаге диалога хранить данные, соответствующие только текущему шагу вне зависимости от длины дедуктивной цепи.

Есть ещё одно соображение общего характера, на которое следует обратить внимание. Если согласиться с тем, что экспертная система в процессе вывода не генерирует факт-заключение, а отыскивает его в имеющейся базе знаний, то тогда машину вывода можно мыслить как некий метод доступа к данным, хранящимся в базе данных. В этом смысле *DiAM* обеспечивает более эффективный доступ. Эффективность проявляется: в высокой скорости (поскольку нет необходимости строить дерево вывода и разрешать конфликтные ситуации) и низком потреблении ресурса памяти (поскольку вне зависимости от длины дедуктивной цепи необходимо хранить информацию только о текущем шаге).

4.5.4. Вопросно-ответная транзакция в контексте логического вывода

Логическая связь между вопросом активного агента и ответом реактивного агента рассматривалась ранее в общем случае в подразделе 1.2 (Логическая структура вопросно-ответной транзакции). В подразделе 2.1 (Вопросно-ответные транзак-

ции в контексте представления знаний) предложены формулы для представления декларативных знаний, ассоциированных с субъектом вопроса и соответствующим ответом. Модифицируем предложенную теорию, рассматривая вопросно-ответную транзакцию с точки зрения логического вывода.

Для получения очередного факта, необходимого для логического вывода, активный агент предъявляет реактивному агенту вопрос, имеющий логическую организацию, моделируемую интеррогативной формулой

$$Que = ? \rho \sigma, \quad (4.8)$$

где:

ρ — требование вопроса;

σ — субъект вопроса;

? — признак интеррогатива.

С субъектом вопроса ассоциированы знания, которые можно представить несколькими различными способами (формулы (2.10)–(2.15)). Ограничимся случаями, когда знания, ассоциированные с субъектом, представляются двумя способами: (1) в виде объекта-вещи и списка объектов-свойств; (2) в виде объекта-свойства и списка объектов-вещей. Таким образом, субъект вопроса будем представлять одной из следующих двух формул:

$$\sigma = x, \{P_\alpha(x)\}; \alpha = 1, \dots, m, \quad (4.9)$$

$$\sigma = P(x), \{x_\alpha\}; \alpha = 1, \dots, m, \quad (4.10)$$

где:

x — объект-вещь;

$P_\alpha(x)$ — предикат « x обладает свойством P_α »;

$P(x)$ — предикат « x обладает свойством P »;

x_α — значения переменной x .

Субъект ли-вопроса хранит родственные и альтернативные факты. Для примера, рассмотренного в предыдущем подразделе, это может быть набор фактов

Место отдыха-горы
 Место отдыха-пляж
 Место отдыха-лес

Приведенный набор фактов можно представить субъектом ли-вопроса при помощи формулы (4.10) следующим образом:

$$\left. \begin{aligned} \sigma &= P(x), \{x_\alpha\}, \alpha = 1, \dots, 3 = \{Fact_1, Fact_2, Fact_3\}, \\ Fact_1 &= P(x)|x = x_1, \\ Fact_2 &= P(x)|x = x_2, \\ Fact_3 &= P(x)|x = x_3, \end{aligned} \right\} \quad (4.11)$$

где:

$P(x)$ — предикат «Вещь x обладает свойством быть местом отдыха»;

x_1 — горы;

x_2 — пляж;

x_3 — лес.

Поскольку для формирования факта-заключения при помощи продукционного правила типа (4.5), на каждом шаге диалога необходимо получить точно один конъюнкт для левой части правила, то на каждом шаге из субъекта вопроса σ необходимо получить точно один ответ-факт. Таким образом, ответ должен представлять собой *один объект-вещь*, либо *один объект-свойство* и может быть представлен одной из следующих формул:

$$Ans / Fact = x, \{P(x)\}, \quad (4.12)$$

$$Ans / Fact = P(x), \{x\}. \quad (4.13)$$

Формулы (4.12) и (4.13) интерпретируются следующим образом:

«Ответ/факт есть вещь x , обладающая свойством P »

«Ответ/факт есть свойство P , присущее вещи x »

Требование вопроса ρ необходимо для детерминирования ответа с точки зрения мощности множества элементов субъекта, переходящих в ответ. Но, поскольку в случае логического вывода ответ-факт представляет собой единичный объект, смысл требования заключается в том, чтобы сообщить реактивному агенту, что субъект вопроса содержит альтернативный список объектов.

Резюмируя, отметим, что однократная транзакция вопросно-ответного диалога может быть использована для получения очередного факта-посылки при формировании логического заключения. При этом вопрос и релевантный ему факт-ответ могут быть представлены либо формулами (4.14), либо формулами (4.15):

$$\left. \begin{aligned} Que &= \rho, x, \{P_\alpha(x)\}; \alpha = 1, \dots, m, \\ Ans / Fact &= x, \{P(x)\}, \end{aligned} \right\} \quad (4.14)$$

$$\left. \begin{aligned} Que &= \rho, P(x), \{x_\alpha\}; \alpha = 1, \dots, m, \\ Ans / Fact &= P(x), \{x\}, \end{aligned} \right\} \quad (4.15)$$

где:

ρ — требование, детерминирующее альтернативность элементов множеств $\{P_\alpha(x)\}$ и $\{x_\alpha\}$;

$P(x)$ — предикат, интерпретируемый как: « x обладает свойством P »;

m — мощность субъекта вопроса.

Однократной вопросно-ответной транзакции соответствует один узел сети *DiAM*. Цепь, состоящая из последовательно соединённых узлов, моделирует вывод в смысле правила (4.5), а вся сеть *DiAM* может рассматриваться как модель *диалоговой базы знаний*, которая хранит *процедурные знания*, перешедшие в разряд автоматических в виде сети «всех возможных выводов».

Рассмотрим формальное описание диалоговой базы знаний *DiAM*, используя Петри-модель вопросно-ответного диалога, описанную в подразделе 3.3.

Пусть *FACT* — комплект фактов-ответов или их идентификаторов, покрывающих некоторую предметную область, которые необходимы и достаточны для формирования диалоговой базы знаний.

Пусть *QUE* — комплект вопросов или их идентификаторов в виде (4.14) или (4.15).

Пусть совокупность переходов, инцидентных *i*-й позиции сети выводов *DiAM* и соединённых с позицией исходящими ветвями, соответствует множеству фактов

$$F^i \cup NF^i,$$

где:

F^i — множество ожидаемых и распознаваемых на *i*-м шаге вывода альтернативных фактов-ответов, принадлежащих субъекту вопроса для данного шага;

NF^i — множество нераспознаваемых на *i*-м шаге вывода фактов-ответов.

Тогда *i*-й шаг вывода на сети выводов *DiAM* моделируется парой

$$Que_i, (F^i \cup NF^i). \quad (4.16)$$

При графической интерпретации сети выводов *DiAM* вопрос Que_i , как и прежде, соответствует позиции, а $(F^i \cup NF^i)$ — совокупности инцидентных ей переходов. Условием запуска перехода в сети выводов *DiAM* является совпадение ответа, полученного от реактивного агента с одним из фактов-ответов, входящих в (4.16).

