

И.М. КОБОЗЕВА, Г.О. СИДОРОВ, А.В. ЦИММЕРЛИНГ
**МОДУЛЬ УПРАВЛЕНИЯ ДИАЛОГОМ В СИСТЕМЕ ОБЩЕНИЯ
ПОЛЬЗОВАТЕЛЯ С ПОДВИЖНЫМ РОБОТОМ-ГИДОМ**

Кобозева И.М., Сидоров Г.О., Циммерлинг А.В. Модуль управления диалогом в системе общения пользователя с подвижным роботом-гидом.

Аннотация. Данная работа посвящена описанию диалогового модуля, разработанного для подвижного робота-гида. Модель управления диалогом представлена как сеть переходов системы из одного состояния в другое в зависимости от двух факторов: коммуникативного и визуального. Описываются детали модуля управления диалогом.

Ключевые слова: робот-гид, моделирование диалога, управление диалогом, речевые акты.

Kobozeva I.M., Sidorov G., Zimmerling A. Module for dialog management in the interaction system between user and mobile robotic guide.

Abstract. The paper presents dialogue management module for a mobile robot as a guide. The model of the dialog management is presented as a network of transitions between states depending on two factors: visual and communicative. Details of the dialogue management module are presented.

Keywords: mobile robot as a guide, mobile robotic guide, dialog modelling, dialog management, speech acts.

1. Введение. В настоящей работе мы представляем алгоритмический подход к проблеме моделирования диалога на естественном языке, используя в качестве примера диалог с подвижным роботом-гидом. Необходимо отметить, что в современной компьютерной лингвистике кроме алгоритмического часто применяется и другой подход — основанный на машинном обучении. Скажем, в случае диалога, метод, основанный на машинном обучении, применялся для центральной проблемы его автоматической обработки: обнаружения в нем речевых актов [21].

Заметим, что для многих задач, связанных с компьютерной лингвистикой, таких как определение плагиата, автоматическое построение резюме, определение полярности текста и т.п., методы машинного обучения на основе размеченных корпусов считаются наиболее подходящими [14]. Для применения этих методов строится модель векторного пространства (vector space model). Вообще говоря, модель векторного пространства это стандартный способ вычислять сходство пар любых объектов, в том числе и текстов, представленных как наборы значений признаков (векторы). Признаки являются координатными осями (измерениями) в таком пространстве. Очевидно, что оно может иметь очень много измерений. В качестве меры сходства вычисляется косинус угла между векторами в таком многомерном пространстве:

чем меньше угол между векторами, тем больше похожи соответствующие объекты. Модель была предложена в [12], ее доступное подробное описание можно найти, например, в [14]. Обычно следуют такой схеме обработки данных: выбор задачи — подготовка корпуса — разметка корпуса (обычно вручную, «золотой стандарт») — выбор признаков и их значений (построение векторного пространства) — уменьшение размерности этого пространства (например, с использованием скрытого семантического анализа, *latent semantic analysis* [14]) — выбор одного или нескольких методов машинного обучения (например, *naive Bayes*, *support vector machines*, *J48* и др.) и их параметров — применение этих методов — выбор методов для сравнения (*baseline*) — оценка результатов (обычно с перекрестным разбиением на 10 частей для размеченного корпуса, *10-fold cross validation*) — интерпретация результатов. Для построения векторного пространства для текстов обычно в качестве признаков используются *n*-граммы разных типов (включая слова, которые являются 1-граммами). В качестве значений признаков в векторном пространстве традиционно берутся разные варианты меры *tf-idf*.

При использовании этих методов мы дополнительно рекомендуем применение синтаксических *n*-грамм, которые строятся, следуя зависимостям в синтаксическом дереве [16]. Заметим, что синтаксические *n*-граммы, примененные для определенного круга задач, впервые позволяют прямо ввести синтаксическую информацию в чисто статистические методы машинного обучения. Важно помнить, что мы не говорим об автоматическом обучении грамматикам для синтаксического анализа (что достаточно традиционно), а о применении результатов этого анализа — представленных в виде синтаксических *n*-грамм — в других задачах, основанных на модели векторного пространства. То есть, мы отвечаем на вопрос, как принять во внимание структуру синтаксического дерева при определении, скажем, полярности текста.

Тем не менее, не для всякой задачи можно применить подход с машинным обучением, например, для задач определения грамматических значений слов, частеречной принадлежности слов, машинного перевода, распознавания речи. В этих случаях необходимы более традиционные алгоритмические подходы и модели. Скажем, в вышеупомянутой работе [21] не было получено убедительных данных, что машинное обучение позволяет хорошо моделировать диалог. И это притом, что речевые акты явно представляют собой потенциально хороший материал для использования машинного обучения. Проблема состоит в том, что признаки речевых актов имеют семантическую и

прагматическую природу и на данном этапе развития теории обработки текста их очень сложно определять.

Как мы уже упоминали, данная статья основана на алгоритмическом подходе к описанию диалогового модуля для взаимодействия с подвижным роботом-гидом в рамках общего подхода, описанного в работе [15]. Мы описываем систему диалога с роботом Pioneer 3DX, которая включает модуль распознавания речи Dragon NaturallySpeaking и систему синтаксического анализа Freeling [9]. Система разрабатывается для испанского языка, тем не менее лингвистические модули (синтаксический анализ, распознавание речи) работают также и для русского языка. Кроме того, важно заметить, что в то время как набор синтаксических ограничений на порядок слов и перемещения элементов, согласно накопленным типологическим данным [22] лингвоспецифичен, инвентарь элементарных речевых актов и коммуникативных стратегий универсален, то есть, модуль управления диалогом не зависит от языка. Так что в дальнейшем примеры в статье мы приводим на русском языке.

Робот Pioneer 3DX, снабженный точным лазерным датчиком измерения расстояний (до возможных препятствий на его пути), является стандартом де факто для разработки приложений автономных подвижных роботов в реальном мире — его размеры сопоставимы с человеком. Робот движется по предварительно подготовленной карте своего окружения, на которой помечаются специально выделенные объекты или места, например, в данной статье мы считаем помеченным офис администрации (Admin) и лабораторию искусственного интеллекта (IA). В дальнейшем эти названия используются в нашем описании модели диалога, но в каждой реализации этой модели они должны быть заменены на другие существенные объекты в специфическом окружении. Робот очень надежно перемещается в таком пространстве, в котором могут появляться новые для него (отсутствующие на карте) объекты, которые он легко объезжает. Робот знает, когда каждый объект становится видимым человеку — это происходит, если между ними нет препятствия, и расстояние стало достаточно небольшим: мы предлагаем рассматривать порог до 3 метров. В нашей модели, когда какой-либо объект X попадает в «поле зрения» (напомним, что не камеры, а датчика), это отражается как элемент `visual_act(X)`.

