

МОДЕЛЬ ЭМОЦИОНАЛЬНОГО РЕЧЕВОГО ПОВЕДЕНИЯ ДЛЯ ВИРТУАЛЬНОГО АГЕНТА РОЛЕВОЙ КОМПЬЮТЕРНОЙ ИГРЫ

MODEL OF EMOTIONAL SPEECH BEHAVIOUR FOR A VIRTUAL AGENT OF A COMPUTER ROLE-PLAYING GAME

A. A. Kotov (kotov@harpia.ru)

В ролевой компьютерной игре играющий решает последовательные задачи (головоломки), управляя действиями некоторого виртуального агента (героя игры). При этом игрок (и его виртуальный агент) могут испытывать успехи или неудачи. В данном докладе мы рассматриваем теоретическую модель, которая бы позволила наделить виртуального агента «речью» и описать возможные высказывания агента в различных ситуациях. Мы рассмотрим как стратегию выбора высказывания из базы данных, так и перспективы семантического синтеза высказываний.

В последние годы в кибернетике и в областях разработки человеко-машинных интерфейсов все большее внимание привлекают задачи моделирования эмоционального поведения и распознавания эмоций пользователя. Предполагается, что компьютер или бытовой робот при взаимодействии с пользователем, с одной стороны, должен распознавать эмоциональное состояние пользователя (и соответствующим образом реагировать на его слова), а с другой стороны, должен сам имитировать некоторые эмоции с помощью жестов, просодических характеристик высказывания или с помощью выражения «лица» (средствами своего интерфейса). Соответственно, современные прикладные исследования в области эмоциональных интерфейсов сосредотачиваются на задачах автоматического синтеза или анализа просодии высказывания [Inanoglu, Young, 2005], создании виртуальных моделей лица с подвижной мимикой (так называемая, область “talking heads”) или фигуры человека [Becker, Kopp et al., 2004; Ochs, Niewiadomski et al., 2005], системах распознавания выражений лиц и т.д. Моделирование эмоциональных процессов также привлекает внимание при исследовании мультиагентных сообществ и даже в философии сознания [Sloman, Chrisley, 2003]. На этом фоне исследования семантики эмоциональных текстов достаточно ограничены – это вызвано как общими трудностями создания автоматического семантического анализа, так и многообразием речевых средств, служащих для передачи эмоций.

Вместе с тем, обработка семантики текста является важнейшей проблемой при создании эмоциональных интерфейсов: адекватный эмоциональный компьютерный агент при обработке команды пользователя должен уметь не только произнести некоторое высказывание с определённой интонацией (выражая заданную эмоцию), но и уметь выбрать соответствующую эмоцию для ответной реакции, а также (что самое важное) – сформировать само ответное высказывание или выбрать его из базы данных.

На прошлой конференции мы представили теоретическую модель для структурного описания высказываний в эмоциональных ситуациях [Котов, 2004]. В данном докладе мы намерены показать перспективы применения этой модели для создания виртуального агента ролевой компьютерной игры. Такой агент должен обрабатывать команды пользователя (например, подаваемые с помощью «мыши»), выполнять некоторые действия и при этом генерировать высказывания (реакции) на естественном языке. В своих ответах агент сможет проявлять некоторый заданный «темперамент» или «настроение» (может быть «оптимистом» или «пессимистом»), находиться в грустном или веселом настроении). Агент может также обращать свои высказывания к другим виртуальным агентам, если они есть в игре.

Создание «говорящего» компьютерного агента позволит обогатить виртуальную среду компьютерной игры. Вместе с тем, в этом докладе мы не затрагиваем вопрос о том, сделает ли «говорящий» компьютерный агент игру более привлекательной для пользователя, а также, какой тип речевого поведения будет наиболее приятен играющему. Мы постараемся рассмотреть максимальное число возможных способов реагирования компьютерного агента, учитывая, что в реальной игре возможное количество способов реагирования может быть сокращено.

