

# КОНВЕРЗАЦИОННЫЙ АГЕНТ В ИНФОРМАЦИОННО-СПРАВОЧНОМ ДИАЛОГЕ<sup>1</sup>

## CONVERSATION AGENT IN INFORMATION DIALOGUE

*М. Э. Койт (mare.koit@ut.ee)*

*Тартуский университет, г. Тарту*

Приводится модель конверсационного агента, которая состоит из нескольких модулей и применяет знания разных типов. Рассматривается представление знаний, в том числе, представление коммуникативных актов в виде фреймов, и регулярные выражения, представляющие структуру информационно-справочного диалога.

### Введение

Рассматривая телефонный диалог клиента и оператора справочной службы, мы можем предполагать, что оператор обладает некоторыми физическими и когнитивными способностями, т.е., что он способен слышать и понимать высказывания клиента, видеть информацию на экране, выбирать релевантную для клиента информацию, а также управлять общением.

Существующие диалоговые системы, общаясь с помощью естественной речи, пока не способны развивать диалог в любой области и могут лишь предлагать информацию в ограниченной предметной области.

Известный логик А. Тьюринг в 1950 году ввел т.н. тест Тьюринга для выяснения, обладает ли машина интеллектом или нет. Он предложил проводить диалог между человеком или машиной на одном конце, и человеком-экспериментатором на другом. Если экспериментатор не может доказать, что он общается с машиной, то машина прошла тест Тьюринга и мы можем сказать, что она обладает интеллектом. Начиная с 1991 года проводятся международные соревнования на приз Лебнера<sup>2</sup>, где участвующие программы (talk bots) стараются пройти тест Тьюринга. Пока это еще не удалось ни одной программе. Тем не менее, ежегодно выдается бронзовая медаль самой “человекоподобной” программе. Нужно подчеркнуть, что единственная цель таких программ – беседовать как можно “разумнее”, но не решить практические проблемы.

Долгое время разработчики ограничились с изучением использования языков программирования и доступа к базам данных, чтобы общаться с компьютерами. Теперь развитие компьютерной технологии дает возможность неопытным пользователям иметь прямой доступ к компьютеру, без поддержки программистов. Тем самым сильно изменилась парадигма человеко-машинного интерфейса. Вместо т.н. традиционных видов коммуникации, включающих манипуляцию иконками и текстовыми меню на экране компьютера, теперь имеются и другие альтернативы: вопросы, ответы или команды на естественном языке в виде текста или речи, общение с компьютером по телефону, настоящий кооперативный диалог, где компьютер помогает человеку решить проблемы.

ЕЯ вводит свою специфику в диалог. Эффективность естественной речи удивительна: если неопытный человек способен в минуту вводить с клавиатуры примерно 20 слов и писать 24 слова, то произносить он может до 150 слов в минуту (Minker, Vennacef 2004: 6). Для современных речевых ДС узким местом является распознавание речи. Это связано с недетерминистским характером процесса, в том числе, трудностями понимания коротких сообщений, сопровождающихся заиканием и восклицанием. Это отражается на понимании высказываний и тем самым повлияет на ход диалога.

Речевые ДС создаются во многих местах Соединенных Штатов, Канады, Японии и Европы. Значительное количество применений ДС связано с телефонными центрами (call-center).

В этой статье мы предлагаем модель системы, общающейся с человеком на ЕЯ, т.н. конверсационного агента. Разные подсистемы нашего агента в разной мере дошли до практической реализации. Например, существуют две ДС, общающиеся с человеком на эстонском языке<sup>3</sup>: Транспортный агент и Театральный агент. В обоих случаях имеет место простой диалог на эстонском языке, где анализ высказывания пользователя ограничен

<sup>1</sup> Работу поддерживает Эстонский научный фонд (грант 5685).

<sup>2</sup> <http://www.loebner.net/Prizef/loebner-prize.html>

<sup>3</sup> <http://www.dialoogid.ee>

с нахождением ключевых слов, необходимых для оформления запроса к базе данных, представляющей собой модель предметной области. Обе системы включают морфологический анализатор и генератор текста (Kaaler 1997), а также синтезатор речи (Mihkla et al. 1999); Театральный агент, кроме того – и экспериментальный распознаватель эстонской речи (Alumäe 2005). Для управления диалогом применяется простая регулярная грамматика.

### Конверзационный агент, общающийся на ЕЯ

Диалог можно рассматривать либо как процесс, в котором участники следуют своим коммуникативным целям, проводя рассуждения, чтобы составлять планы, и применяя коммуникативные стратегии, чтобы реализовать составленные планы, либо как текст или речь, являющиеся результатом этого процесса.