Лингвистическое обеспечение естественного диалога человека с подвижным роботом является одним из самых трудных вариантов уже давно поставленной общей задачи лингвистического обеспечения диалога на естественном языке (ЕЯ) в человеко-машинных системах, см. обзор [24]. Трудность связана с мультимодальностью информации,

которую робот должен «понимать», с принципиальными различиями между языками репрезентации пространства в «сознании» робота и в сознании человека, не говоря уже об общей для всех систем человеко-машинного общения проблеме понимания ЕЯ на уровне, позволяющем разрешать референцию и правильно интерпретировать интенции собеседника [10,11,13].

Развернувшиеся в последние двадцать лет работы по моделированию общения человека с подвижным роботом, как правило, носят исследовательский характер и углубленно анализируют тот или иной аспект комплексной проблемы, например, влияние способности робота жестикулировать на успешность коммуникации, важность направления взгляда и ситуационных знаний для эффективного взаимодействия [6, 25]. С другой стороны, имеются и разработки действующих подвижных роботов, способных выполнять те или иные функции и вести диалог с пользователем на приемлемом уровне естественности и надежности, напр., [1, 2, 7, 17, 19].

2. Функции мобильного робота. Среди функций робота, для поддержания которых разрабатываются диалоговые модули, можно назвать следующие:

1) функция проводника (например, гида в музее или на выставке; пример — робот-гид Robobox (подробное описание представлено на сайте <http://lis2.epfl.ch/resources/podcast/2008/01/roland-siegwart-autonomous-robots.html>), созданный Р. Зигвартом (Roland Siegwart), работавший на швейцарской национальной выставке Expo02);

2) функция ученика, которого обучают понимать свое окружение (см. [4, 5, 20]);

3) функция транспортера, как, например, у инвалидной коляски, которой пользователь управляет при помощи голосовых команд;

4) функция посыльного, движущегося к какой-либо цели самостоятельно или по командам, исходящим от пользователя:

i. в ситуации, когда робот «знает» путь;

ii. в ситуации, когда он не знает пути и, чтобы пройти по нему, должен общаться с пользователем в удаленном доступе [2];

5) функция репортера, который сообщает, что он в данный момент воспринимает;

6) функция официанта в ресторане;

7) функция помощника при выборе объектов, например, в магазине при покупке;

8) функция манипулятора, который выполняет команды пользователя по перемещению объектов (например, самый первый робот Т. Винограда SHRDLU).

Разумеется, робот может быть предназначен для нескольких функций. Часто сочетаются функции посыльного с функцией манипулятора: робот должен по команде пользователя переместиться в нужное место и произвести то или иное действие с указанным объектом. Функции ученика и репортера могут сочетаться и между собой и с функцией посыльного, как в [18], где моделируется диалог для робота, способного перемещаться по команде пользователя, рапортовать о том, что он воспринимает, и научиться от пользователя именам воспринимаемых объектов.

Каждая функция предъявляет определенные требования к структуре диалога и репертуару коммуникативных актов (реплик и/или невербальных реакций). Так, в зависимости от функций, инициатором и контролером хода диалога выступает либо пользователь, либо робот, либо инициатива и контроль могут переходить от робота к пользователю по мере развития диалога. В функции гида робот выступает инициатором и контролером хода диалога, а в остальных случаях диалог инициируется и контролируется пользователем или обоими коммуникантами попеременно. Чем больше функций у робота, тем шире репертуар коммуникативных актов, которые он должен порождать и понимать, тем важнее наличие в архитектуре системы концептуального и интенционального модуля (энциклопедической информации и модели рационального поведения).

В нашем случае диалоговый модуль разрабатывается для системы, не имеющей такого модуля «обширных знаний», и поэтому в качестве моделируемой функции избирается сравнительно простая функция робота-гида, предполагающая контроль за ходом диалога со стороны робота и сужение благодаря этому репертуара необходимых коммуникативных актов (КА).

3. Специфика диалога «гид–посетитель». Под гидом в данном случае понимается не только представитель определенной профессии, а любой агент, который в определенной ситуации решает задачу ознакомления посетителей с тем или иным местом, в которое они прибыли, чтобы узнать о нем больше, чем они знали ранее, и/или с теми объектами, которые находятся в этом месте. Таким местом может быть та или иная местность (напр., заповедник), населенный пункт (напр., город), чье-то владение (напр., усадьба помещика) или интерьер какого-то здания или его части (напр., музей, школа, этаж большого учреждения, квартира), а гидом может быть хозяин дома, риэлтор, сотрудник научного учреждения, местный житель, и вообще любой человек, взявшийся выполнить указанную социальную функцию. В любом случае исполнение функции гида начинается с того момента, как агент

принимает на себя эту роль, что в свою очередь предполагает заранее установленным, что собеседник выступает в роли «посетителя незнакомого места» и имеет соответствующее этой роли намерение. В случае с роботом ответственность за уместность принятия им на себя роли гида может лежать на человеке, который запускает (triggers) процесс в нужный момент, то есть, человек может, например, вызвать соответствующую программу и робот начнет исполнять свою роль.

Общая задача гида — ознакомить посетителей с местом, — разбивается на ряд подзадач, каждая из которых состоит в том, чтобы подвести их к объекту (в широком смысле), представляющему для них интерес и, возможно, сообщить тот или иной объем сведений об этом объекте. В ходе решения каждой из подзадач гид может дополнительно обращать внимание на встречающиеся по пути к очередной цели «второстепенные» объекты, например, ведя путников по лесу к озеру, которое они хотели увидеть, отметить встретившуюся по дороге поляну, на которой можно расположиться на отдых.

В свою очередь, решение каждой подзадачи предполагает совершение коммуникативных (вербальных и/или невербальных) актов и актов перемещения к заранее намеченному объекту интереса по обычно заранее известному для гида пути.

4. Речевые акты и речевые блоки в диалоге робота-гида с посетителем. В нашем случае модель диалога разрабатывается для робота-гида, проводящего посетителей по одному из этажей института. Для этой конкретной ситуации в репертуар действий робота включается три вида актов — речевые акты, речевые блоки и мультимедальные блоки.

1) *Речевые акты* мы понимаем стандартным для теории речевых актов способом — как акты произнесения языкового выражения, обычно равного отдельному простому предложению, имеющего определенную иллокутивную силу, т.е. выражающего (прямо или косвенно) цель говорящего наряду с рядом других прагматических компонентов смысла [13].

Набор речевых актов робота, релевантных для данной ситуации, включает в себя:

- вопрос о принятии предложения о туре (Tour-Question),
- предложение выбрать первый объект осмотра из набора альтернатив, в нашем случае, офис администрации (Admin) и лабораторию ИИ (AI) (Question_Altern(Adm, AI)),
- завершение диалога в случае отказа от предложения (Closure).

