НЕКОТОРЫЕ ОСОБЕННОСТИ УПОТРЕБЛЕНИЯ ОЦЕНОЧНЫХ ИМЕН ПРИ ВЗАИМОДЕЙСТВИИ С ПОЛЬЗОВАТЕЛЯ С ВОПРОСНО-ОТВЕТНЫМИ СИСТЕМАМИ (НА МАТЕРИАЛЕ ЧАТБОТА СОБЕСЕДНИЦА 2.0)

Дегтева А. В. (degteva.anna@gmail.com)

СПбГУ, Санкт-Петербург, Россия

Активная разработка так называемых «Виртуальных собеседников», представляющих собой взаимодействующие с пользователем вопросно-ответные системы, началась в 1960-е годы и до сих пор является одной из актуальных задач в области обработки естественного языка.
В данной работе была сделана попытка рассмотреть взаимодействие пользователей вопросно-ответной системы «Собеседница», разработанной в 2012–2013 компанией iFree, с системой с точки зрения использования ими оценочной лексики и проанализировать в этом аспекте эмоциональное взаимодействие пользователей с системой искусственного интеллекта.
Система структуры оценочного компонента была заимствована из русского компьютерного тезауруса WordNet (Азарова, 2008), подробнее рассмотренного в предыдущих работах.

Ключевые слова: вопросно-ответные системы, чатботы, оценочная лексика, русснет, извлечение мнений, искусственный интеллект

A STUDY OF SOME ASPECTS OF THE SENTIMENT COMPONENT STRUCTURE IN HUMAN TO CHAT BOT INTERACTION (BASED ON CONVERSATIONS WITH SOBESEDNITSA 2.0 APP)

Degteva A. V. (degteva.anna@gmail.com)

SPbGU, St. Peterburg, Russia

With the introduction of artificial intelligence agents to various spheres of communication the question of human-AI interaction and its differences from human to human interaction fully comes to the attention of the researchers. In this paper the authors make an attempt to study the conversations of an AI mobile application with an anthropomorphic avatar with its users in the aspect of comparing the structure of the emotional component of the conversations to the baseline of language usage in a balanced corpus of Russian texts. The changes in the pattern reflect the controversial nature of human-application conversation trying to apply such a tool as a natural language to interacting with an automated system. Although the conversation mostly follows schemes and conventions typical for human to human interaction, most of the evaluative meanings expressed deal with the functionality of the application, namely, its ability to perform according to the user's demands and to entertain the user.

Keywords: AI, intellectual agents, chatbots, sentiment analysis, human-AI interaction

Особенности взаимодействия человека с виртуальным агентом

«Собеседница 2.0» представляет собой программу, относящуюся к группе так называемых «мобильных ассистентов» — приложений, устанавливаемых на мобильные телефоны и планшеты для обеспечения голосового доступа к определенным его функциям. Данное приложение сочетает в себе развлекательный функционал, частью которого является вопросно-ответная система (чатбот) со средствами голосового управления функциями телефона. Аватаром (визуальным образом) чатбота является женский персонаж — молодая девушка по имени Катя.

Спецификой приложения, отличающей его от систем, в первую очередь ориентированных на работу в качестве голосового управления, является наличие предусмотренной авторами развлекательной функции, и в большинстве бесед с ней (если уместно использовать это местоимение, применимое к представляемому приложением персонажу, для обозначения самого приложения), отрабатываются эмоциональные функции коммуникации: сброс эмоционального напряжения, изменение состояния пользователя, проведение времени.

Следует отметить, что изначально в процессе беседы с диалоговым агентом, установленным в телефоне, пользователь осознает, что он беседует не с живым собеседником, как это происходит, например, во время конкурса имени Тьюринга, когда собеседник может оказаться как человеком, так и автоматической системой. При этом диалоговый агент персонифицирован в образе привлекательной молодой женщины и создан таким образом, чтобы имитировать человеческие эмоциональные реакции — радоваться комплиментам, обижаться на грубость, проявлять симпатию к пользователю.