Петри-модель *DiAM*, которая рассматривается как диалоговая база знаний, хранящая процедурные знания, перешедшие в разряд автоматических в виде сети «всех возможных выводов», зададим четвёркой

$$DiAM = (QUE, FACT, NextQue, NextFact). \quad (4.17)$$

NextFact — функция последующих и ожидаемых фактов-ответов, задающая отображение комплекта вопросов (или их идентификаторов) в комплект ожидаемых фактов.

$$NextFact : QUE \rightarrow FACT. \quad (4.18)$$

NextQue — функция последующих вопросов (или их идентификаторов), задающая отображение комплекта ожидаемых фактов в комплект вопросов.

$$NextQue : FACT \rightarrow QUE. \quad (4.19)$$

Специфика вопросно-ответного диалога, заключающаяся в том, что активный агент за один раз генерирует и передает реактивному агенту только один вопрос, накладывает на функцию *NextQue* следующее ограничение. Для каждого факта-ответа из комплекта ожидаемых фактов-ответов функция последующих вопросов детерминирует единственный вопрос из комплекта вопросов.

Таким образом, *i*-я позиция сети выводов *DiAM* может быть описана следующими формулами:

$$NextFact(Que_i) = (F^i \cup NF^i), \quad (4.20)$$

$$NextQue(Fact_i) = Que_i. \quad (4.21)$$

На рис. 4.33 приведено графическое представление одного шага Петри-модели сети вывода *DiAM*, которое иллюстрирует введенные понятия.

Сопоставление дедуктивных связей в экспертной системе, основанной на правилах, и в системе, основанной на диалоговой базе знаний, позволяет предложить концепцию экспертной системы с распределённой архитектурой.

Организация экспертной системы, основанной на правилах, подчиняется некоторому стандарту de facto, в соответствии с

которым архитектура такой системы включает следующие базовые компоненты [140,141]:

- глобальную базу фактов,
- множество правил продукции или базу правил,
- систему управления, или интерпретатор правил.

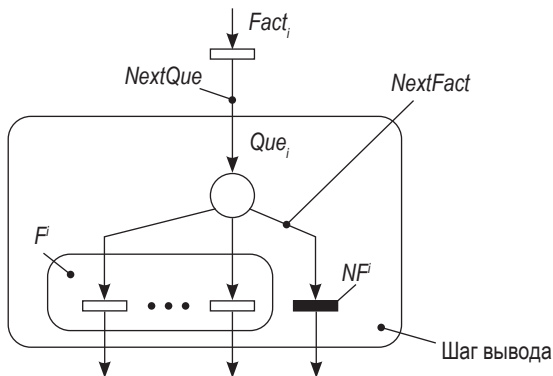


Рис. 4.33. Графическая интерпретация основных понятий Петри-модели сети выводов *DiAM*

Кроме отмеченных базовых компонентов в экспертную систему входит также развитый интерфейс с пользователем. Эволюция архитектуры продукционных экспертных систем направлена на совершенствование каждого из перечисленных компонентов в отдельности. Иными словами, архитектура современных экспертных систем, основанных на правилах, базируется на концепции сосредоточения базовых функций системы в отдельных блоках. Такая «сосредоточенная» архитектура является универсальной и позволяет независимо совершенствовать отдельные компоненты. Однако она порождает ряд проблем, таких, например, как разработка эффективной стратегии интерпретатора правил, учитывающей потенциальную конфликтность набора правил (режим вывода с возвращением и др.).

Под распределённой архитектурой будем понимать такую организацию системы, когда все её функции распределены на множестве однотипных и структурно подобных элементов, каждый из которых содержит фрагмент: глобальной базы фактов, базы правил, системы управления и интерфейса с пользователем. Процесс функционирования распределённой экспертной системы заключается в пошаговой интерпретации отмеченных однотипных элементов в соответствии с корректировкой цели, вводимой в процессе диалога с пользователем. Таким образом, в распределённой экспертной системе диалог с пользователем является существенной частью процесса решения задачи. Основу организации распределённой экспертной системы может составлять диалоговая база знаний, представленная *DiAM* и памятью вопросов.

СПИСОК ЛИТЕРАТУРЫ

1. Платон. Сочинения в четырёх томах. Под общ. ред. А. Ф. Loseva и В. Ф. Asmusa; Пер. с древнегреч. СПб.: Изд-во С.-Петерб. ун-та; Изд-во Олега Абышко, 2006–2007.
2. Сочинения Григория Саввича Сковороды под ред. Проф. Вагалея Д. И. Издание типографии губернского правления. Харьков. 1894.
3. Платон. Сочинения в четырёх томах. Т.1. Под общ. ред. А. Ф. Loseva и В. Ф. Asmusa. Пер. с древнегреч. СПб.: Изд-во С.-Петерб. ун-та; Изд-во Олега Абышко, 2006. 632 с.
4. Schaffer J. Knowing the answer. *Philosophy and Phenomenological Research*. 2007. Vol. LXXV No. 2, September 2007. P. 383–403.
5. Kahneman D. Attention and Effort. 1973. Prentice-Hall, Inc., Englewood Cliffs, New Jersey.
6. Belnap N. B., Steel T. The Logic of Questions and Answers. New Haven and London: Yale University Press, 1976.
7. Epistemology: an anthology. 2nd edition. Edited by Ernest Sosa, Jaegwon Kim, Jeremy Fantl, and Matthew McGrath. Blackwell Publishing, 2008. 917 p.
8. Wenger E. Artificial Intelligence and Tutoring Systems. Computational and Cognitive Approaches to the Communication of Knowledge. Morgan Kaufmann Publishers Inc., 1987. 486 p.
9. Kubinski T. An Outline of the Logical Theory of Questions. Akademie-Verlag. Berlin, 1980. 143 p.
10. Hintikka J., Knowledge and Belief. An Introduction to the Logic of the Two Notions. King's College Publications reprint edition, 2005. 133 p.