Лингвистическая модель

Используемая лингвистическая модель основывается на инвентаре «сценариев» - процедур обработки текста (точнее, семантических графов, построенных при анализе текста). Среди сценариев выделяются *доминантные сценарии (д-сценарии)*, заданные списком; д-сценарии служат для анализа и синтеза эмоциональной речи.¹ При описании речевых стратегий в конкретной ситуации может быть выделен набор *рациональных сценариев (р-сценариев)*; р-сценарии описывают возможные не эмоциональные (рациональные) высказывания и речевые реакции. Д-сценарии были предложены для описания воздействия текстов СМИ [Котов, 2003] и далее использованы для анализа эмоциональной речи и эмоционального речевого поведения. Описание каждого д-сценария включает набор семантических признаков, отличающих эмоциональные высказывания; например, признак 'интенсивности предиката' является различительным для высказываний *Он вставляет ключ в скважину* и *Ну и куда ты суёшь этот ключ?* (см. также [Апресян, 2003; Гловинская, 2004]).

Экспериментальное исследование и сегментация ситуации

В экспериментальном исследовании [Kotov, 2005] мы попросили около 500 испытуемых предложить возможные ответы для ряда эмоциональных ситуаций – когда человеку причиняют неудобства (например, его забрызгала проезжающая машина) или когда сам человек причиняет неудобства другим (например, разбивает в гостях вазу). Нами использовался ограниченный набор ситуаций из теста Розенцвейга [Практич. психология..., 2003]; мы просили испытуемых также указать, что бы они подумали в этой ситуации и что бы они сделали/сказали, если бы не сдерживались. Высказывания испытуемых были организованы в базу данных и размечены с помощью д/р-сценариев.

Полученная база данных может использоваться для формирования набора возможных высказываний виртуального агента: если агент попадает в такую же ситуацию, как ситуация теста Розенцвейга, мы можем использовать имеющиеся в базе высказывания; если ситуация в игре отличается, мы должны будем модифицировать имеющиеся высказывания из базы.

Вместе с тестом Розенцвейга мы предлагали испытуемым другие психологические тесты (разные – для разных подгрупп испытуемых: тест Лири, тест Личко, тест «большая пятёрка») и устанавливали корреляции между ответами испытуемых и их личностными характеристиками. Полученные соответствия могут использоваться для имитации определённых личностных характеристик при выборе агентами возможного высказывания.

При систематизации высказываний, полученных в ходе эксперимента, мы разделили высказывания на классы в зависимости от того, какой компонент ситуации приводит к активизации д- или р-сценария. Так, если мы активизируем д-сценарий НЕАДЕКВ ('некто ведёт себя неадекватно и это вызывает нашу эмоциональную реакцию'), то исходной ситуаций может быть 'я веду себя неадекватно', 'я доставляю неудобство другим – это неправильно', 'другие не понимают меня / того, что я хотел сделать', 'другой ведёт себя неадекватно', 'другой доставляет неудобства мне' и т. д. При разметке высказываний мы используем общую схему взаимодействия между субъектом речи (агентом) и некоторым контрагентом (Схема 1.); в этой схеме выделяются сегменты, каждый из которых может быть основанием для активизации д-сценария и для синтеза последующей речевой реакции.

Возможные игровые сценарии и речевые реакции

В ролевой компьютерной игре (игре типа "quest") играющий руководит действиями некоторого агента в виртуальном пространстве. Играющий подаёт агенту команды «мышью», и агент, следуя этим командам, может передвигаться по экрану (ходить, прыгать), осуществлять некоторые действия с объектами мира (открывать дверь, включать свет), брать некоторые предметы и применять взятые предметы к объектам мира (например, отвинчивать винт отвёрткой). Руководя действиями агента, игрок должен решить ряд головоломок, чтобы перейти к следующей стадии игры. При решении головоломки игрок может испытывать успех или неудачу (ключ может подойти или не подойти к двери), причём это может выглядеть как успех или неудача виртуального агента, соответственно, агент может «переживать» воодушевление или фрустрацию и обозначать их с помощью речевых реакций.