Каждому речевому акту соответствует набор реплик, которые варьируются, чтобы избежать повторения одной и той же фразы. Примеры реплик приводятся ниже.

2) *Речевые блоки* мы понимаем как сложные речевые акты, состоящие из последовательности речевых актов (РА) одного говорящего, связанные между собой единством общего речевого замысла в одно целое и предполагающих реакцию адресата на это целое, а не на отдельные составляющие его РА.

В нашем случае используется три вида речевых блоков: (а) Пограничные блоки (BB, от Borderline Blocks), (b) Нарративные блоки (NB от Narrative blocks) и (c) Второстепенные блоки (SB от Secondary Blocks).

2а) *Пограничные блоки* включают Вводный блок (IB от Introductory Block) и Терминальный блок (TB от Terminal Block). Каждый блок состоит из последовательности элементарных речевых актов (РА).

Вводный блок (IB) — это последовательность вида (в заданном порядке):

Приветствие (Greeting) + Представление себя (Introduction) + Предложение осмотра помещения (Tour-Proposal).

Этот блок запускается в исходной фазе, которая принадлежит типу фаз говорения (о типах фаз см. ниже).

Терминальный блок (TB) — это последовательность вида:

Объявление о конце осмотра (Announce_End) + Этикетная самооценка (Evaluation) + Прощание (Farewell).

Этот блок запускается в конечной фазе, которая определяется с опорой на историю диалога. Ниже в формальной репрезентации диалога переход в фазу произнесения терминального блока обусловлен контекстным ограничением вида «NB(X)» \in HD, где HD — протокол данного диалога (history of the dialogue), а X принимает одно из двух значений — Adm или AI, т.е. отслеживается тот факт, что оба объекта осмотра уже пройдены.

2b) *Нарративные блоки* (NB) представляют собой тексты или презентации, содержащие определенный объем информации об объектах показа (в нашем случае — о лаборатории ИИ и офисе администрации). Нарративные блоки запускаются по достижении роботом объекта показа, точнее, после того, как робот завершит мультимодальный блок типа MMB2 (см. ниже). В нарративный блок можно заранее добавлять нужную информацию.

2c) *Второстепенный блок* (SB от Secondary block) произносится роботом в ситуации, связанной с возможностью в ходе осмотра поме-

щения воспользоваться удобствами (туалет и кулер), и идет в связке с второстепенным мультимодальным блоком SMB (см. ниже).

Второстепенный блок — это отвлечение от основной цели: заявление о намерении подождать, пока пользователь воспользуется удобством (Promise_Wait) + просьба пользователю сообщить о своем возвращении (Request_Return).

Этот блок запускается положительным ответом пользователя на предложение воспользоваться удобством.

3) *Мультимодальные блоки* — это последовательности речевых актов (РА) и неречевых действий, совершаемых в фиксированном порядке. Они делятся на основные и второстепенные.

3а) *Основные мультимодальные блоки* (MMB от Main Multimodal Blocks). В системе представлены два вида MMB:

MMB1: просьба (команда) двигаться вслед за роботом, которую тот произносит (Imper_Move) + движение робота к цели (Move(X)), где $X \in \{Adm, AI\}$.

Заметим, что мы предполагаем, что пользователь будет следовать за роботом, то есть выполнять команду (Imper_Move), поскольку он согласился на осмотр. При необходимости проверять, что пользователь все еще следует за роботом, можно ввести в нашу модель дополнительное подтверждение: робот может спрашивать о подтверждении продолжения осмотра каждые 5-10 минут, и в случае отсутствия ответа или отрицательного ответа прекращать тур.

Предполагается, что робот-гид всегда начинает тур с фиксированной исходной точки L (в нашем конкретном приложении — от входа на этаж с лестницы) и движется к первой принятой цели кратчайшим путем, а ко второй цели — по маршруту, предполагающему прохождение точки L, при этом, напомним, робот способен обходить препятствия, которых нет на его карте.

Первый в ходе диалога блок данного иллокутивного типа запускается ответом пользователя на речевой акт (РА) робота типа Question_Altern(Adm, AI) (см. выше). Второй блок данного типа запускается отрицательным ответом пользователя на вопрос робота о желании воспользоваться некоторым удобством и произносится по завершении первого нарративного блока.

MMB2: остановка робота у цели (Stop_X) + команда остановиться (Imper_Stop). Этот блок запускается визуальным актом, представляемым в модели как visual_act(X), где $X \in \{Adm, AI\}$.

3б) *Второстепенные мультимодальные блоки* (SMB от Secondary Multimodal Blocks) — это последовательность действий, связанная с второстепенными объектами на пути (в нашем случае это

удобства: туалет и кулер, репрезентируемые в модели как WC и Cooler соответственно).

SMB: остановка около второстепенного объекта (Stop(X)) + привлечение внимания к второстепенному объекту Announce_Utility(X) + вопрос о желании воспользоваться удобством X, Question_Utility(X), где $X \in \{\text{Cooler, WC}\}$. SMB, как и MMB2 запускается визуальным актом, представляемым в модели как visual_act(X), где $X \in \{\text{Cooler, WC}\}$, то есть робот «видит» данный объект, отмеченный на карте, и находится достаточно близко от него.

Таким образом, среди действий робота есть как иллокутивно-независимые, так и иллокутивно-зависимые в смысле Баранова и Крейдлина (1992) [23]. К иллокутивно-независимым, т.е. запускаемым по «инициативе» робота, относятся речевые блоки типа BB, NB, мультимодальные блоки MMB2 (Stop(X) + Imper_Stop) и SMB. К иллокутивно-зависимым, т.е. запускаемым при условии реализации пользователем того или иного речевого акта (РА), относятся речевой блок SB и речевые акты Question_Altern и Ask_Confirm, а также MMB1 (Imper_Move + Move(X)).

Каждому иллокутивному типу речевого акта или речевого блока, а также речевого компонента из мультимодального блока в модуле синтеза текста ставится в соответствие набор речевых шаблонов (реплик), из которых робот делает выбор, запоминая, какой вариант уже был выбран с тем, чтобы в следующую итерацию выбрать другой шаблон того же типа и запомнить его и т.д. Это делает поведение робота более естественным. Для примера приведем возможные реализации некоторых из выше охарактеризованных типов по-русски:

– Тип “Приветствие (Greeting)”: *Здравствуйте; Привет; Приветствую Вас в нашем институте; Добро пожаловать* и т.д.

– Тип “Предложение осмотра помещения (Tour-Proposal)”: *Я могу провести экскурсию по этажу; Я могу показать Вам отделы, которые находятся здесь, и рассказать о них; Я подведу Вас к отделам института, находящимся на этом этаже* и т.п.

– Тип “Этикетная самооценка (Evaluation)”: *Надеюсь, я справился со своей задачей; Надеюсь, Вам не было скучно со мной; Я очень старался, чтобы Вам было интересно* и пр.