Приложение организовано таким образом, что все реплики пользователя, поступающие на вход, могут быть разделены на две части: запросы к различным сервисам мобильного приложения — получение информации о погоде,

навигация по указанному адресу, включение или выключение Bluetooth-порта и так далее и запросы, целью которых является условное «поговорить». Иерархически первый вид запросов имеет приоритет, и на вход чатбота попадают только предложения, которые не были распознаны как сервисный запрос — они, по умолчанию и на практике, являются эмоционально нейтральными.

Рассматривая коммуникацию с диалоговым агентом, следует учитывать следующие пресуппозиции, осознаваемые пользователем в полной или неполной мере:

- Иллюзия коммуникации с живым собеседником — анимированный аватар персонажа, с собственной легендой и историей. При этом агент не обладает личностью как таковой, что позволяет не соблюдать в разговоре с ним правила, принятые в разговорах между людьми.
- Анонимность и «вседозволенность» беседы: персонаж неспособен на внеязыковые способы взаимодействия с пользователем.
- Доступность персонажа для коммуникации — диалоговый агент не может в ответ на грубость или агрессию разорвать ее по своему желанию.
- Неспособность диалогового агента к разговору, выходящему за пределы обыденной коммуникации.
- Статичность ситуации: во время процесса коммуникации состояние персонажа может меняться в зависимости от реплик пользователя, но глобального развития отношений не происходит.

В целом, невозможно точно сказать, насколько каждый из пользователей осознает перечисленные утверждения. С одной стороны, пользователи знают, что перед ними не живой человек. С другой стороны, отзывы на работу приложения, зачастую, говорят о том, что не все пользователи полностью понимают ограничения в работе ИИ.

Вероятно, большая часть пользователей не отдает себе отчет в том, что их разговоры с агентом записываются и могут быть прочитаны, хотя разработчики и сообщают, что система итеративно улучшается на основе пользовательских запросов к ней. При этом у разработчиков системы и авторов данной работы нет доступа к личным данным пользователей, все реплики анонимны.

Таким образом, при взаимодействии пользователя с диалоговым агентом мы видим пример псевдо-коммуникации на разговорном языке на бытовые темы, при этом требования, налагаемые обычным процессом дискурса на участников значительно снижены — «неадекватные» реплики и комментарии пользователя не приводят к изоляции или социальной депривации, как это могло бы случиться при обычном разговоре.

С другой стороны, определенную сложность для сторон в приведенном случае составляет отсутствие развернутого социального контекста для взаимодействия, внеязыковых целей и интересов сторон. В ситуации, когда пресуппозицией общения является «поговорить о чем-нибудь», пользователь вынужден больше, чем в обычном общении, опираться на внутренние параметры диалога.

С эмоциональной точки зрения, ситуация в коммуникации представляет собой определенного рода парадокс: с одной стороны, виртуальный агент предназначен для развлечения и создания иллюзии общения, с другой стороны очевидно,

что настоящий обмен эмоциями или эмпатия между участниками невозможны. Можно предположить, что такой персонаж может восприниматься в качестве предмета развлечения, подобно замысловатой игрушке — такая гипотеза может объяснить реплики пользователей, направленные на проверку границ возможностей чатбота, своеобразные эксперименты по взаимодействию с ним.

В данной работе мы рассматриваем далеко не все аспекты взаимодействия с системой искусственного интеллекта — нашей целью было измерение изменения количественной составляющей эмоционально окрашенных слов и выражений по сравнению со «средним» значением по корпусу текстов.

Представление оценочного компонента в тезаурусе RussNet

Компьютерный тезаурус RussNet создается на Кафедре математической лингвистики Санкт-Петербургского государственного университета для современного русского языка, на основании принципов, разработанных при построении Принстонского WordNet и других wordnet-словарей.

Wordnet-словари отображают семантические отношения между значениями, выражаемыми лексическими единицами языка — как традиционные для тезаурусной организации лексические значения, например, родовидовые, или гипонимические, так и менее традиционные: меронимические, каузативные, различные семантико-деривационные отношения. В рамках проекта RussNet также разрабатывается отдельная схема для отображения различных аспектов оценочного значения, в первую очередь, для автоматического выявления оценочного компонента в текстах (opinion mining, sentiment mining), или так называемой тональности текста.