Рассмотрим следующий игровой сценарий. Агент находится под платформой, на которую ему нужно взобраться. Платформа имеет изогнутую форму: с некоторых точек запрыгнуть на неё нельзя – агент прыгает, но не достаёт до верха платформы и падает обратно (это – фрустрация №1). Над платформой летает птица, которая может сбросить агента, если он уже взобрался на платформу (фрустрация №2). Агент может использовать приведённую схему взаимодействия (Схема 1.) по отношению к (а) пользователю, (б) враждебному контрагенту – птице и (в) дружественному контрагенту. В каждой из этих схем, теоретически, он может выделить указанный

¹ Инвентарь д-сценариев доступен по адресу: <http://www.harpia.ru/d-scripts.html>

набор из 8 сегментов, и для каждого сегмента активизировать д- или р-сценарии (на практике, многие из теоретически возможных вариантов не будут реализованы). Возможный набор высказываний для схемы взаимодействия с пользователем и фрустрации №1 может выглядеть следующим образом:

S1*НЕАДЕКВ²: *Я так криво прыгаю!* **S1*ТЕЩТН:** *У меня никогда не получится туда запрыгнуть!*
S1*СУБЪЕКТ: *Как увижу гору, сразу надо на неё лезть!* **S1*р-сценарии:** *Как бы туда запрыгнуть? Может, ты подумаешь, как туда лучше запрыгнуть? Извините, у меня не получилось! Мне нужно быть сосредоточеннее! Я же не специально упал! Я всегда спокоен! Тебе не стоит из-за этого переживать, это – ерунда! Еще получится! Это даже лучше – посидим внизу! Тут просто высоко, вот я и не достал!*

S3*ВИНА: *Так жалко, что у тебя из-за меня ничего не получается!* **S1*р-сценарии:** *Мне бы не хотелось, чтобы ты из-за меня сильно переживал! Извини, я тебя, наверно, напугал! Я не хотел тебя расстроить!*

S5*СУБЪЕКТ: *Ты меня специально заставляешь прыгать/гоняешь туда-сюда?* **S5*НЕПОСЛЕД:** *Ну как так можно?! То в одну сторону иди, то совсем в противоположную!* **S5*ОПАСН:** *Тебе вообще меня не жалко? Ты меня вообще не ценишь?* **S5*р-сценарии:** *Да ладно, ты же не специально меня уронил! Пошёл бы ты отдохнул! Тебе нужно быть со мной аккуратнее! д/р-комбинированные: Я надеюсь, теперь-то ты убедился, как криво ты играешь! Очень жаль, но я уже шишку набил!*

S6*НЕАДЕКВ: *Ты соображаешь куда мышкой жмешь?* **S6*НЕАДЕКВ:** *Ты думаешь, я смогу туда взобраться?* **S6*СУБЪЕКТ:** *Ты думаешь, я туда побегу, куда ты нажал?* **S6*БЕЗДЕЙСТВ:** *Давай, делай уже что-нибудь! Тебе что, трудно? Нам играть надо!*

S7*ФРУСТР: *Чёрт, жалко! Чёрт, опять куда-то прыгать надо!* **S1*р-сценарии:** *Ну ладно, будем прыгать! В конце-концов, это не так страшно!*

S8*НЕАДЕКВ: *Я совершенно не могу понять, что ты хочешь, чтоб я сделал!*

Агент может применять приведённую схему взаимодействия к дружественным контрагентам – *Ну куда ты попёрся? Тебе что, сложно мне помочь?* – или враждебным контрагентам (здесь – к ‘птице’) – *Что это за уроды тут летают? или Почему я всё время на всех натыкаюсь?* Агент может применять данную схему взаимодействия по отношению к одному контрагенту и при этом обращаться к другому, например, он может комментировать действия ‘птицы’ при обращении к пользователю или действия пользователя по отношению к дружественному контрагенту (*Ты посмотри, как он вообще играет?! Что он вообще делает!*).