– Тип “Заявление о намерении подождать (Promise_Wait)”: *Я подожду Вас здесь; Жду Вас здесь; Я буду ждать Вас; Я подожду Вашего возвращения* и т.п.

– Тип “Команда двигаться вслед (Imper_Move)”: *Следуйте за мной!; Пойдемте!; Вперед! В путь, друзья мои!* и т.п.

– Тип “Вопрос о принятии предложения о туре (Tour-Question)”: *Хотите?; Вам это интересно?; Вы принимаете мое предложение? И т.п.*

– Тип “Предложение выбрать первый объект осмотра из набора альтернатив (Question_Altern(Adm, AI))”: *С чего начнем — с офиса администрации или с лаборатории искусственного интеллекта?; Куда направимся сначала — к офису администрации или к лаборатории искусственного интеллекта?*

– Тип “Завершение диалога в случае отказа от предложения (Closure)”: *Ну что ж, в таком случае до свидания; Тогда до свидания; Извините за беспокойство и до новых встреч!* и т.п.

Каждый иллокутивно-независимый речевой акт и речевой блок робота в моделируемом типе диалога налагает определенное коммуникативное обязательство (commitment) на адресата при условии, что он в целом соблюдает аксиомы коммуникации Грайса [3], т.е. принимает общую цель и направление развития диалога, не пытается намеренно вводить адресата (робота) в заблуждение и т.д. Все речевые акты пользователя в данном типе диалога являются иллокутивно-зависимыми: в силу указанных выше ограничений не предполагается возможности для пользователя проявить инициативу в диалоге, например, задать роботу какой-либо вопрос, даже касающийся осматриваемого места. Пользователь должен реагировать в соответствии с иллокутивной целью речевых действий робота, а если он позволяет себе «вольности», то робот будет трактовать их как ошибку и пытаться «добиться» реакции требуемого типа (см. ниже). В дальнейшем планируется снять это ограничение с диалога.

5. Речевые акты пользователя. При вышеуказанных ограничениях репертуар речевых актов ожидаемых от пользователя, сводится к следующим РА, в сокращенных названиях которых присутствует префикс uSA (user Speech Act):

– Согласие uSA(Yes) или Отказ uSA(No) как реакции на предложение осмотра, входящее в блок IB и предложение воспользоваться второстепенным объектом, входящее в блок SB1.

– Ответ с выбранной альтернативой на вопрос робота, с какого объекта начать осмотр (см. выше Question_Altern(X), где $X \in \{Adm, AI\}$, в нашем конкретном случае):

– выбор офиса администрации uSA(Adm),

– выбор лаборатории искусственного интеллекта uSA(AI).

– Сообщение о возвращении на место после отлучки к второстепенному объекту uSA(Ready).

– Любой непредусмотренный системой в данном типе диалога РА пользователя трактуется как ошибка: Непредусмотренный (неинтерпретабельный) речевой акт пользователя uSA(Error).

Если система не может подвести очередной речевой акт пользователя ни под один из пяти возможных иллокутивных типов, то робот возвращается в фазу, непосредственно предшествовавшую появлению речевого акта, распознанному как uSA(Error), и повторяет тот речевой акт, после которого пользователь выдал речевой акт uSA(Error).

Если в отношении иллокутивного типа высказывания пользователь ограничен рамками коммуникативных обязательств, налагаемых на него речевыми актами робота, то в отношении способа оформления речевых актов разрешенного типа он относительно свободен. Анализатор речи опознает единицы, маркированные в словаре системы по пропозициональному содержанию или иллокутивной функции. Модуль управления диалогом припишет реплике пользователя в данной точке диалога тот иллокутивный тип из набора возможных, который соответствует маркеру распознанных единиц. Так, лексемы *офис* и *администрация* в реплике, поданной пользователем в ответ на речевой акт робота типа «Предложение выбрать первый объект осмотра из набора альтернатив Question_Altern(Adm, AI)» будет маркером иллокутивного типа «Выбор офиса администрации uSA(Adm)». В будущем можно реализовывать и более сложный анализ, потому что синтаксический анализатор позволяет работать с деревом разбора каждой фразы диалога.

6. Модель диалога. Модель диалога можно представить как сеть переходов робота из одного диалогического состояния в другое в зависимости от полученной им речевой или визуальной информации, истории диалога и намерений пользователя. Мы представляем эти состояния как правила, другая возможность состоит в построении конечного автомата [11].

В нашей модели диалогические состояния бывают следующих типов:

- состояние говорения (talk, talking state),
- состояние восприятия (perc., perception state),
- состояние движения (move, moving state),
- конечное состояние (final, final state).

Переход из одной фазы в другую осуществляется при помощи совершения роботом тех или иных действий (речевых и/или двигательных). Эти действия производятся в качестве реакции на речевой акт пользователя uSA(X), где X принимает значение из набора «пользовательских» иллокутивных типов РА, или визуальный акт воспри-

ятия роботом некоторого объекта Y , $\text{visual_act}(Y)$, где Y принимает значение из набора объектов, отмеченных на внутренней карте робота.

Соответственно, действие робота в зависимости от его текущего состояния задается парой «(входной речевой акт пользователя или визуальный акт) → выходной акт робота».

Таким образом, модель диалога может быть представлена в виде набора фаз и действий, переводящих диалог из одной фазы в другую. Можно использовать рекурсию, то есть вызывать одну и ту же функцию внутри самой себя, но во избежание заикливания возможная глубина рекурсии должна быть ограничена.

Далее мы приводим список правил, которые описывают нашу модель. Правила имеют вид:

состояние робота & «условие (если выполняется) : действие (то сделай)» → новое состояние робота

То есть, чтобы применить правило, робот должен быть в данном состоянии и должно выполниться условие. В этом случае робот будет выполнять соответствующее действие и перейдет в новое состояние. После каждого применения правила добавляются в историю данного диалога HD (history of dialogue). Иногда условие может отсутствовать: это значит, что условие истинно, то есть действие должно выполняться во всех случаях. Иногда не требуется действие, то есть нет необходимости, чтобы робот что-либо делал. Эти случаи обозначаются символом \emptyset . Заметим, что после применения правила, робот всегда переходит в новое состояние. Примеры правил приведены ниже:

talk ₁	&	\emptyset : IB	→ talk ₂
talk ₂	&	\emptyset : Tour-Question	→ perc ₁
perc ₁	&	uSA(Yes) : Question_Altern(Adm, AI)	→ perc ₂
perc ₁	&	uSA(No) : Closure	→ final
perc ₁	&	uSA(Error) : \emptyset	→ talk ₂
perc ₂	&	uSA(Adm) : MMB1(Adm)	→ move ₁
perc ₂	&	uSA(AI) : MMB1(AI)	→ move ₂
perc ₂	&	uSA(Error) : \emptyset	→ talk ₃
talk ₃	&	\emptyset : Question_Altern(Adm, AI)	→ perc ₂
move ₁	&	visual_act(Adm) : MMB2(Adm)	→ talk ₄
move ₁	&	visual_act(Cooler) : SMB(Cooler)	→ perc ₃
perc ₃	&	uSA(Yes) : SB	→ perc ₄
perc ₃	&	uSA(No) : MMB1(Adm)	→ move ₁