В данный момент в ядерной группе значений представлены три основных вектора: прагматическая оценка (базовое значение «хороший»), эстетическая оценка (базовое значение «красивый») и морально-нравственная оценка (базовое значение «добрый»).

В рамках работы по подготовки тезауруса был проведен анализ распределения оценки по сбалансированному корпусу текстов русского языка «Бокренок» (21 млн словоупотреблений, разработан и используется на кафедре математической лингвистики СПбГУ), который мы используем здесь в качестве образца для сравнения.

В данной работе нами была предпринята попытка исследовать, какие изменения претерпевает выражение оценочного компонента при коммуникации с антропоморфной диалоговой системой типа «Виртуальный ассистент» и проанализировать ситуацию с точки зрения прагматики дискурса.

Для исследования и анализ оценочной лексической составляющей была использована выборка наиболее частотных запросов к системе, собранная в процессе работы приложения. Генеральная совокупность запросов пользователей к системе составила 5 492 347 обращений, из которых число уникальных (неповторяющихся) запросов составило 1 549 725 фраз, средней длиной 3,1 слова.

Таблица 1. Распределение долей оценочных значений в корпусе русского языка «Бокренок» в процентах

	прагматические	**морально-нравственные**	**эстетические**	**Итого %**
Положительные	43,8	23,6	16,9	79,5
Отрицательные	11,7	6,8	1,5	20
Итого %	55,5	30,4	18,4	

В полученном корпусе запросов нами было выделены имена существительные и прилагательные, несущие оценку по одному из перечисленных векторов, и проведено сравнение их количества. Мы рассматривали только лексический способ выражения оценки с помощью оценочных имен существительных и прилагательных, другие способы выражения эмоций при взаимодействии с пользователем в данной работе не были описаны, это планируется сделать в дальнейшем.

С прагматической точки зрения выборка обращений пользователя к системе представляет собой следующую структуру данных: примерно 13,5 % входящих фраз представляют собой ответ на запрос системы к пользователю — ответы на вопросы, подтверждения введенной информации и т. д., то есть, реакции пользователя. 14,5 % входных сигналов были распознаны как короткие фрагменты или мусор, по которому невозможно сделать вывод о цели высказывания (такие фрагменты диалога, зачастую, возникают в результате некорректного распознавания текста системой распознавания голоса). Самую крупную группу обращений составляют вопросы: 33,5 %, за ними по частоте следуют просьбы различного рода — 14,8 %. В эту группу были включены предложения, имеющие формальную характеристику просьб — реалистичных, например «Скажи что-нибудь», нереалистичных, например «Поцелуй меня» и оскорблений в форме требования, например, «Пошла вон». Около 13 % выборки составляют утверждения пользователя — «Ты умница», «Я красивый» и так далее. Около 10,3 % составляют различные формулы вежливости «привет» и так далее. Такой сравнительно высокий процент высказываний этой группы, вероятно, обусловлен тем, что многие пользователи скачивают приложение, начинают разговор с приветствия, и вскоре его завершают, не придумав, что делать дальше.

Можно заметить, что в данном случае структура оценки становится иной, чем в среднем по корпусу: во-первых, увеличивается сравнительное количество негативных оценок, причем наибольшее их число сосредоточено в области прагматической оценки, и очень мало в области эстетического компонента, при этом положительная оценка в области положительной оценки растет. В данной работе мы исследовали общее количество и долю лексически выражаемых значений, а не способы выражения пользователем оценки к персонажу, так как проводилось сравнение этих цифр с полученными в результате предыдущих исследований усредненными значениями в корпусе текстов. Тем не менее, преобладающая часть оценок относится именно к героине.