Приведённые высказывания непосредственно взяты из базы данных или получены путём модификации высказываний из базы данных. В экспериментальном исследовании нами были собраны ответы на вопросы ‘что бы Вы подумали?’ и ‘что бы Вы сделали, если бы не сдерживались?’, что позволяет нам в игре приписывать агенту некоторые суждения в виде «мыслей» - эти суждения могут появляться в выводах или проговариваться агентом «себе под нос».

В игру может быть включен дополнительный персонаж (агент #2), который будет озвучивать команды пользователя, подаваемые мышью, а также произносить собственные команды или высказывания. Если мы командуем агенту #1 переместиться на новое место, то агент #2 при этом будет говорить *Иди вон туда!* (такие высказывания агента #2 имеет смысл включать в игру только если агент #1 на них будет отвечать). В одной из рассмотренных игровых ситуаций пользователь отдаёт агенту #1 команду включить свет (нажать кнопку на выключателе). Команда озвучивается агентом #2, который говорит *Включи свет!*. Агент #1 нажимает кнопку на выключателе, но в этот момент выключатель взрывается и из него летят искры.

Сама фраза агента #2 *Включи свет!* уже может вызывать спектр возможных речевых реакций агента #1. Например: *Чего ты раскомандовался?! Ну и как я его включу?! и т. д.* Ситуация включения света (взрыв выключателя) также может вызывать спектр возможных высказываний агента #1:

S1*НЕАДЕКВ: *Я все делаю не так! Из меня никудышний электрик! Чёрт, я его ненароком сломал, надо незаметно отсюда смотаться!* **S1*р-сценарии:** *Как же теперь это починить? Может, отвёрткой? Не надо мне лазить, куда не нужно! Надо быть внимательнее / передохнуть! Ну ладно, в конце-концов это не так страшно; я же не специально это сделал! Это – всего лишь какой-то выключатель!*

S2*СУБЪЕКТ/ЭМОЦ (обращаясь к агенту #2): *Чего ты так дёргаешься? Не стоит переживать из-за какого-то выключателя! Тебе нужно быть спокойнее. Попей водички!*

S7*ФРУСТР: *О! Чёрт! Ой! Как это она так взорвалась?! Почему эти выключатели всё время взрываются?!*

S8*НЕАДЕКВ: *Чего я так пугаюсь?!* **S8*р-сценарии:** *Надо мне быть спокойнее!*

S3*НЕАДЕКВ: *Опять я тебя подвёл! Это я не проверил выключатель! Я всё время ломаю выключатели, а потом все сидят без света! Как стыдно, надо отсюда тихо исчезнуть!* **S3*р-сценарии:** *Прости, пожалуйста! Не волнуйся, я как-нибудь это починю / я как-нибудь включу свет! Я же не специально! Я его только тронул, а он*

² При разметке мы указываем сегмент (S1, S2 и т.д.) и название д-сценария (НЕАДЕКВ, СУБЪЕКТ и т.д.); названия р-сценариев здесь не указываются.

– вот! Ну что делать; да, теперь он сломан! Ну да, сломал, но жизнь-то продолжается / это не самое страшное в жизни! Да я вообще его не трогал; это не я!

S4*НЕАДЕКВ: Я этот выключатель не трогал, а ты уже на меня злишься! Неужели ты думаешь, что я теперь его должен починить? Ты чего, думаешь, его теперь нельзя починить? Ага, и свет, типа, уже никогда не включить! (в последнем случае использована ирония и несобственная прямая речь)

S5*ОПАСН: Меня чуть током не убило из-за твоего выключателя / чуть твой выключатель током не убил! Это что, шутка? Ты это подстроил?