perc ₄	&	uSA(Ready) : MMB1(Adm)	→	move ₁
talk ₄	&	∅ : NB(Adm)	→	talk ₅
talk ₅	&	NB(AI) ∉ HD: MMB1(AI)	→	move ₂
talk ₅	&	NB(AI) ∈ HD:TB	→	final
move ₂	&	visual_act(AI) : MMB2(AI)	→	talk ₆
move ₂	&	visual_act(WC) : SMB(WC)	→	perc ₅
perc ₅	&	uSA(Yes) : SB	→	perc ₆
perc ₅	&	uSA(No) : MMB1(AI)	→	move ₂
perc ₆	&	uSA(Ready) : MMB1(AI)	→	move ₂
talk ₆	&	∅ : NB(AI)	→	talk ₇
talk ₇	&	NB(Adm) ∉ HD : MMB1(Adm)	→	talk ₇
talk ₇	&	NB(Adm) ∈ HD:TB	→	final

Первые два правила интерпретируются так: находясь в исходной фазе говорения робот совершает два инициативных, т.е. не вызванных внешними причинами, речевых действия, произнося сначала «Вводный блок (IB)», а затем вопрос о желании совершить тур Tour-Question, после чего переходит в фазу восприятия (входной информации) perc₁.

7. Примеры диалогов. Приведем пример диалога, соответствующего данной модели. *P* это робот, *П* — пользователь, после каждой реплики приводится ее описание в терминах предлагаемой модели. Заметим, что одно и то же правило в этих примерах может начинаться до реплики или продолжаться после реплики, что показано многоточием и открывающими и закрывающими кавычками на разных строках.

P: Здравствуйте. Я робот-гид. Я могу провести для Вас экскурсию по этажу. Хотите?

talk₁ & «∅ : IB» → talk₂;

talk₂ & «∅ : Tour-Question» → perc₁

П: Да, спасибо.

perc₁ & «uSA(Yes)...

P: С какого отдела начнем — с офиса администрации или с лаборатории искусственного интеллекта?

... : Question_Altern(Adm, AI)» → perc₂

П: Конечно, с лаборатории.

perc₂ & «uSA(AI)...

P: Следуйте за мной!

... : MMB1(AI)» → move₂

Двигаются в сторону лаборатории. Поравнявшись с туалетом, робот останавливается и произносит:

P: Вот здесь у нас туалет. Не желаете воспользоваться?

$move_2 \& \langle \text{visual_act(WC) : SMB(WC)} \rangle \rightarrow perc_5$

П: Спасибо, нет.

$perc_5 \& \langle uSA(No) \dots$

P: Пойдем.

$\dots : MMB1(AI) \rangle \rightarrow move_2$

Двигаются далее в сторону лаборатории. Дойдя до цели робот останавливается и произносит:

P: Мы у цели.

$move_2 \& \langle \text{visual_act(AI) : MMB2(AI)} \rangle \rightarrow talk_6$

Робот начинает рассказ или показывает презентацию о лаборатории. Закончив и поменяв свое состояние ($talk_6 \rightarrow talk_7$), он произносит:

P: Следуйте за мной!

$talk_7 \& \langle \emptyset : MMB1(Adm) \rangle \rightarrow move_1$

На пути ко второму объекту показа — офису, поравнявшись с кулером с питьевой водой робот останавливается и произносит:

P: Обратите внимание на кулер. Хотите выпить воды?

$move_1 \& \langle \text{visual_act(Cooler) : SMB(Cooler)} \rangle \rightarrow perc_3$

П: Хорошо бы.

$perc_3 \& \langle uSA(Yes) \dots$

P: Жду вас здесь. Когда вернетесь, скажите мне об этом.

$\dots : SB \rangle \rightarrow perc_4$

Пользователь отходит к кулеру, пьет воду, возвращается и произносит:

П: Я вернулся. Я готов.

$perc_4 \& \langle uSA(Ready) \dots$

P: Пойдемте.

$\dots : MMB1(Adm) \rangle \rightarrow move_1$

Двигаются в сторону офиса. Дойдя до него, робот останавливается и произносит:

P: Вот и офис.

$move_1 \& \langle \text{visual_act(Adm) : MMB2(Adm)} \rangle \rightarrow talk_4$

После этого идет рассказ или презентация о деятельности администрации. Завершив нарративный блок и поменяв свое состояние ($talk_4 \rightarrow talk_5$), робот произносит:

P: Тур по этажу окончен. Надеюсь, я не слишком Вас утомил. До свидания.

$talk_5 \& \langle \emptyset : TB \rangle \rightarrow final$

Приведем для иллюстрации еще один фрагмента диалога:

...
P: *Какой отдел посетим в первую очередь — офис администрации или лабораторию искусственного интеллекта?*

[perc₁ & «uSA(Yes) :] Question_Altern(Adm, AI)» → perc₂

П: *Не знаю.*

perc₂ & «uSA(Error) : Ø» → talk₃

P: *С чего начнем: с офиса администрации или лаборатории искусственного интеллекта?*

talk₃ & «Ø : Question_Altern(Adm, AI)» → perc₂

П: *Ну, давайте с офиса.*

perc₂ & «uSA(Adm)...

...
Заметно, что реплики робота не во всех точках диалога звучат естественно. Чтобы сделать их более естественными, предстоит усовершенствовать модель диалога, например, перейдя от пропозициональной модели к предикатной, т.е. учесть не только иллюкутивную семантику базовых речевых актов, но и предикатно-аргументную структуру реплик, а также разработать шаблоны межфразовых связей.

8. Заключение. В данной статье мы представили описание модуля управления диалогом при общении с подвижным роботом, выполняющим функции гида. Модуль включает в себя модель диалога, описание речевых актов и речевых блоков, из которых можно строить каждый компонент модели, и готовые шаблоны реплик, соответствующих всем речевым актам, имеющимся в модели. Вся модель за исключением шаблонов является независимой от языка.

Хотя в статье мы описали только ограниченное применение модели, предложенная модель по своей архитектуре универсальна и предоставляет средства для формального моделирования диалога с автономным подвижным роботом на разном языковом материале, поскольку инвентарь речевых актов и коммуникативных стратегий говорящего в отличие от лингвоспецифических ограничений на синтаксическую структуру является общим для всех естественных языков.