Нами разработана процедуральная модель диалога (Koit, Õim 2000, Койт 1999). Согласно этой модели, участник общения, или конверзационный агент А, можно рассматривать как программу, состоящую из шести компонентов (модулей):

$A = (ПЛ, РЗ, УД, ИНТ, ГЕН, ЛП)$ ,

где ПЛ – планировщик, РЗ – решатель задач, УД – управляющий диалогом, ИНТ - интерпретатор, ГЕН - генератор, ЛП – лингвистический процессор. ПЛ выступает как “центральный процессор”, который дает команды как управляющему диалогом, так и решающему задачи. ЛП выполняет морфологический и синтаксический анализ высказываний партнера, а также синтаксический и морфологический синтез высказываний самого агента. Задачей ИНТ является семантический анализ высказываний партнера, а ГЕН – семантический синтез высказываний агента. Если общение происходит в речевом регистре, то агент должен иметь еще два модуля: РР – распознаватель речи и СР – синтезатор речи. Агент применяет базу целей БЦ и базу знаний БЗ, которая состоит из четырех компонентов:

$BZ = (BZ_M, BZ_L, BZ_D, BZ_C)$ .

Здесь  $BZ_M$  – знание о предметной области (мире),  $BZ_L$  – лингвистическое знание,  $BZ_D$  – знание о диалоге,  $BZ_C$  – знание о субъектах общения (себе и других агентах).  $BZ_M$  содержит определения объектов и их отношений в предметной области (декларативное знание) и алгоритмы решения задач (процедуральное знание).  $BZ_L$  включает лексиконы применяемых языков (декларативное знание) и алгоритмы анализа/синтеза текста и речи (процедуральное знание).  $BZ_D$  содержит определения коммуникативных актов, а также алгоритмы, применяемые агентом для достижения коммуникативных целей, т.н. коммуникативные стратегии и тактики.  $BZ_C$  включает, с одной стороны, (реальные или предполагаемые) оценки субъектов к разным аспектам мира (например, что является приятным, полезным или вредным) и алгоритмы, применяемые для составления планов действий на основе оценок.

Необходимое условие для того, чтобы общение могло иметь место – существование разделенного знания: участники должны иметь общую цель (общение), общий язык, общее представление о мире и нормах общения, а также хотя бы частично правильное представление о других участниках. Если участниками общения являются А и Б, то это означает, что  $BZ^A \cap BZ^B \neq \emptyset$ ,  $BZ_L^A \cap BZ_L^B \neq \emptyset$ ,  $BZ_M^A \cap BZ_M^B \neq \emptyset$ ,  $BZ_D^A \cap BZ_D^B \neq \emptyset$ ,  $BZ_C^{AB} \cap BZ_C^B \neq \emptyset$ ,  $BZ_C^{BA} \cap BZ_C^A \neq \emptyset$ .

РР переводит высказывание партнера (устную речь) в текст, затем ЛП выполняет морфологический и синтаксический анализ высказывания. Полученное синтаксическое представление идет в следующий блок – ИНТ, который выполняет семантический анализ, а также определяет коммуникативную структуру высказывания партнера, т.е. распознает коммуникативный акт высказывания (приветствие, запрос, предложение, ответ на вопрос и т.д.). Тем самым интерпретатор выдвигает цели двух типов. Во-первых, коммуникативную цель, которая зависит от типа коммуникативного акта. Например, за вопросом, как правило, следует ответ, за предложением – его принятие или отклонение. Таким образом, коммуникативная цель – ответить на предложение или вопрос партнера, задать вопрос партнеру для получения информации, и т.д. Во-вторых, предметная цель зависит от семантической структуры высказывания и заключается в решении определенной задачи. Например, для того что ответить на вопрос “Сколько стоит билет на поезде из Тарту в Москву, с пересадкой на станции Тапа?”, нужно вычислить сумму двух известных (найденных в базе данных) цен. Обработкой целей первого типа занимается УД, второго типа – РЗ. В результате работы УД определяется коммуникативная структура ответного высказывания, а в результате РЗ – его семантическое содержание. ГЕН образует семантическое представление ответного высказывания, ЛП оформляет его в виде текста на ЕЯ, а СР переводит текст в речь.

### Коммуникативные стратегии и тактики

Знание о диалоге  $BZ_D$ , применяемое модулем управления диалогом, состоит, с одной стороны, из правил общения, и, с другой стороны, из правил построения и сочетания коммуникативных актов.