-
11. Wisniewski A. The Posing of Questions. Logical Foundations of Erotetic Inferences. Kluwer Academic Publishers, Dordrecht, 1995. 246 p.
 12. Белнап Н., Стил С. Логика вопросов и ответов. М.: Прогресс, 1981. 110 с.
 13. Uyemov A. The Ternary Description Language as a Formalism for the Parametric General Systems Theory. Part I. *Int. J. of General Systems*. 1999. Vol. 28(4-5). P. 351–366.
 14. Uyemov A. The Ternary Description Language as a Formalism for the Parametric General Systems Theory. Part II. *Int. J. of General Systems*. 1999. Vol. 31(2), P. 131–155.
 15. Uyemov A. The Ternary Description Language as a Formalism for the Parametric General Systems Theory. Part III. *Int. J. of General Systems*. 2003 Vol. 32(6). P. 583–623.
 16. Leonid Leonenko. The Language of Ternary Description and its founder. *Modern Logic*. 2000-2001. Vol. 8, N. 3–4. P. 31–52.
 17. Закусило О. К., Верлань А. Ф., Чмырь И. А. Моделирование диалоговых процессов. Киев: Издательско-полиграфический центр «Киевский университет». 2014. 255 с.
 18. Chimir I., Abu-Dawwas W. Modeling of a Procedural Knowledge by a Dialogue Knowledge Base. *The International Arab Journal of Information Technology*. 2004. Vol. 1, N 1. P. 128–134.
 19. Hintikka J. Questions about questions. In M. K. Munitz and P. K. Unger (Eds.). *Semantics and Philosophy*. New York: University Press, 1974. P.103–158.
 20. Turing A. Computing machinery and intelligence. *Mind*, 59. 1950. P. 433–460.
 21. Рассел С., Норвиг П. Искусственный интеллект. Современный подход. Второе издание. Москва, Санкт-Петербург, Киев: Издательство Вильямс, 2006. 1407 с.
 22. Brainerd C. Piaget's Theory of Intelligence. Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1978. 326 p.
 23. Найссер У. Теория восприятия. В сборнике *Психология внимания*. Под редакцией Ю. Б. Гиппенрейтера и В. Я. Романова. Москва: Издательство «ЧеРо» НОУ Московский психолого-социальный институт, 2001. С. 640-648.
 24. Edmund Gettier. Is Justified True Belief Knowledge? *Analysis*. 1963. Vol. 23. P. 121-123.
 25. DeRose Keith, and Warfield Ted. *Skepticism. A Contemporary Reader*. Oxford: Oxford University Press, 1999.

-
26. Hintikka Jaakko. Different Constructions in Terms of the Basic Epistemological Verbs: A Survey of Some Problems and Proposals: *The Intensions of Intentionality and Other New Models for Modalities* (Dordrecht), 1975.
 27. Lewis David. Whether Report. In Tom Pauli (ed.) *Philosophical Essays Dedicated to Lennart Åqvist on his Fiftieth Birthday* (Uppsala): 1982, P. 194–206.
 28. Boër Steven and William Lycan. *Knowing Who* (Cambridge), 1986.
 29. Higginbotham James. The Semantics of Questions. In Shalom Lappin (ed.). *The Handbook of Contemporary Semantic Theory* (Oxford), 1996, P. 361–383.
 30. Stanley Jason and Timothy Williamson. Knowing How. *Journal of Philosophy*, 2001. 98. P. 411–414.
 31. Anderson J. R. Rules of the Mind. — Hillsdale, N. J.: Lawrence Erlbaum Associates, 1993. 410 p.
 32. Минский М. Фреймы для представления знаний. М: Энергия, 1979. 50 с.
 33. Eysenck M. W., Keane M. T. *Cognitive Psychology. A Student's Handbook*. 3rd edition. Lawrence Erlbaum Associates, 1995. 542 p.
 34. Солсо Р. Л. Когнитивная психология. М.: Триволта, 1996. 598 с.
 35. Kellogg R. T. *Cognitive Psychology*. SAGE publications Ltd., 1995. 523 p.
 36. Tolman E. C. Cognitive maps in rats and men. *Psychological Review*. 1975. Vol. 55. P. 189–208.
 37. Chimir I., Horney M. New Visions of Old Models. In W. M. Beynon, C. L. Nehaniv, & K. Dautenhahn (eds.), *Cognitive Technology: Instruments of Mind* (Proc. 4th International Conference on Cognitive Technology, CT 2001 — University of Warwick, U. K.). Springer Lecture Notes in Artificial Intelligence, 2001. Vol. 2117. P. 157–163.
 38. Чмырь И. А., Раид Муса Аль-Кавасми. Цикл перцепции Нейсера: формальное представление и практическое применение. *Искусственный интеллект*. 2003. №1. С. 107–116.
 39. Верлань А. Ф., Чмырь И. А. Объектно-ориентированное моделирование когнитивных процессов. *Электронное моделирование*. 2002. Т.24, №4. С. 53–64.
 40. Chimir I. A., Abu-Dawwas W. A., Horney M. A. Neisser's Cycle of Perception: Formal Representation and Practical Implementation. *Journal of Computer Science*. 2005 Special issue on Intelligence, Communication and Computer Interaction. P. 106–111.

-
41. Брауэр В. Введение в теорию конечных автоматов. М.: Радио и связь, 1987. 219 с.
 42. Аверкин А., Батыршин И. и др. Нечёткие множества в моделях управления и искусственного интеллекта. М.: Наука, 1986. 219 с.
 43. Jang J.-S. R., Sun Ch.-T., Mizutani E. Neuro-Fuzzy and Soft Computing. N. J.: Prentice Hall, 1997. 603 p.
 44. Чмырь И. А., Чмырь А. И., Абу-Давваз. Нечёткая автоматная модель диалогового поведения. *Тепловые режимы и охлаждение радиоэлектронной аппаратуры (ТРуО)*. 1998. № 1. С. 84–90.
 45. Чмырь И. А., Ус М. Ф., Пискун А. В. Применение нечёткой логики для определения когнитивного типа пользователя в интеллектуальных обучающих системах. Моделювання та інформаційні технології: Збірник наукових праць. Київ: Інститут проблем моделювання в енергетиці НАН України, 2002. Вип. 14. С. 98–104.
 46. Ричмонд У. К. Учителя и машины. Введение в теорию и практику программированного обучения. М.: Мир, 1968. С. 57–64.
 47. Волож Б. Б. ДИМО — гибкие сценарии диалога и их конструирование. *Пакеты прикладных программ. Инструментальные системы*. М.: Наука, 1987. С. 18–29.
 48. Чмырь И. А., Попов П. С. Сценарий диалога и его структурные элементы. *Вопросы радиоэлектроники. Серия ТРТО*. 1991. Вып. 3. С. 87–95.
 49. Чмырь И. А., Клименко М. А. Объектная модель для диалогового решателя задач. *Збірник наукових праць Інституту проблем моделювання в енергетиці*. — Київ, 1998. Вип. 4. С. 152–158.
 50. Цикридис Д., Лоховски Ф. Модели данных. М.: Финансы и статистика, 1985. 254 с.
 51. Питерсон Дж. Теория сетей Петри и моделирования систем. М.: Мир, 1984. 264 с.
 52. Мурата Т. Сети Петри: свойства, анализ, приложения. *ТИИЭР*. 1989. Т.77, №4. С. 41–85.
 53. Котов В. Е. Сети Петри. М.: Наука, Главная редакция физико-математической литературы, 1984. 164 с.
 54. Raven J. Mental tests used in genetic studies: The performance of related individuals on tests mainly educative and mainly reproductive. MSc Thesis, University of London, 1936.
 55. Chmyr Igor. Dialogue of Partners as a Method For Non-Formal Problem Solving. In Maddy D. Brouwer-Janse and Thomas L. Harrington (eds.) collection: *Human-Machine Communication*