Литература

1. Dialog with Robots // AAAI 2010. Fall Symposium. November 2010, Arlington VA. URL: http://hci.cs.wisc.edu/aaai10/?page_id=169.
2. Green S.A., Billinghurst M., Chen X. Q., Chase G.J. Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design // International Journal of Advanced Robotic Systems. 2007. vol. 5. pp. 1–18.
3. Grice P. Logic and conversation // Syntax and Semantics, 3: Speech Acts, ed. P. Cole & J. Morgan, New York: Academic Press. 1975. (Reprinted in: Studies in the Way of Words, ed. H. P. Grice, Cambridge, MA: Harvard University Press. 1989. pp. 22–40).

4. *Kruijff G.J.M., Zender H., Jensfelt P., Christensen H.I.* Clarification dialogues in human-augmented mapping // Human Robot Interaction'06, Salt Lake City, Utah, USA, 2006.
5. *Kruijff G.J.M., Lison P., Benjamin T., Jacobsson H., Zender H., Kruijff-Korbayova I., Hawes N.* Situated dialogue processing for human-robot interaction // Cognitive Systems. 2010. pp. 311–364.
6. *Kuzuoka H., Yamazaki K., Yamazaki A., Kosaka J., Suga Y., Heath C.* Dual ecologies of robot as communication media: Thoughts on coordinating orientations and projectability // CHI'04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2004. pp 183–190.
7. *Lemaignan S., Ros R., Alami R.* Dialogue in situated environments: A symbolic approach to perspective-aware grounding, clarification and reasoning for robot // Proc. Robotics, Science and Systems, Grounding Human-Robot Dialog for Spatial Tasks workshop, 2011.
8. *Marge M., Pappu A., Frisch B., Harris Th.K., Rudnický A.* Exploring Spoken Dialog Interaction in Human-Robot Teams // Proceedings of Robots, Games, and Research: Success stories in USA, RSim IROS Workshop, St. Louis, MO, USA, 2009.
9. *Padró L., Collado M., Reese S., Lloberes M., Castellón I.* FreeLing 2.1: Five Years of Open-Source Language Processing Tools // Proceedings of 7th Language Resources and Evaluation Conference (LREC 2010), ELRA, La Valletta, Malta, 2010.
10. *Pakray P., Barman U., Bandyopadhyay S., Gelbukh A.* A Statistics-Based Semantic Textual Entailment System // MICAI 2011, Lecture Notes in Artificial Intelligence 7094, Springer, 2011, pp. 267–276.
11. *Pineda L.A., Estrada V.M., Coria S.R., Allen J.F.* The obligations and common ground structure of practical dialogues // Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial. 2007. vol. 11(36). pp. 9–17.
12. *Salton G., Wong A., Yang C.S.* A vector space model for automatic indexing // Communications of the ACM. 1975. vol. 18(11). pp. 613–620.
13. *Searle J.* Indirect speech acts // Syntax and Semantics, 3: Speech Acts, ed. P. Cole & J. L. Morgan, New York: Academic Press. 1975. pp. 59–82. (Reprinted in: Pragmatics: A Reader, ed. S. Davis, Oxford: Oxford University Press, 1991. pp. 265–277).
14. *Sidorov G.* Non-linear construction of n-grams in computational linguistics: syntactic, filtered, and generalized n-grams // Sociedad Mexicana de Inteligencia Artificial. 2013. 166 p.
15. *Sidorov G.* Development of an application for dialog in natural language with a mobile robot [Desarrollo de una aplicación para el diálogo en lenguaje natural con un robot móvil (in Spanish)] // Research in computing science. 2013. vol. 62. pp. 95–108.
16. *Sidorov G., Velasquez F., Stamatatos E., Gelbukh A., Chanona-Hernández L.* Syntactic N-grams as Machine Learning Features for Natural Language Processing // Expert Systems with Applications. 2014. vol. 41(3). pp. 853–860.
17. *Sisbot E.A., Ros R., Alami R.* Situation assessment for human-robot interaction // Proc. of 20th IEEE International Symposium in Robot and Human Interactive Communication. 2011.
18. *Skubic M., Perzanowski D., Schultz A., Adams W.* Using Spatial Language in a Human-Robot Dialog // IEEE International Conference on Robotics and Automation, Washington, D.C., 2002.
19. *Svanaes D., Seland G.* Putting the users center stage: role playing and low-fi prototyping enable end users to design mobile systems // Proceedings CHI 2004, Vienna, Austria, Association for Computing Machinery. 2004. pp. 479-486.
20. *Topp E.A., Huttenrauch H., Christensen H.I., Severinson E. K.* Bringing Together Human and Robotic Environment Representations—A Pilot Study // Proceedings of the IEEE/RSJ Conference on Intelligent Robots and Systems (IROS), Beijing, China. 2006.

21. *Wilks Y., Webb N., Setzer A., Hepple M., Catizone R.* Machine learning approaches to human dialogue modelling // J.C.J. van Kuppevelt et al. (eds.), *Advances in Natural Multimodal Dialogue Systems*. 2005. pp. 355–370.
22. *Аркадьев П.М., Сердобольская Н.В., Циммерлинг А.В.* Проект типологической базы по синтаксическим ограничениям на передвижения // *Компьютерная лингвистика и интеллектуальные технологии, по материалам международной конференции «Диалог 2010»*. М.: РГГУ. 2010. Вып. 9(16). С. 437–441.
23. *Баранов А.Н., Крейдлин Г.Е.* Иллокутивное вынуждение в структуре диалога // *Вопросы языкознания*. 1992. №2. С. 84–99.
24. *Кобозева И.М., Лауфер Н.И., Сабурова И.Г.* Моделирование общения в человеко-машинных системах // *Лингвистическое обеспечение информационных систем*. М.: Изд-во ИНИОН АН СССР, 1987.
25. *Прищепа М.В.* Разработка профиля пользователя с учетом психологических аспектов взаимодействия человека с информационным мобильным роботом // *Труды СПИИРАН*. 2012. Вып. 2(21). С. 56–70.