Правила общения задаются *коммуникативной стратегией* – алгоритмом, применяемым агентом для

достижения своей коммуникативной цели. Агент может реализовать свою стратегию через различные коммуникативные тактики. Например, если целью агента А является, привести партнера Б к принятию решения выполнить некоторое действие Д, то А может завлекать, убеждать или угрожать Б принимать это решение. В таком случае мы будем говорить, что А применяет, соответственно, тактику завлечения, убеждения или. Коммуникативная тактика – это алгоритм для построения очередного коммуникативного акта. На рис. 1 приведена тактика угрозы.

```

/* угроза */
ЕСЛИ найдется аргумент для увеличения меры наказания, ТО
    выбрать аргумент,
В ПРОТИВНОМ СЛУЧАЕ
    неудача.
    
```

Рис. 1. Тактика угрозы

У агента Б также имеются различные коммуникативные тактики, например, тактика сопротивления (рис. 2).

```

/* сопротивление */
ЕСЛИ найдется аргумент к отказу делать Д, ТО
    выбрать аргумент,
В ПРОТИВНОМ СЛУЧАЕ
    неудача.
    
```

Рис. 2. Тактика сопротивления

Коммуникативные тактики, применяемые в информационно-справочных диалогах, рассматриваются и другими авторами, например, в конструктивной модели диалога (КМД, Jokinen 1996). Каждое высказывание строится для того, чтобы приближаться к некоторой коммуникативной цели. Согласно КМД, коммуникативная тактика<sup>4</sup> определяется четырьмя факторами: 1) ожидания – соответствовало ли высказывание партнера ожиданиям говорящего или нет; 2) центральное понятие – соответствовало ли высказывание партнера теме или нет; 3) цели – достигнуты ли все цели говорящего или нет (компьютер, будущий кооперативным партнером, принимает цели клиента как свои собственные); 4) инициативы – имеет ли говорящий инициативу или нет. Анализ эстонских устных информационно-справочных диалогов (звонков на справочный телефон) показывает, что оператор наиболее часто применяет тактику „продолжение старой темы“, что выражает реакцию на ожидаемый акт, продолжающий текущую тему, цели еще не достигнуты, и инициативу удерживает слушающий, но не говорящий (например, оператор отвечает на вопрос клиента). Это понятно, так как оператор справочного телефона обычно выполняет пассивную роль.

Выбор тактики зависит от точки в т.н. пространстве общения, в которой агент А находится в данный момент общения с Б. Коммуникативное пространство определяется такими параметрами как коммуникативное расстояние между А и Б, кооперативность, персональность, модальность, интенсивность общения (Koiv, Oim 2000). На рис. 3 представлена коммуникативная стратегия инициатора общения. Начиная диалог, агент имеет некоторую цель, для достижения которой он выбирает тактику, а также начальную точку в коммуникативном пространстве. В ходе общения у него могут возникать новые цели, являющиеся подцелями исходной. Для хранения целей агент применяет магазин, откуда он берет очередную цель (которой будет следовать в своем очередном высказывании) в обратном порядке возникновения целей. Диалог заканчивается, если либо магазин пуст, либо агент решил отказаться от первоначальной цели. В ходе общения агент может переместиться в коммуникативном пространстве из одной точки в другую, а также переменить тактику.

```

Выбрать коммуникативную цель, занести ее в магазин целей
выбрать точку в коммуникативном пространстве
выбрать тактику
генерировать высказывание
ПОВТОРИТЬ, пока магазин не пуст или не решено сдаться
    интерпретировать высказывание партнера
    если цель достигнута, то удалить ее из магазина
    если партнер выдвиг новую цель, то занести в магазин
    переменить коммуникативную точку, если нужно
    переменить тактику, если нужно
    генерировать свое высказывание
    
```

Рис. 3. Коммуникативная стратегия инициатора общения

<sup>4</sup> Коммуникативная стратегия в смысле Екинен.

В ходе общения агент может перемещаться в коммуникативном пространстве с одной точки в другую, а также переменить тактику. Например, имея цель заставить партнера Б принять решение о выполнении некоторого действия, агент А может начинать общение дружелюбно и персонально, стараясь убеждать Б, а затем, когда кончаются убедительные аргументы, перейти на враждебность и угрожение.

В информационно-справочном диалоге инициатором является клиент, который задает запрос оператору. Оператор (роль которого может выполнять ДС) может работать по аналогичному алгоритму, с той разницей, что в кооперативном диалоге не выбирает цели самостоятельно, но принимает цели клиента в качестве своих.