-
- for Educational Systems Design*, NATO ASI Series F129 Springer-Verlag, Berlin, 1994. P. 221–228.
56. Мейер Д. Теория реляционных баз данных. М.: Мир, 1987. 331 с.
 57. Burt P. J. Smart sensing within a pyramid vision machine. *Proceedings of the IEEE*. 1988. Vol. 76, iss. 8. P 1006-1015.
 58. Харитоненко И. О. Организация средств визуализации изображений в компьютерных системах поддержки диалога: дисс. канд. техн. наук: спец. 05.13.05. Элементы и устройства вычислительной техники и систем управления. Защищена в Одесском политехническом университете. Одесса, 1993.
 59. Taylor Edwin F. (ed.) ELIZA: a skimmable report on the ELIZA conversational tutoring system. Educational Res. Center, MIT, Cambridge, Mass., Mar. 1968.
 60. Шенк Р. Обработка концептуальной информации. М.: Энергия, 1980.
 61. Stevens A. L. and Collins A. The goal structure of a Socratic tutor. *Proceedings of the National ACM Conference*. Seattle. Washington. Association for Computing Machinery. New York, 1977. P. 256-263.
 62. Brown J. S., Burton R. R. and Bell A. G. SOPHIE: a step towards a reactive learning environment. *Int. J. Man-Machine Studies*. 1975. Vol. 7. P. 675–696.
 63. Harmon Paul, King Davis. *Expert Systems. Artificial Intelligence in Business*. John Willey, 1985.
 64. Newell A. Remarks on the relationship between artificial intelligence and cognitive psychology. In R. Banerji and J. D. Mesarovich (eds.) *Theoretical Approaches to Non-Numerical Problem Solving*. New York: Springer-Verlag, 1970. P. 363–399.
 65. Anderson J. R., Corbet A. T., Koedinger K. R., Pelletier R. Cognitive Tutors: Lessons Learned. *The Journal of the Learning Sciences*. — 1995. Vol.4, N. 2. P. 167–207.
 66. Carbonell J. R. *Mixed-Initiative Man-Computer Instructional Dialogues: Doctoral dissertation*. Massachusetts Institute of Technology. Cambridge, Massachusetts, 1970.
 67. Collins A., Warnock E. H., Aiello N., and Miller M. L. Reasoning from incomplete knowledge. In Bobrow D. and Collins A. (eds.) *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press, 1975.
 68. Collins A., Warnock E. H. and Passafiume J. Analysis and synthesis of tutorial dialogues. In Bower G. (ed.). *The Psychology of Learning and Motivation* (vol. IX). New York: Academic Press, 1975.

-
69. Collins A. Processes in Acquiring Knowledge. In Anderson R. C., Spiro R. J. and Montague W. E. (eds.). *Schooling and the Acquisition of Knowledge*. Lawrence Erlbaum Associates. Hillsdale. New Jersey, 1977.
 70. Collins A. and Stevens A. L. Goals and strategies for inquiry teachers. In Glaser R. (ed.) *Advances in Instructional Psychology* (vol. II). Lawrence Erlbaum Associates. Hillsdale. New Jersey, 1982.
 71. Stevens A. L., Collins A. and Goldin S. Misconception in students' understanding. *Int. J. Man-Machine Studies*. 1979. Vol. 11. P. 146-156.
 72. Stevens A. L., Collins A. Multiple models of a complex system: In Snow R., Frederico P. and Montague W. (eds.). *Aptitude, Learning, and Instruction* (vol. II). Lawrence Erlbaum Associates. Hillsdale. New Jersey, 1980.
 73. Brown J. S., Burton R. R. and Bell A. G. SOPHIE: a sophisticated instructional environment for teaching electronic troubleshooting. BBN Report 2790. Bolt Beranek and Newman Inc., Cambridge. Massachusetts, 1974.
 74. Brown J. S. Burton R. R. and Bell A. G. SOPHIE: a step towards a reactive learning environment. *Int. J. Man-Machine Studies*. 1975. Vol. 7. P. 675-696.
 75. Brown J. S., Burton R. R. Multiple representation of knowledge for tutorial reasoning. In Bobrow D. and Collins A. (eds.) *Representation and Understanding: Studies in Cognitive Science*. New York: Academic Press, 1975.
 76. Brown J. S., Rubinstein R. and Burton R. R. Reactive learning environment for computer-assisted electronic instruction. BBN Report 3314. Bolt Beranek and Newman Inc., Cambridge. Massachusetts, 1976.
 77. Brown J. S., Burton R. R. and de Kleer J. Pedagogical, natural language, and knowledge engineering techniques in SOPHIE I, II, and III. In Sleeman D. H. and Brown J. S. (eds.). *Intelligent Tutoring Systems*. Academic Press. London, 1982.
 78. Burton R. R. Semantically-centered parsing. Doctoral dissertation. University of California. Irvine. California, 1975.
 79. Burton R. R., Brown, J. S. An investigation of computer coaching for informal learning activities. *Int. J. of Man-Machine Studies*. 1979. Vol. 11. P. 5-24.
 80. Yob G. Hunt the Wumpus. *Creative Computing*. 1975. September/October. P. 51-54.

-
81. Stansfield J. C., Carr B., Goldstein I. P. Wumpus advisor I: a first implementation of a program that tutors logical and probabilistic reasoning skills. Cambridge, Massachusetts, Massachusetts Institute of Technology, 1976. 98 p. (AI Lab Memo 381.).
 82. Carr B., Goldstein, I. P. Overlays: a theory of modeling for computer-aided instruction. Cambridge, Massachusetts, Massachusetts Institute of Technology, 1977. 115 p. (AI Lab Memo 406)
 83. Goldstein I. P. The genetic graph: a representation for the evolution of procedural knowledge. *Int. J. Man-Machine Studies*. 1979. Vol. 11. P. 51-57.
 84. Anderson J. R. Language, memory, and thought. Hillsdale, N. J.: Lawrence Erlbaum Associates, 1976. 287 p.
 85. Anderson J. R. The Architecture of Cognition. Cambridge, Massachusetts: Harvard University Press, 1983. 311 p.
 86. Boyle C. F., Anderson J. R. Acquisition and automated instruction of geometry skills. Annual Meeting of the American Educational Research Association, New Orleans. Washington, D. C., 1984. 21 p.
 87. Anderson J. R., Boyle C. F., Yost G. The geometry tutor. Proc. of the International Joint Conference on Artificial Intelligence, Los Angeles. 1985. P. 1-9.
 88. Farell R. G., Anderson J. R., Reiser B. J. An interactive computer-based tutor for LISP. Proc. of the National Conference on Artificial Intelligence; Austin, Texas, 1984. P. 106-109.
 89. Reiser B. J., Anderson J. R., Farell R. G. Dynamic student modeling in an intelligent tutor for LISP programming. Proc. of the Ninth International Joint Conference on Artificial Intelligence; Los Angeles, 1985. P. 8-14.
 90. Corbett A. T., Anderson J. R. LISP Intelligent Tutoring System: Research in Skill Acquisition. Computer-Assisted Instruction and Intelligent Tutoring Systems. Shared Goals and Complementary Approaches. In Jill H. Larkin and Ruth W. Chabay (eds.). N. J., Hillsdale, Lawrence Erlbaum Associates, 1992. P. 48-55.
 91. Baker R. S., Corbett A. T., Wagner A. Z., Koedinger K. R. Off-Task Behavior in the Cognitive Tutor Classroom: When Students "Game the System". Proc. of the SIGCHI conference on human factors in computing systems, 2004. Vol.6. N.1 P. 383-390.
 92. Alevan V., McLaren B., Roll I., Koedinger K. Towards Tutoring Help Seeking: Applying Cognitive Modeling to Meta-Cognitive Skills. Proc. of Intelligent Tutoring Systems Conference, 2004. P. 129-136.