References

1. *Dialog with Robots*. AAAI 2010. Fall Symposium. November 2010, Arlington VA. Available at: http://hci.cs.wisc.edu/aaai10/?page_id=169.
2. Green S.A., Billinghurst M., Chen X. Q., Chase G.J. *Human-Robot Collaboration: A Literature Review and Augmented Reality Approach in Design*. *International Journal of Advanced Robotic Systems*. 2007. vol. 5. pp. 1–18.
3. Grice P. *Logic and conversation. Syntax and Semantics, 3: Speech Acts*, ed. P. Cole & J. Morgan, New York: Academic Press. 1975 (Reprinted in: *Studies in the Way of Words*, ed. H. P. Grice, Cambridge, MA: Harvard University Press. 1989. pp. 22–40).
4. Kruijff G.J.M., Zender H., Jensfelt P., Christensen H.I. *Clarification dialogues in human-augmented mapping*. *Human Robot Interaction'06*, Salt Lake City, Utah, USA, 2006.
5. Kruijff G.J.M., Lison P., Benjamin T., Jacobsson H., Zender H., Kruijff-Korbayova I., Hawes N. *Situated dialogue processing for human-robot interaction*. *Cognitive Systems*. 2010. pp. 311–364.
6. Kuzuoka H., Yamazaki K., Yamazaki A., Kosaka J., Suga Y., Heath C. *Dual ecologies of robot as communication media: Thoughts on coordinating orientations and projectability*. *CHI'04 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2004, pp 183–190.
7. Lemaignan S., Ros R., Alami R. *Dialogue in situated environments: A symbolic approach to perspective-aware grounding, clarification and reasoning for robot*. *Proc. Robotics, Science and Systems, Grounding Human-Robot Dialog for Spatial Tasks workshop*, 2011.
8. Marge M., Pappu A., Frisch B., Harris Th.K., Rudnický A. *Exploring Spoken Dialog Interaction in Human-Robot Teams*. *Proceedings of Robots, Games, and Research: Success stories in USA, RSim IROS Workshop*, St. Louis, MO, USA. 2009.
9. Padró L., Collado M., Reese S., Lloberes M., Castellón I. *FreeLing 2.1: Five Years of Open-Source Language Processing Tools*. *Proceedings of 7th Language Resources and Evaluation Conference (LREC 2010)*, ELRA, La Valletta, Malta, 2010.
10. Pakray P., Barman U., Bandyopadhyay S., Gelbukh A. *A Statistics-Based Semantic Textual Entailment System*. *MICAI 2011, Lecture Notes in Artificial Intelligence 7094*, Springer. 2011. pp. 267–276.
11. Pineda L.A., Estrada V.M., Coria S.R., Allen J.F. *The obligations and common ground structure of practical dialogues*. *Inteligencia Artificial, Revista Iberoamericana de Inteligencia Artificial*. 2007. vol. 11(36). pp. 9–17.
12. Salton G., Wong A., Yang C.S. *A vector space model for automatic indexing* // *Communications of the ACM*. 1975. vol. 18(11). pp. 613–620.

13. Searle J. Indirect speech acts. *Syntax and Semantics, 3:Speech Acts*, ed. P. Cole & J. L. Morgan, New York: Academic Press. 1975. pp. 59–82. (Reprinted in: *Pragmatics: A Reader*, ed. S. Davis, Oxford: Oxford University Press. 1991. pp. 265–277).
14. Sidorov G. Non-linear construction of n-grams in computational linguistics: syntactic, filtered, and generalized n-grams. 2013. 166 p.
15. Sidorov G. Development of an application for dialog in natural language with a mobile robot [Desarrollo de una aplicación para el diálogo en lenguaje natural con un robot móvil] *Research in computing science*. 2013. vol. 62. pp. 95–108. (In Spanish).
16. Sidorov G., Velasquez F., Stamatatos E., Gelbukh A., Chanona-Hernández L. Syntactic N-grams as Machine Learning Features for Natural Language Processing. *Expert Systems with Applications*. 2014. vol. 41(3). pp. 853–860.
17. Sisbot E.A., Ros R., Alami R. Situation assessment for human-robot interaction. *Proc. of 20th IEEE International Symposium in Robot and Human Interactive Communication*. 2011.
18. Skubic M., Perzanowski D., Schultz A., Adams W. Using Spatial Language in a Human-Robot Dialog. *IEEE International Conference on Robotics and Automation*, Washington, D.C., 2002.
19. Svanaes D., Seland G. Putting the users center stage: role playing and low-fi prototyping enable end users to design mobile systems. *Proceedings CHI 2004*, Vienna, Austria, Association for Computing Machinery. 2004. pp. 479–486.
20. Topp E.A., Huttenrauch H., Christensen H.I., Severinson E. K. Bringing Together Human and Robotic Environment Representations—A Pilot Study. *Proceedings of the IEEE/RJSJ Conference on Intelligent Robots and Systems (IROS)*, Beijing, China. 2006.
21. Wilks Y., Webb N., Setzer A., Hepple M., Catizone R. Machine learning approaches to human dialogue modeling. Edited by J.C.J. van Kuppevelt et al., *Advances in Natural Multimodal Dialogue Systems*. 2005. pp. 355–370.
22. Arkadiev P.M., Serdobolskaya N.V., Zimmerling A.V. [Project of a typological database of syntactic constraints on movement]. *Komp'yuternaja lingvistika i intellektual'nye tehnologii, po materialam mezhdunarodnoj konferencii «Dialog 2010»* [Computational linguistics and intellectual technologies, proceedings of the international conference “Dialogue. 2010”]. Moscow. 2010. vol. 9(16). pp. 437–441. (In Russ.).
23. Baranov A.N., Kreidlin G.E. [Illocutive forcement in dialogue structure]. *Voprosy jazykoznanija – Problems of Linguistics*. 1992. vol. 2. pp. 84–99. (In Russ.).
24. Kobozeva I.M., Laufer N.I., Saburova I.G. [Modelling conversation in man-machine systems] *Lingvisticheskoe obespechenie informacionnyh sistem – Linguistic support of information systems*. M.: INION of Academy of Sciences of USSR, 1987. (In Russ.).
25. Prischepa M.V. [Development of the user profile based on the psychological aspects of human interaction with the informational mobile robot]. *Trudy SPIIRAN – SPIIRAS proceedings*. 2012. vol. 2(21). pp. 56–70. (In Russ.).

Кобоzeва Ирина Михайловна — д-р филол. наук, доцент, профессор кафедры теоретической и прикладной лингвистики филологического факультета МГУ имени М. В. Ломоносова. Область научных интересов: лингвистическая семантика и прагматика, автоматический семантический анализ текста на ЕЯ. Число научных публикаций — 150. kobozeva@list.ru; МГУ, Ленинские горы, г. Москва, 119991, РФ; p.t. +7(495)939-2601, факс +7(495)939-5596.

Kobozeva Irina Mikhajlovna — Ph.D., Dr. Sci., Assoc. professor, professor of the Dept. of Theoretical and Applied Linguistics, Philological Faculty of Moscow State Lomonosov University. Research interests: linguistic semantics and pragmatics, natural language processing.

Number of publications: 150. kobozeva@list.ru; Moscow State Lomonosov University, Leninskije gory, Moscow, 119991, Russia; office phone +7(495)939-2601, fax +7(495)939-5596.