### Структура диалога

Общую структуру простого информационно-справочного диалога можно представить с помощью регулярного выражения (где исключены ритуалы, обычно присутствующие в начале и конце общения):

**диалог = (запрос (обмен информацией)\* ответ (объяснение)\*)<sup>+</sup>**

Ядро диалога образуется смежной парой коммуникативных актов “запрос (клиента) – ответ (оператора)”. Получив запрос, оператор может начинать т.н. поддиалог “обмен информацией”, задавая уточняющие вопросы, с целью понимать желание клиента (см. рис. 4; пример из корпуса эстонских диалогов<sup>5</sup>). Получив ответ, клиент, в свою очередь, может начинать поддиалог “объяснение” (см. рис. 5).

(звонок)  
 О(ператор) : справочная служба  
 Керсти  
 здравствуйте  
 К(лиент) : здравствуйте,  
 /\* запрос \*/  
 я хотела бы узнать номер телефона салона красоты в Доме Э-торговли  
 О : да  
 один момент  
 /\* поддиалог “обмен информацией” \*/  
 вы думаете Э-салон красоты?  
 К : да  
 /\* ответ \*/  
 О : а отмечено, что он больше не действует  
 К : ага,  
 спасибо  
 О : да пожалуйста

**Рис. 4.** Диалог, включающий поддиалог “обмен информацией” (звонок на справочный телефон, перевод с эстонского)

Диалог, содержащий переговоры, можно представить как регулярное выражение:

**диалог = (предложение (переговоры)\* [принятие | отклонение])<sup>+</sup>**

/---/  
 /\* ответ \*/  
 О : восемнадцать сто, платный короткий номер.  
 /\* объяснение \*/  
 К : а семь в начале не нужно набрать?  
 О : нет  
 прямо набираете  
 К : восемнадцать сто?  
 О : да  
 К : спасибо  
 О : пожалуйста

**Рис. 5.** Диалог, включающий поддиалог “объяснение” (звонок на справочный телефон, перевод с эстонского)

Таким образом, рассматриваемые диалоги могут включать поддиалоги трех видов (табл. 1).

<sup>5</sup> <http://www.ut.ee/~koit/Dialoog/EDiC/>

А: запрос /*обмен информацией*/ Б: А: Б: ответ	А: запрос Б: ответ /*объяснение*/ А: Б:	А: предложение /*переговоры*/ Б: А: Б: принятие/ отклонение
---	---	--

Табл. 1. Виды поддиалогов

Наряду с соседними парами “запрос – ответ”, ядро диалога может содержать и соседние пары “предложение – принятие/отклонение”. За предложением может следовать поддиалог, где взвешиваются положительные и отрицательные стороны его принятия и отклонения, т.н. переговоры (пример рис. 6).

### Коммуникативные акты

Общаясь друг с другом, конверсационные агенты обмениваются коммуникативными актами (запрос, предложение и т.п.). Мы в нашей модели применяем ограниченное количество коммуникативных актов, для представления которых используем фреймы. Каждый коммуникативный акт включает статическую и динамическую части. Статическая (декларативная) часть состоит из 1) предпосылок, 2) цели, 3) содержания и 4) результатов акта.

```

/---/
/* запрос */
К: я заинтересован в поездках в страны Скандинавии
/* ответ */
О: да
но дело в том, что мы их, в крайней мере сейчас, не организуем
К: ага
/*предложение */
О: но у нас есть поездка в Лапландию – вы были бы заинтересованы?
это в декабре
/* переговоры */
К: мм да
а что это точнее?
О: поедем в гости к Деду Морозу
К: да
О: длина путешествия 4 дня
К: угу
О: и даты 27 по 30 декабря
/---/

```

Рис. 6. Диалог, включающий поддиалог “переговоры” (звонок в бюро путешествий, перевод с эстонского)

Динамическая (процедуральная) часть содержит процедуры двух видов: 1) запускаемые автором для порождения акта и 2) запускаемые адресатом для интерпретирования акта партнера и порождения своего ответного акта. В качестве примера на рис. 7 приведен фрейм “Общий вопрос”.

```

Общий вопрос
Статическая часть
Предпосылки:
А хочет узнать, имеет ли место р
А считает, что Б знает, имеет ли место р
Цель: Б знает, что А хочет узнать, имеет ли место р
Содержание: А сообщает Б, что хочет узнать, имеет ли место р
Результат: Б знает, что А хочет узнать, имеет ли место р
Динамическая часть
Процедуры порождения, применяемые А: образовать вопрос “р?”
Процедуры интерпретирования-порождения, применяемые Б: распознать тип вопроса
по вопросительному слову, обратиться к базе данных для нахождения ответа,
ответить либо “да”, либо “нет”.