Если высказывание агента #1 было эмоциональным (было порождено в результате активизации д-сценария), то агент #2 может ответить агенту #1, например: **S2*ЭМОЦ:** Чего ты так из-за всего нервничаешь? В результате мы получим диалог, например, такой:

Агент #1 (ломает выключатель): О! Чёрт! Я всё время всё ломаю!

Агент #2: Чего ты так из-за всего нервничаешь?

Группировки

Агент может использовать различный набор субъектов высказывания (точнее, различный набор семантических субъектов в д/р-сценарии). В качестве субъекта могут выступать (а) сам агент – У меня ничего никогда не получается! (б) агент и игрок – У нас опять ничего не получилось! (в) игрок – У тебя опять ничего не получается! На какие кнопки ты вообще жмешь? (г) игрок и другие враждебные агенты (птицы) – Вы меня все уже достали! При выборе определённых речевых стратегий агент может группировать себя с враждебными агентами, в этом случае его высказывание будет ироничным (обращаясь к игроку после того, как птица в очередной раз сбрасывает агента с платформы): Мы тебя, наверно, уже достали с этими птицами!

Дальнейшие перспективы: семантический синтез высказываний

Интеллектуальный агент должен иметь возможность действовать не только в фиксированных ситуациях (как играх типа “quest”), но и в динамической, нефиксированной игровой реальности (как, например, в играх типа “action” или играх-стратегиях). Для такого развития интеллектуальных агентов выбор высказывания из базы данных может оказаться неприемлемым, и должен быть заменён процедурой синтеза семантического графа с последующим построением поверхностного представления (возможно, с дополнительным привлечением статистических методов). Как представляется, в перспективе именно этот подход должен быть применён для создания роботов, взаимодействующих с человеком в реальном мире. Мы полагаем, что использование модели д/р-сценариев позволяет не только размечать высказывания при их организации в базу данных, но и видоизменять высказывание (если оно представлено в виде семантического графа). Д/р-сценарии действуют на валентной структуре: каждый д-сценарий принимает на вход предикат с актантами и, далее, может модифицировать эту структуру двумя способами: (а) добавляя семантические признаки эмоциональной речи и (б) выбирая производную речевую стратегию для формирования ответа. Если агент А предпринимает некоторое действие (например, ‘идёт’), то он может передавать валентную структуру ‘А идёт’ агенту В. В свою очередь, агент В может активизировать д-сценарий и (i) дополнить исходную валентную структуру, в частности, признаком интенсивности предиката – ‘А интенсивно идёт’, (ii) выбрать из словаря лексемы для признаков ‘интенсивно идти’ – бежать, топтать, (iii) выбрать производную речевую стратегию, например, вопрос к А – Ну и куда ты прёшься/потопал?

Личностные характеристики агента и выбор конкретного высказывания

Компьютерный агент может обладать определённым темпераментом или менять своё настроение во время игры. В нашем случае настроение и темперамент агента могут влиять на выбор высказывания из предлагаемого инвентаря. Так, если птица сбивает агента с платформы, он может (кроме прочего) сказать *Что это за уроды тут летают?* или *Почему я всё время на всех натыкаюсь?* – выбор будет зависеть именно от настроения и темперамента агента. Настроение агента может повышаться в результате удачи, или, наоборот, портиться из-за неудачных действий пользователя, а также вследствие нескольких фрустраций. Для описания динамики изменения настроения в существующих интерфейсах уже применяются динамические психологические модели, например, модель ОСС – Ortony-Clore-Collins [Ортони, Клоур et al., 1996]. Совокупность «субъективных» факторов, влияющих на эмоциональную реакцию агента, рассматривается как иерархия *контрольных состояний* [Allen, 2001: 11-15]. На вершине этой иерархии находятся такие “статичные” факторы как *личностные характеристики*, далее следуют *отношения* (наша реакция на некоторый объект зависит от нашего отношения к нему) и *правила поведения*, еще ниже – относительно динамичные факторы: *настроения, эмоции, желания и намерения*. Хотя на реакцию агента оказывают влияние все перечисленные факторы, в конкретной эмоциональной ситуации те или иные факторы могут преобладать.