Сидоров Григорий Олегович — к-т филол. наук, профессор лаборатории естественно-го языка и обработки текста Центра Компьютерных Исследований Национального Политехнического института (Мексика), национальный исследователь Мексики (SNI 3), главный редактор научного журнала “Computación y sistemas” (Scopus, ISI). Область научных интересов: компьютерная лингвистика, автоматический анализ текста. Число научных публикаций — 190. sidorov@cic.ipn.mx; CIC-IPN, Av. Juan de Dios Bátiz, s/n, Zacatenco, 07738, México DF, Mexico; p.t. +52(55)5729-6000 доб. 56518.

Sidorov Grigori — Ph. D., professor, research professor of Natural Language and Text Processing Laboratory, Center for Computing Research, Instituto Politécnico Nacional, Mexico City, Mexico; National researcher of Mexico (SNI 3). Editor-in-Chief of the research journal “Computación y sistemas” (Scopus, ISI). Research interests: computational linguistics, natural language processing. Number of publications: 190. sidorov@cic.ipn.mx; CIC-IPN, Av. Juan de Dios Bátiz, s/n, Zacatenco, 07738, México DF, Mexico; office phone +52(55)5729-6000 ext. 56518.

Циммерлинг Антон Владимирович — д-р филол. наук, профессор кафедры русского языка Московского государственного гуманитарного университета имени М. А. Шолохова (МГГУ им. М. А. Шолохова), заведующий лабораторией общей и компьютерной лингвистики МГГУ им. М. А. Шолохова. Область научных интересов: лингвистическая типология, теория грамматики, компьютерная лингвистика. Число научных публикаций — 185. fagraey64@hotmail.com. МГГУ им. М. А. Шолохова, ул. Верхняя Радищевская 16-18, Москва 109240, p.t. +7(495)672-0490.

Zimmerling Anton — Ph.D., Dr. Sci., professor of the Russian department, Philological Faculty of Sholokhov Moscow State University for the Humanities (SMSUH), head of the Laboratory of General and Computational Linguistics SMSUH. Research interests: linguistic typology, grammar theory, computational linguistics. Number of publications: — 185. fagraey64@hotmail.com. SMSUH, Verkhnyaya Radishchevskaya str. 16-18, Moscow 109240 office phone +7(495)672-0490.

Поддержка исследований. Работа выполнена при финансовой поддержке правительства Мексики (SNI), Национального Политехнического Института (Мексика) (SIP 20120418, 20131441, 20144274, COFAA), Правительства города Мехико (ICYT-DF PICCO10-120), Европейской комиссии FP7-PEOPLE-2010-IRSES грант 269180, и Министерства образования и науки Российской Федерации, федеральный проект (госзадание) 2685 «Параметрическое описание грамматических систем».

Acknowledgements. This research is done under partial support of Mexican government (SNI), Instituto Politécnico Nacional, Mexico (projects SIP 20120418, 20131441, 20144274, COFAA), Mexico City government (ICYT-DF project PICCO10-120), FP7-PEOPLE-2010-IRSES: Web Information Quality - Evaluation Initiative (WIQ-EI) European Commission project 269180 and Ministry of education and science of Russian Federation, federal project 2685 “Parametric description of grammar systems”.

РЕФЕРАТ

Кобозева И.М., Сидоров Г.О., Циммерлинг А.В. **Модуль управления диалогом в системе общения пользователя с подвижным роботом-гидом.**

В статье представлен алгоритмический подход к моделированию диалога на естественном языке с подвижным роботом, выполняющим функцию гида. Модель управления диалогом представлена как сеть переходов системы из одного состояния в другое в зависимости от двух факторов: коммуникативного и визуального. Описываются детали модуля управления диалогом. Система диалога с подвижным роботом Pioneer 3DX включает модуль распознавания речи Dragon Naturally Speaking и модуль синтаксического анализа Freeling. Система протестирована для испанского языка, но по своей идеологии не зависит от языкового материала и может быть настроена на другие естественные языки, включая русский. Чем больше функций у робота, тем больше репертуар коммуникативных актов и тем важнее роль интенционального модуля (модели рационального поведения, энциклопедическая информация). В диалоговую модель описываемого робота-гида включены три вида актов — речевые акты, речевые блоки и мультимодальные блоки. Речевые блоки определяются как последовательности речевых актов одного говорящего, связанные единством замысла и предполагающих реакцию адресата на блок как целое. Выделяются пограничные блоки, нарративные блоки и второстепенные блоки. Пограничные блоки запускаются в начальной и конечной фазе диалога с роботом. Нарративные блоки являются текстами, содержащими информацию об объектах показа. Второстепенные блоки произносятся роботом в ситуации переключения с основной цели показа на второстепенную. Мультимодальные блоки — фиксированные последовательности речевых актов и неречевых действий. Модель диалога с роботом, в целом, следует аксиомам коммуникации Грайса. Среди действий робота есть как иллокутивно-зависимые, так и иллокутивно-независимые, т.е. запускаемые по инициативе самого робота. Все действия человека в данном типе диалога являются иллокутивно-зависимыми от действий робота. Любой непредусмотренный в данном типе диалога речевой акт пользователя интерпретируется как ошибка, в этом случае робот возвращается в фазу диалога, непосредственно предшествующему тому речевому акту пользователя, который был распознан как ошибка, и повторяет свой речевой акт.

SUMMARY

Kobozeva I.M., Sidorov G., Zimmerling A. **Module for dialog management in the interaction system between user and mobile robotic guide.**

This paper presents a dialogue management module for a mobile robot acting as a guide. The model of the dialog management is presented as a network of transitions between states depending on two factors: visual and communicative. Details of the dialogue management module are presented. We adopt algorithmic approach to dialogue modeling. The dialogue system with a mobile robot Pioneer 3DX contains speech recognition module Dragon Naturally Speaking and syntactic parser Freeling. The system is tested for Spanish, but its architecture is language-independent and can be adapted to other natural languages, including Russian. The more functions a robot has, the wider the inventory of communicative acts and the more salient the intentional module (models of rational behavior, encyclopedic information) is. The inventory of the described robot-guide includes three types of acts: speech acts, speech blocks, and multimodal blocks. Speech blocks are defined as sequences of speech acts by the same speaker sharing the same communicative intention and implying that the addressee reacts to the block as a whole unit. We distinguish Borderline Blocks, Narrative Blocks and Secondary Blocks. Secondary blocks are texts containing information about the demonstration objects. Secondary blocks are pronounced by the robot in situations where shifting from the main demonstration task to secondary tasks takes place. Multimodal blocks are fixed sequences of speech acts and nonverbal actions. The described dialogue model with a robot-guide largely conforms to Grice's communication axioms. The robot's actions both include illocutionary-dependent and illocutionary-independent actions, i.e. actions launched by the robot itself. All actions of a human user in this type of dialogue are illocutionary-dependent from the robot's actions. Each unprovided speech act of the user is interpreted as error; in this case the robot returns back to the dialogue phase that immediately preceded the user's speech act recognized as error and repeats its speech act.