-
93. Draper S. W. Constructivism, other theories of the teaching and learning process, and their relationships. Discussion paper for NATO advanced studies institute at Heriot Watt University. Edinburgh, 1994.
 94. Rumelhart D. E. and Norman D. A. Accretion, tuning and restructuring: three modes of learning. In Cotton, J. W. and Klatzky, R. L. (eds.). *Semantic factors in cognition*. Hillsdale, N. J.: Erlbaum, 1978. P. 37–53.
 95. von Glasersfeld E. An introduction to radical constructivism. In Watzlawick P. (ed.). *The invented reality: How do we know what we believe we know? Contributions to constructivism*. W. W. Norton: New York, 1984. P. 17–40.
 96. Bartlett F. C. *Remembering*. Cambridge University Press, 1932.
 97. Putnam H. The meaning of meaning. In *Mind, language and reality*. Cambridge University Press. Cambridge, U. K., 1975.
 98. Scribner S., Cole M. The cognitive consequences of formal and informal education. *Science*. 1973. Vol. 182. P. 553–559.
 99. Newell A. *Unified Theories of Cognition*. Harvard University Press. Cambridge Massachusetts, 1994.
 100. Laurillard D. *Rethinking university teaching: A framework for the effective use of educational technology*. London: Routledge, 1993.
 101. Frisby A. W. *Teaching English: Notes and Comments on Teaching English Overseas*. London: Longman, 1957.
 102. Halliday M. A. K., McIntosh A. and Strevens P. *The Linguistic Science and Language Teaching*. London: Longman, 1964.
 103. Davies P., J. Roberts and R. Rossner. *Situational Lesson Plans*. Mexico City: Macmillan, 1975.
 104. Byrne D. *Teaching Oral English*. London: Longman, 1976.
 105. Brooks N. *Language and Language Learning: Theory and Practice*. 2nd ed. New York: Harcourt Brace, 1964.
 106. Skinner B. F. *Verbal Behavior*. New York: Appleton-Century-Crofts, 1957.
 107. Brown H. D. *Principles of Language Learning and Teaching*. Englewood Cliffs. N. J.: Prentice-Hall, 1980.
 108. Rivers W. M. *The Psychologist and Foreign Language Teacher*. Chicago: University of Chicago Press, 1964.
 109. Fries C. C., Fries A. C. *Foundation for English Teaching*. Tokyo: Kenkyusha, 1961.
 110. Palmer H., D. *English Through Actions*. 1925. Reprint ed. London: Longman Green, 1959.

-
111. Katona G. Organizing and Memorizing: Studies in the Psychology of Learning and Teaching. New York: Columbia University Press, 1940.
 112. Asher J. Learning Another Language Through Action: The Complete Teacher's Guide Book. Los Gatos, Calif.: Sky Oaks Productions, 1977. (2nd ed. 1982)
 113. Мельников В. М., Ямпольский Л. Т. Введение в экспериментальную психологию личности. М.: Просвещение, 1985. 319 с.
 114. Общая психодиагностика. Основы психодиагностики, немедицинской психотерапии и психологического консультирования. Под ред. А. А. Бодалева, В. В. Столина. М.: Изд-во МГУ, 1987. 304 с.
 115. Пирьев Г. Д. Классификация методов в психологии. Братислава: Психодиагностика в социалистических странах. 1985. С. 19–25.
 116. Бурлачук Л. Ф., Морозов С. М. Словарь-справочник по психологической диагностике. — Киев: Наукова думка, 1989. 200 с.
 117. Шмелев А. Г., Похилько В. И. Анализ пунктов при конструировании и применении тест-опросников: ручные и компьютерные алгоритмы. *Вопросы психологии*. 1985. № 4 С. 126–134.
 118. Березин Ф. Б., Мирошников М. П., Рожанец Р. В. Методика многостороннего исследования личности. М.: Медицина, 1976.
 119. Wechsler D. The Measurement and Apraisal of Adult Intelligence. Baltimore: Williams and Wilkins, 1958. 297p.
 120. Румянцев Г. Г. Опыт применения метода незаконченных предложений в психиатрической практике. Исследование личности в клинике и в экспериментальных условиях. Л.: НИИ психоневрологии им. В. М. Бехтерева, 1969. С. 266–275.
 121. Беспалько И. Г., Гильяшева И. Н. Проективные методы. Методы психологической диагностики и коррекции в клинике. Л.: Медицина, 1971. С. 116–144.
 122. Норакидзе В. Г. Методы исследования характера личности. Тбилиси: Мецниереба, 1975. 243 с.
 123. Pittman G. Teaching Structural English. Brisbane: Jacaranda, 1963. 281 p.
 124. Francoise G. Developing Reading Skills: A practical guide to reading comprehension exercises. Cambridge University Press, 1992. 201 p.
 125. Watson B. Mini-Effort, Maxi>Returns. Odessa: The British Council, English Teaching Resource Center, 1996. 56 p.
 126. Представление и использование знаний. Под ред. Х. Уэно и М. Исидзука. М.: Мир, 1985. 153 с.
 127. Кондаков Н. И. Логический словарь-справочник. М.: Наука, 1976. 586 с.

-
128. Winston P. H. Artificial Intelligence. Addison-Wesley, 1984. 359 p.
 129. Попов Э. В. Экспертные системы. М.: Наука, 1987. 211 с.
 130. Levin R. I., Drang D. E., Edelson B. A Comprehensive Guide to AI and Expert Systems: Turbo Pascal Edition. McGraw-Hill, 1988. 182 p.
 131. Шеннон К. Символический анализ релейных и переключательных схем. Работы по теории информации и кибернетике. М.: Изд. ин. лит., 1963. 829 с.
 132. Eysenck M. W., Keane M. T. Cognitive Psychology. A Student's Handbook. 3rd edition. Lawrence Erlbaum Associates, 1995. 542 p.
 133. Schneider W., Shiffrin R. M. Controlled and automatic human information processing: Detection, search, and attention. *Psychological Review*. 1977. N 92. P. 424–428.
 134. Shiffrin R. M., Schneider W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*. 1977. N 84. P. 127–190.
 135. Cheng P. W., Holyok K. J. Pragmatic reasoning schemas. *Cognitive Psychology*. 1985. N 17. P. 391–416.
 136. Norman D. A., Shallice T. Attention to action: Willed and automatic control of behavior. Davidson R. J., Schwartz G. E., Shapiro D. (eds.), *The design of everyday things*. New York: Doubleday, 1988. P. 89–97.
 137. Logan G. D. Toward an instance theory of automatization. *Psychological Review*. 1988. N 95. P. 492–527.
 138. Чмырь И. А., Попов П. С. Экспертные системы на базе сетевой диалоговой модели. *Вопросы радиоэлектроники. Серия ТРТО*. 1992. Вып. 3-4. С. 42–51.
 139. Франклин Дж. Э., Кармоди К. Л., Келлер К., Левит Т. С., Бюто Б. Л. Технология экспертных систем для военных применений: избранные примеры. *ТМЭР*. 1988. Т. 76, №10. С. 18–69.
 140. Осука С. Обработка знаний. М.: Мир, 1989. 149 с.
 141. Нильсон Н. Принципы искусственного интеллекта. М.: Радио и связь, 1985. 252 с.