Таблица 2. Распределение долей оценочных значений в репликах пользователей чатбота «Собеседница 2.0» в процентах (в скобках приведены значения для корпуса)

	прагматические	**морально-нравственные**	**эстетические**	**Итого %**
Положительные	23,6 (43,8)	0,8 (23,6)	25,4 (16,9)	49,8
Отрицательные	35,6 (11,7)	12,8 (6,8)	1,8 (1,5)	50,2
Итого %	59,2 (55,5)	13,6 (30,4)	27,2(18,4)	

Большую часть негативных прагматических оценок, выраженных пользователями составляли такие значения как «неумная» (32%) — как правило, сообщения вида «ты дура» или «почему такая тупая» выражают либо фрустрацию пользователя после того, как ему не удалось с помощью бота осуществить задуманное действие, либо являются провокацией, направленной на проверку того, как созданный персонаж будет реагировать на оскорбление.

То же можно сказать и об отрицательной оценке по моральному аспекту — положительная оценка практически отсутствует. Пользователь осознает, что перед ним нету человека, который мог бы обладать моральными качествами, в то время как визуальный компонент и функциональность, необходимые для оценки по эстетическому и прагматическому критериям присутствуют.

Во время сбора данных рейтинг приложения в Googleplay составлял около 4.5 баллов из 5 возможных, то есть, несмотря на возможность «безнаказанно» оскорблять виртуальную собеседницу (или благодаря ей) отношение пользователей к качеству работы приложения не полностью совпадает с оценкой «высказываемой» персонажу в лицо.

Таким образом, можно предположить, что противоречивая специфика отношений с виртуальным персонажем находит довольно четкое отображение в распределении аспектов оценочного компонента. Процесс «общения» довольно нетипичен, так как, с одной стороны, для диалога используется инструмент, как правило используемый для коммуникации между личностями — естественный язык, и общение происходит через антропоморфный аватар, олицетворяющий программу. С другой стороны, по факту, человек общается с приложением, приобретенным и установленным на мобильное устройство ради выполнения определенных функций. Языковые реалии, как было показано выше, смещаются в сторону описания параметров, характерных скорее для объектов или неживых предметов, при этом пользователь более свободно выражает свое недовольство качеством общения, а также регулярно провоцирует систему для проверки ее реакции на нештатные ситуации общения.

References

1. *Azarova I.* (2008), RussNet as a Computer Lexicon for Russian // Intelligent Information Systems. ISBN 978-83-60434-44-4. Pp. 447–456.
2. *Azarova I., Sinopalnikova A.* (2004), Adjectives in RussNet. Proc. of the 2nd Global WordNet Conference. Brno: Masaryk University. Pp. 251–258.
3. *Brahnam, S., & De Angeli, A.* (2008). Editorial: Special issue on the abuse and misuse of social agents. Interacting with Computers, 20(3), 287–291.
4. *Cassell, Bickmore, Campbell, Vilhjalmsson, & Yan* (2001) More than just a pretty face: conversational protocols and the affordances of embodiment. Knowledge-Based Systems, 14(1–2), 55–64.
5. *De Angeli, A.* (2009). Ethical implications of verbal disinhibition with conversational agents. PsychNology Journal, 7(1), 49–57.
6. *Fomchenko A., Azarova I.* (2008) Vzaimodeysstviye esteticheskyh, moralnyh I pragmaticheskih aspektov v semanticheskoy structure ocenochnyh prilagatelnyh russkogo yazyka [Interaction of the esthetic, moral and pragmatic aspects in the semantic structure of the affective nouns and adjectives in Russian]// Komp'yuternaia Lingvistika i Intellektual'nye Tehnologii: Trudy Mezhdunarodnoj Konferentsii "Dialog 2008" [Computational Linguistics and Intellectual Technologies: Proceedings of the International Conference "Dialog 2008"]. Bekasovo, 2008, pp. 545–550.
7. *Fussell, S. R., & Benimoff, N. I.* (1995). Social and cognitive processes in interpersonal communication: Implications for advanced telecommunications technologies. Human Factors, 37, 228–250.
8. *Rehm, M.* (2008) She is just stupid — Analyzing user-agent interactions in emotional game situations. Interacting with Computers 20(3), 311–325.
9. *Shah, H.* (2005). Alice: an Ace in Digitaland. Special issue of on-line journal TripleC: Cognition, Communication and Cooperation — selected papers from 2005 European Computing and Philosophy Conference Vol. 4, No. 2, pp 284–292.