```

Рис. 7. Фрейм “Общий вопрос”

В случае общего вопроса клиент, узнав у оператора, например, что автобус на X. отправляется в 6.00, может инициировать поддиалог "объяснение", задавая (общий) вопрос "Это экспресс?". Здесь р="автобус на X, отправляющийся в 6.00, является экспрессом", и оператор первым делом должен разрешить анафор "это", а затем обратиться к расписанию автобусов, чтобы найти подходящий ответ "да" или "нет".

Существует несколько методов для автоматического распознавания коммуникативных актов, начиная с применения простых лингвистических "сигналов" до применения методов распознавания планов говорящего (Jurafsky, Martin 2000).

Для эстонского языка проведен эксперимент распознавания коммуникативных актов в текстах (транскрипциях устных) диалогов, с помощью нейронных сетей с применением методов машинного обучения (Fishel 2005). Построенная нейронная сеть учитывает результаты морфологического анализа высказываний, пытаясь найти для каждого распознаваемого акта подходящие ключевые слова.

### **Знания о мире**

В случае хорошо определенной предметной области, знания, нужные для интерпретирования высказываний, удобно представлять в виде фреймов. Например, если справочная система настроена на сообщение информации о театральных спектаклях, то фрейм "посещение театра" может иметь слоты "город", "название театра", "название спектакля", "число", "время начала", "время окончания", "цена билета", "существование свободных мест" и т.п. В случае планирования путешествия применяется фрейм "путешествие", который содержит слоты "откуда", "куда", "вид транспорта". Таким образом, можно применять т.н. семантическую грамматику, которая дает возможность упрощать анализ, а также генерацию высказываний (Minker, Bennacef 2004). Вместо полного цикла анализа высказывания достаточно найти по ключевым словам исполнителей слотов основного фрейма. Такой подход принят в двух выше упомянутых экспериментальных ДС Транспортный агент и Театральный агент, дающих пользователю информацию, соответственно, о рейсах самолетов и театрах Эстонии.

### **Заключение**

В данной статье мы представили конвезационного агента – программу, состоящую из нескольких модулей, каждая из которых выполняет свою подзадачу в информационно-справочном диалоге. Управление диалогом происходит по простому регулярному выражению, которая представляет ядро информационно-справочного диалога в виде смежной пары диалоговых актов, возможно дополненных т.н. поддиалогами (и подподдиалогами), которые, в свою очередь, представляют собой смежные пары.

Наша следующая работа – совершенствование существующих экспериментальных диалоговых систем и их обобщение – построение такой системы (конвезационного агента-оператора), которая способна, по запросу клиента, настраиваться на нужную предметную область.

### **Список литературы**

- Alumäe T. *Phonological and morphological modeling in large vocabulary continuous Estonian speech recognition system* // Proc. of Second Baltic Conference on Human Language Technologies, Tallinn, 2005, 249-254.
- Fishel M. *Dialogue act recognition in Estonian using artificial neural networks* // Proc. of the 2<sup>nd</sup> Baltic Conference on Human Language Technologies, Tallinn, 2005, 231-236.
- Jokinen K. *Cooperative response planning in CDM: Reasoning about communicative strategies* // Proc. of COLING. Stanford, 1996, 444-447.
- Jurafsky, D., Martin, J.H. *An introduction to natural language processing, computational linguistics, and speech recognition*. Prentice Hall, 2000.
- Kaalep H.-J. *An Estonian Morphological Analyser and the Impact of a Corpus on Its Development* // Computers and the Humanities. 1997, 115-133.
- M. Koit, H. Õim. *A model of dialogue and its application: building a communication trainer* // Тр. конф. КИИ'2000. М.: Изд. Физико-математической литературы, 2000. Т. 1. С. 345-353.
- Mihkla, M., Eek, A., Meister, E. *Text-to-Speech Synthesis of Estonian* // Proc. of the 6<sup>th</sup> European Conference on Speech Communication and Technology. Budapest, 1999. Vol. 5, 2095-2098.
- Minker W., Bennacef S. *Speech and Human-Machine Dialog*. Boston/Dordrecht/London: Kluwer Academic Publishers, 2004.
- Койт М. *О реализации диалоговой системы* // Обработка текста и когнитивные технологии. Сборник. Вып. 2, Москва-Пушино, 1999. С. 32-50.