ТЕЗАУРУС

1. ЕСТЕСТВЕННЫЙ ДИАЛОГ

Диалоговые агенты.

Диалоговые агенты — это участники диалогового процесса, которые формируют диалоговую транзакцию.

Диалоговая транзакция.

Диалоговая транзакция представляет собой элементарный завершённый цикл обмена знаниями между диалоговыми агентами. Хотя участников диалога может быть много, диалоговую транзакцию всегда формируют только два диалоговых агента.

Роль диалогового агента.

Диалоговый агент в процессе диалогового взаимодействия со своим партнёром может исполнять одну из двух альтернативных ролей: роль активного диалогового агента либо роль реактивного диалогового агента.

Активный диалоговый агент.

Активный диалоговый агент — это спрашивающий агент. Та часть диалоговой транзакции, которую формирует ак-

тивный агент, по отношению к реактивному агенту имеет статус вопроса.

Реактивный диалоговый агент.

Реактивный диалоговый агент — это отвечающий агент, и та часть диалоговой транзакции, которую он формирует, имеет статус ответа по отношению к активному агенту.

Внешний диалог.

Диалоговая транзакция при внешнем диалоге формируется двумя личностями. Одна из личностей исполняет роль активного агента, а другая — роль реактивного агента.

Внутренний диалог.

Диалоговая транзакция может быть сформирована одной личностью, попеременно исполняющей роль активного или реактивного агентов. Такой диалог называется внутренним диалогом.

Протокол диалога.

Диалоговый процесс может быть записан в виде протокола диалога, представляющего собой последовательную цепь диалоговых транзакций.

Информационная посылка активного агента.

Информационная посылка активного агента — это часть диалоговой транзакции, которую формирует активный агент.

Степень интеррогативности посылки активного агента.

Информационная посылка активного агента всегда имеет статус вопроса/интеррогатива, но с разной степенью интеррогативности.

Информационная посылка реактивного агента.

Информационная посылка реактивного агента — это часть диалоговой транзакции, которую формирует реактивный агент.

Информационная посылка активного агента типа «поисковое предписание».

Информационная посылка активного агента типа «поисковое предписание» детерминирует область долговременной памяти реактивного агента, в которой находятся запрашиваемые знания. Предполагается, что реактивный агент уже обладает запрашиваемыми знаниями.

Информационная посылка активного агента типа «задача».

Для формирования знаний, запрашиваемых при помощи посылки активного агента типа «задача», реактивный агент должен решить задачу, ассоциированную с этой посылкой. Ответом является вариант решения, полученный реактивным агентом. Предполагается, что, знания, запрашиваемые посылкой типа «задача», не хранятся в памяти реактивного агента в готовом виде, и требуется использование ментального ресурса для их формирования.

Неопределённое поисковое предписание.

Неопределённое поисковое предписание посылки активного агента означает, что ответная посылка реактивного агента может принадлежать одному из нескольких классов ответов.

Определённое поисковое предписание.

Определённое поисковое предписание посылки активного агента означает, что запрашиваются знания, принадлежащие только одному классу ответов.

Степень неопределённости поискового предписания.

Степень неопределённости посылки активного агента типа «поисковое предписание» характеризует количество классов, к которым могут относиться запрашиваемые знания. Чем больше это количество, тем выше степень неопределённости посылки типа «поисковое предписание».

Субъект поискового предписания.

Субъектом посылки активного агента типа «поисковое предписание» называется компонент определённого поискового предписания, детерминирующий класс запрашиваемых знаний.

Ли-посылка.

Посылка активного агента типа «поисковое предписание», субъект которой детерминирует класс ответов с конечным количеством экземпляров, называется ли-посылка.

Какой-посылка.

Посылка активного агента типа «поисковое предписание», субъект которого детерминирует класс ответов с бесконечным количеством экземпляров, называется какой-посылка.

Прямая посылка реактивного агента.

Прямой посылкой реактивного агента называется ответ реактивного агента, при помощи которого реактивный агент непосредственно ссылается на знания, запрашиваемые активным агентом.

Непрямая посылка реактивного агента.

Непрямой посылкой реактивного агента называется ответ реактивного агента, при помощи которого реактивный агент опосредованно ссылается на знания, запрашиваемые активным агентом.

Кодированная прямая посылка реактивного агента.

Кодированная прямая посылка реактивного агента — это прямая посылка реактивного агента, представленная компактным кодом, смысл которого определяется контекстом диалоговой транзакции.

Диалоговое поведение.

Диалоговым поведением называется реализация метода решения проблемы при помощи диалогового процесса.

Метод вопросов Сократа.

Диалоговое поведение, используемое Сократом и описанное в диалогических произведениях Платона, часто называют методом вопросов Сократа. Метод вопросов Сократа предполагает, что метод решения проблемы известен активному агенту диалога.

Wh-вопрос.

Wh-вопрос — это наименование типа вопросов, обобщающего шесть подтипов специальных вопросов английского языка: (1) who/кто-вопрос; (2) what/что-вопрос; (3) when/когда-вопрос; (4) where/где-вопрос; (5) how/как-вопрос; (6) why/почему-вопрос.

Знания-wh.

Знания-wh — это знания, выражаемые ответами на wh-вопросы. Знания-wh формируют шесть подтипов декларативных знаний: (1) who/кто-знания; (2) what/что-знания; (3) when/когда-знания; (4) where/где-знания; (5) how/как-знания; (6) why/почему-знания.

Прямой ответ.

Прямой ответ — это знания, которые необходимы активному агенту для продолжения диалога и которые он хочет получить, формулируя вопрос.

Субъект вопроса.

Субъект вопроса явно или неявно определяет множество утверждений. Утверждения понимаются в смысле ассерторических суждений, в которых констатируется наличие или отсутствие у предмета некоторого признака. В субъекте вопроса могут быть как истинные, так и ложные утверждения.

Требование вопроса.

Требование вопроса представляет собой набор ограничений, накладываемых на возможные ответы. Требование необходимо

для специфицирования того, в каком виде должен быть представлен ответ, формируемый реактивным агентом из субъекта.

Спецификация размера выборки.

Компонент требования вопроса, который специфицирует количество элементов субъекта, переходящих в ответ.

Спецификация требования полноты.

Компонент требования вопроса, который специфицирует долю истинных элементов субъекта, переходящих в ответ, по отношению ко всему множеству истинных элементов субъекта.

Спецификация требования различения.

Компонент требования вопроса, который индицирует, необходимо ли в ответе указывать, являются ли элементы субъекта реально или номинально различными. Реальные элементы субъекта — это некоторые сущности мира, а номинальные — их имена.

Ли-вопрос.

Если субъект вопроса детерминирует конечное или небольшое и обозримое количество элементов, то вопрос с таким субъектом называется ли-вопросом.