На основании проведённого нами экспериментального исследования можно говорить о связи между личностными (психологическими) характеристиками реальных испытуемых и их речевыми ответами в эмоциональной ситуации. Определённые типы ответов (сегмент ответа, отношение к д- или р-сценарию) коррелируют с определёнными личностными характеристиками [Kotov, 2005: 299-300]. Это позволяет при конструировании речевого поведения агента создать портрет того или иного характера. Например, отмеченная в эксперименте корреляция *нейротизма* (по тесту Личко) и количества высказываний, относящихся к д-сценариям, позволяет регулировать уровень “нейротизма” агента: предпочитая активизацию д-сценариев, мы тем самым создадим портрет “более нервного агента”.

Литература

- Апресян В. Ю. *ИмPLICITная агрессия в языке // Компьютерная лингвистика и интеллектуальные технологии: Тр. Междунар. конференции Диалог'2003 (Протвино, 11-16 июня 2003 г.) / Ред. И. М. Кобозевой, Н. И. Лауфер, В. П. Селегей.- М.: Наука, 2003.- С. 32-35.*
- Гловинская М. Я. *Скрытая гипербола как проявление и оправдание речевой агрессии // Сокровенные смыслы: Слово. Текст. Культура.- М.: Языки славянской культуры, 2004.- С. 69-76.*
- Котов А. А. *Механизмы речевого воздействия в публицистических текстах СМИ: Дис. ... канд. филол. наук; 10.02.19; - Защищена 23.06.03.- М., 2003.*
- Котов А. А. *Аспекты моделирования эмоционального речевого поведения // Компьютерная лингвистика и интеллектуальные технологии: Тр. Междунар. конференции Диалог'2004 ("Верхневолжский", 2-7 июня 2004 г.) / Ред. И. М. Кобозева, А. С. Нариньяни, В. П. Селегей.- М.: Наука, 2004.- С. 346-355.*
- Ортони А., Клоур Д., Коллинз А. *Когнитивная структура эмоций // Язык и интеллект.- М.: Прогресс, 1996.- С. 314-381.*
- Практическая психология в тестах, или как научиться понимать себя и других.- М.: АСТ-Пресс Книга, 2003.*
- Allen S. R. *Concern Processing in Autonomous Agents, Ph.D thesis.- University of Birmingham, 2001.*
- Becker C., Kopp S., Wachsmuth I. *Simulating the Emotion Dynamics of a Multimodal Conversational Agent // E. André et al. (Eds.): ADS 2004, LNAI 3068,- Berlin, Heidelberg: Springer-Verlag, 2004.- С. 154-165.*
- Inanoglu Z., Young S. *Intonation Modelling and Adaptation for Emotional Prosody Generation // J. Tao, T. Tan, and R.W. Picard (Eds.) ACII 2005, LNCS 3784.- Berlin, Heidelberg: Springer-Verlag, 2005.- С. 286-293.*
- Kotov A. A. *Application of Psychological Characteristics to D-Script Model for Emotional Speech Processing // J. Tao, T. Tan, and R.W. Picard (Eds.) ACII 2005, LNCS 3784.- Berlin, Heidelberg: Springer-Verlag, 2005.- С. 294-302.*
- Ochs M., Niewiadomski R., Pelachaud C., et al. *Intelligent Expressions of Emotions // J. Tao, T. Tan, and R.W. Picard (Eds.) ACII 2005, LNCS 3784.- Berlin, Heidelberg: Springer-Verlag, 2005.- С. 707-714.*
- Sloman A., Chrisley R. *Virtual Machines and Consciousness // Journal of Consciousness Studies.- 2003.- 10 (4-5).- С. 133-172.*