Какой-вопрос.

Если субъект вопроса детерминирует бесконечное или очень большое количество элементов, то вопрос с таким субъектом называется какой-вопросом.

Интеррогативная формула.

Формально-логическая модель ли- или какой-вопроса называется интеррогативной формулой.

Ли-интеррогатив.

Ли-интеррогативом называется интеррогативная формула, при помощи которой моделирует ли-вопрос.

Ли-субъект.

Ли-субъект называется субъект ли-вопроса. Ли-субъект представляется списком утверждений, относительно которого вводятся следующие ограничения: (1) среди элементов списка нет повторяющихся утверждений; (2) список может включать как простые утверждения, так и их конъюнкции.

Вопрос типа да/нет.

Вопрос типа да/нет является частным случаем ли-вопроса. Субъект вопроса типа да/нет состоит из двух простых взаимно противоположных утверждений.

Матрица и категорные условия какой-вопроса.

Матрица и категорные условия являются описывают субъект какой-вопроса. Матрица представляет собой естественно-языковое предложение, в которое включены символы переменных. Категорные условия представляют собой естественно-языковые предложения, которые ограничивают значения переменных в матрице.

Какой-субъект.

Какой-субъектом называется субъект какой-вопроса. Какой-субъект представляется: (1) матрицей какой-вопроса; (2) множеством переменных в матрице какой-вопроса; (3) категорным отображением.

Категорное отображение какой-субъекта.

Категорное отображение какой-вопроса ставит в соответствие переменной в матрице какой-вопроса соответствующее категорное условие.

Категорно свободная переменная какой-вопроса.

Категорно-свободной переменной какой-вопроса называется переменная из матрицы какой-вопроса, которой не поставлено в соответствие категорное условие. Категор-

но-свободная никак не ограничена и может принимать произвольные значения.

Какой-интеррогатив.

Какой-интеррогативом называется интеррогативная формула, при помощи которой моделирует какой-вопрос.

2. ВОПРОСНО-ОТВЕТНЫЕ ТРАНЗАКЦИИ

Вопросно-ответный диалог.

Вопросно-ответный диалог является одним из типов естественного диалога, транзакции которого ограничены вопросно-ответными парами. В вопросно-ответном диалоге информационные послышки активного агента имеют стопроцентную степень интеррогативности и часто (но не всегда) представляют собой вопросы в лингвистическом смысле слова, а информационные послышки реактивного агента — ответы на эти вопросы.

Объект-вещь, объект-свойство, объект-отношение.

Объект-вещь, объект-свойство и объект-отношение — это формы существования объекта в сущностях, моделируемых средствами Языка Тернарного Описания.

Определённый, неопределённый, произвольный объекты.

Определённый объект, неопределённый объект и произвольный объект — это категории объектов в зависимости от степени неопределённости знаний об объекте в Языке Тернарного Описания.

Знания-that.

Эпистемологическое наименование знаний, выражаемых пропозицией, которыми обладает некоторый субъект.

Обоснованное и истинное убеждение (Justified True Belief — JTB).

Словесная формула, выражающая необходимые и достаточные условия знаний-that.

Конвергентные вопросы.

Конвергентные вопросы — это различные вопросы с одинаковыми истинными ответами.

Конвергентные знания.

Конвергентные знания — это знания, которые содержатся в ответах на конвергентные вопросы.

3. ПРОБЛЕМНО-НЕЗАВИСИМЫЕ МОДЕЛИ ДИАЛОГА

Когнитивная схема.

Когнитивная схема — это ментальное представление знаний о сущностях внешнего мира в долговременной памяти человека.

Когнитивная карта.

Когнитивная карта — это ментальное отображение фрагмента внешнего мира в долговременной памяти человека. Когнитивная карта представляет собой совокупность когнитивных схем.

Когнитивный цикл вопросно-ответного диалогового процесса.

Цикл, базирующийся на перцептивном цикле Найссера, в котором зрительный и слуховой анализаторы человека подключены к искусственной окружающей среде, формируемой потоками сенсорных событий, генерируемых противоположным агентом диалога.

Сценарий активного агента диалога.

Управляющий компонент когнитивного цикла диалогового процесса, являющийся аналогом когнитивной карты цикла перцепции Найссера.

Метод доступа к памяти вопросов, диалоговый метод решения проблемы.

Метод доступа к памяти вопросов — это метод доступа к памяти вопросов активного агента диалога на основании

текущего ответа реактивного агента и истории диалога. Является «хранилищем» метода достижения цели диалога или диалоговым методом решения проблемы.

Гармоничный диалога.

Гармоничным диалогом называется вопросно-ответный диалог, в котором диалоговые агенты удовлетворены своими ролями и не намерены их изменять.

Функция прямоты вопроса.

Функция прямоты вопроса является компонентом нечёткой автоматной модели вопросно-ответного диалога. Значение функции прямоты вопроса отражает «близость» текущего вопроса активного к прямому вопросу.

Функция ожидаемой правдивости ответов.

Функция прямоты вопроса является компонентом нечёткой автоматной модели вопросно-ответного диалога. Значение функции ожидаемой правдивости вопроса отражает «близость» текущего ответа реактивного агента к истинному ответу.

Унифицированный цикл вопросно-ответного диалога.

Компонент модели вопросно-ответного диалога в виде сети Петри. Представляет собой циклический процесс, управляющий вопросно-ответным диалогом и осуществляющий стандартную обработку любого шага диалога. Алгоритм унифицированного цикла реализует когнитивный цикл вопросно-ответного диалога.

Процессы-демоны.

Компоненты модели вопросно-ответного диалога в виде сети Петри. Процессы-демоны вызываются в том случае, когда формирование текущего вопроса или индекса последующего вопроса требует специальных средств, которые отсутствуют в алгоритме унифицированного цикла вопросно-ответного диалога. Существуют два класса процес-

сов-демонов: (1) процессы, используемые для формирования текущего вопроса; (2) процессы, используемые для формирования индекса следующего вопроса.

Узел прямой навигации.

Компонент сетевой даталогической модели вопросно-ответного диалога. Это узел сети диалогового метода доступа к памяти вопросов, для которого индекс следующего узла детерминируется только текущим ответом реактивного агента.

Узел условной навигации.

Компонент сетевой даталогической модели вопросно-ответного диалога. Это узел сети диалогового метода доступа к памяти вопросов, для которого индекс следующего узла детерминируется и текущим ответом реактивного агента и содержимым памяти истории ответов.

Целевой узел.

Компонент сетевой даталогической модели вопросно-ответного диалога. Это узел сети диалогового метода доступа к памяти вопросов, которым завершается диалоговый процесс решения проблемы.

Интеррогативный стимул.

Сложная информационная посылка активного агента, включающая вербальные и невербальные компоненты. Интеррогативный стимул воспринимается как зрительным, так и слуховым анализаторами реактивного агента.

Сцена.

Даталогическое представление интеррогативного стимула и ассоциированных с ним ожидаемых реакций реактивного агента.

Видео-объект сцены.

Компонент сцены, предназначенный для восприятия зрительным анализатором реактивного агента.

Аудио-объект сцены.

Компонент сцены, предназначенный для восприятия слуховым анализатором реактивного агента.

Атрибутивное описание объекта сцены.

Атрибутивное описание объекта сцены — это описание объекта при помощи набора атрибутов, сформулированных автором сцены.

Неатрибутивное описание объекта сцены.

Неатрибутивное описание объекта сцены соответствует информации, полученной от прототипа объекта, например, оцифрованного графического изображения, в случае видео-объекта.

Плохо формализуемые проблемы.

К плохо формализуемым относятся проблемы, обладающие одной или несколькими из следующих характеристик: (1) проблемы не могут быть заданы в числовой форме; (2) цели не могут быть выражены в терминах точно определённых целевых функций; (3) не существует алгоритмического решения проблемы; (4) алгоритмическое решение существует, но его нельзя использовать в связи с ограниченностью компьютерных ресурсов.

4. КОГНИТИВНЫЕ ТЬЮТОРЫ

Тьюторская система.

Тьюторская система — это система, предназначенная для решения плохо формализуемой проблемы обучения и передачи знаний при помощи диалогового процесса, в котором учеником является один из агентов диалога.

Идеальная тьюторская система (когнитивный тьютор).

Идеальная тьюторская система — это система, которая: (1) непрерывно адаптируется к текущим знаниям ученика

и степени понимания им учебного материала; (2) выбирает учебный материал, для которого имеет место несоответствие между текущими и требуемыми знаниями и/или степени понимания ученика; (3) автоматически, или по запросу ученика дополняет учебный материал тематически связанной информацией (например, информацией справочного характера); (4) уделяет больше внимания «трудным местам» и пропускает уже известный материал; (5) тестирует ученика и определяет его когнитивные характеристики и когнитивный профиль; (6) проводит процесс обучения методом, соответствующим текущим когнитивным способностям и когнитивному профилю ученика.

Система коммуникации знаний.

Наименование системы, предназначенной для решения проблемы обучения при помощи диалогового процесса, в ряде публикаций. Коммуникация знаний определяется, как способность инициировать и поддерживать целенаправленный обмен знаниями между двумя носителями знаний (тьютором и учеником) при помощи ограниченного набора коммуникационных ресурсов.

Модель предметной области.

Компонент системы коммуникации знаний. Модель предметной области реализует двойную функцию. С одной стороны, она представляет собой источник предметных знаний, «порционно» представляемых ученику, а с другой — желаемый стандарт знаний ученика (или цель обучения), который может использоваться для оценки его текущих знаний.

Модель ученика.

Компонент системы коммуникации знаний. Модель ученика предназначена, главным образом, для того, чтобы предсказывать диалоговое поведение ученика с целью его интерпретации. Идеальная модель ученика должна делать точные предсказания поведения любого ученика в любом контексте.

сте предметной области. Например, в виде множества возможных реакций/ответов на текущую обучающую сцену/вопрос.

Метод обучения.

Компонент системы коммуникации знаний. Метод обучения определяет какая обучающая сцена и на каком шаге диалога должна быть передана ученику.

Структурализм.

Одно из направлений в философии обучение, которое исходит из того, что знания существуют в виде множества «кусков» или фрагментов, а не в виде гомогенной структуры.

Конструктивизм.

Одно из направлений в философии обучение, которое исходит из того, что знания человека есть результат взаимодействия двух факторов: внешнего (обучающего) стимула и предыдущего ментального опыта. Упрощенный конструктивизм отрицает необходимость и возможность обучения путём трансмиссии знаний и предполагает, что единственным методом обучения является экспериментирование с предметом изучения на основе накопленного ментального опыта.

Обучение с ориентацией на деятельность.

Одно из направлений в философии обучение, которое исходит из того, что обучение осуществляется не в процессе пассивного восприятия обучающего материала учеником, а в процессе активного взаимодействия ученика с обучающим материалом.

Обучение с множественными связями.

Одно из направлений в философии обучение, которое исходит из того, что чем больше связей у изучаемого материала с ранее накопленными знаниями, тем больше шансов у этого

материала быть изученным, запомненным и впоследствии реконструируемым.

Компьютерная лингводидактика.

Компьютерная лингводидактика — это направление в теории обучения, которое занимается моделированием и разработкой компьютерных программ обучения иностранным языкам.

Ситуативный метод обучения.

Ситуативный метод обучения иностранным языкам основан на предположении, что знание языковых структур должно быть увязано с ситуациями, в которых они употребляются.

Аудио-лингвистический метод обучения.

Ситуативный метод обучения иностранным языкам основан на предположении о том, что первичной языковой средой является устная речь, поскольку во многих языках отсутствует письменность.

Метод общего физического отклика.

Метод общего физического отклика — это метод преподавания иностранного языка, опирающийся на координацию речи и моторной активности ученика.

Обучающая сцена.

Обучающая сцена — это сцена диалогового метода обучения, которая используется для проектирования обучающих эпизодов когнитивного тьютора.

Диагностическая сцена.

Диагностическая сцена — это сцена диалогового метода обучения, которая используется для проектирования эпизодов когнитивного тьютора, осуществляющих тестирование ученика с целью определения текущего уровня знаний, когнитивных способностей и когнитивного стиля.

Динамическая обучающая сцена.

В динамической обучающей сцене, как объекты обучающего стимула, так и объекты ожидаемых реакций появляются на экране не одновременно, а последовательно. Динамическая обучающая сцена постепенно разворачивается на экране, что позволяет ученику легко воспринимать даже сложную сцену.

Книга присвячена формальному моделюванню діалогових процесів типу «питання-відповідь» і використанню цих моделей для синтезу штучних діалогових агентів. У першій частині книги пропонується онтологічна модель природних діалогових процесів. На основі понять, що становлять онтологічну модель природного діалогу, будуються кілька формальних моделей діалогу типу «питання-відповідь», що відрізняються складністю і сферою застосування. Мережева даталогічна модель є теоретичною основою для синтезу штучного діалогового агента. У другій частині книги розглядається застосування штучного діалогового агента для вирішення погано формалізованих проблем. Увага фокусується на вирішенні проблем лінгводидактики. При вирішенні цих проблем штучний діалоговий агент розглядається як когнітивний т'ютор, що реалізує лінгводидактичну методику досвідченого викладача

Для фахівців та аспірантів у галузі штучного інтелекту, когнітивних наук та комп'ютерного навчання.

Наукове видання

ВЕРЛАНЬ Анатолій Федорович

ЧМИР Ігор Олексійович

ВЕРЛАНЬ Дмитро Анатолійович

ДІАЛОГОВІ ПРОЦЕСИ І КОГНІТИВНІ Т'ЮТОРИ

Російською мовою

Київ, Науково-виробниче підприємство
«Видавництво “Наукова думка” НАН України», 2022

Підп. до друку 00.00.0000. Формат 60×84/16. Папір офс. № 1.
Гарн. Таймс. Друк офс. Ум. друк. арк. 17,44.
Тираж 50 прим. Зам. № 00-00

ТОВ «Друкарня “Рута”»,
вул. Руслана Коношенка, 1, м. Кам'янець-Подільський, 32305.
Тел. (03849)42250