

Когнитивная архитектура робота Ф-2, поддерживающего коммуникацию с человеком

А.А. Котов, Н.А. Аринкин, Л.Я. Зайдельман, А.А. Зинина

НИЦ «Курчатовский Институт», Москва; email: kotov_aa@nrcki.ru

Один из подходов к исследованию когнитивных функций человека – это создание компьютерных моделей, воспроизводящих отдельные механизмы эмоций, мышления, логического вывода или поведения. Актуальным приложением такого метода является разработка «человекоподобных» роботов, моделирующих не только функции мышления, но и средства общения с человеком.

Робот Ф-2 разрабатывается как исследовательский проект с целью построить модель когнитивной архитектуры человека, достаточную для поверхностного понимания текста и поддержания коммуникации, включая речевое, жестовое и мимическое общение. Робот получает на вход тексты на русском языке в письменной или акустической форме (для распознавания используется Yandex Speech API) и на выходе демонстрирует речевые высказывания, мимику и жесты, соответствующие реакции на поступивший текст.

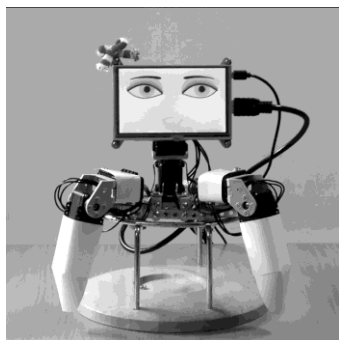


Рис. 1. Робот Ф-2, предназначенный для поддержания речевого, жестового и мимического общения с человеком

В составе робота находятся три основных компонента:

- **лингвистический компонент** выполняет анализ текста и строит его семантическое представление;
- **компонент сценариев** выбирает одну или несколько реакций на построенное семантическое представление;
- **компонент управления** выполняет на роботе речевые, мимические или жестовые паттерны, определённые для конкретной реакции, а также обеспечивает баланс между конкурирующими реакциями.

Хотя возможность компьютерного понимания текста вызывает большие споры в области философии сознания, в частности, это дискуссии вокруг «аргумента китайской комнаты» Дж. Сёрла [1], мы рассматриваем понимание текста как возможность построить семантическое представление в лингвистическом компоненте, выбрать реле-

вантную реакцию в компоненте сценариев и выполнить её на роботе с помощью компонента управления.

Лингвистический компонент

В составе компонента последовательно выполняются три этапа лингвистической обработки: морфологический, синтаксический и семантический анализ. Каждый этап является необходимой ступенькой для следующего. В созданной нами системе обработки текстов каждый модуль передаёт результаты своей работы следующему модулю, а также может сохранить эти результаты в базе данных. Пользователь может выбирать набор модулей в системе в зависимости от конкретной задачи. Рассмотрим каждый из этапов анализа и связанные с ним задачи.

В основе **морфологического модуля** лежит словарь, содержащий более 48 тысяч лексем со всеми словоформами, грамматическими характеристиками и семантическими признаками. Словарь сформирован на основе материалов проекта OpenCorpora [2]. Морфологический модуль ищет каждую словоформу текста в словаре и приписывает ей морфологические характеристики. Результатом работы данного этапа является морфологический разбор предложения, который может быть сохранён в виде JSON-структуры. В данной структуре для каждой словоформы предложения указана её нормальная форма, а также грамматические характеристики (род, число, падеж, время и др.). Кроме того, в структуре отображаются все случаи лексической омонимии – количество наборов грамматических характеристик и начальной формы словоформы-омонима соответствует числу гипотез, порождённых алгоритмом морфологического разбора. Все несловарные лексемы получают в структуре соответствующую помету. Морфологический этап является не только подготовительным шагом для синтаксического анализа, но и самостоятельным процессом, который может использоваться для поиска текстов по ключевым словам, а также для выявления неизвестных системе лексем с целью пополнения словаря.

Синтаксический модуль системы хранит описание грамматики русского языка в виде 490 грамматических правил на специально разработанном языке syntXML [3]. Каждое правило может менять состав признаков одного слова, либо устанавливать синтаксическую связь между двумя и более словами. Последовательное применение синтаксических правил позволяет построить для предложения синтаксическое дерево. В случае омонимии для одного предложения строится два и более деревьев. Кроме того, для части предложе-

ний может быть не построено ни одного дерева – такое возможно, если в предложении присутствуют не описанные правилами синтаксические структуры или несловарные лексемы с неизвестной сочетаемостью. Синтаксическое дерево далее может передаваться в семантический компонент для конструирования смысла, либо сохраняться в базу данных в формате treeJSON – этот формат фиксирует структуру синтаксического дерева, а также морфологические и синтаксические признаки каждой словоформы.

В задачу **семантического модуля** входит конструирование семантического представления текста, которое далее передаётся в компонент сценариев и может быть сохранено в базу данных. Семантический модуль работает с набором из 560 семантических признаков: слова получают определения в виде ранжированного множества признаков. Всего с помощью признаков размечены 28600 слов словаря. В качестве семантического представления односоставного предложения используется семантическая предикация: глагол и множество его актантов. В семантической предикации признаки каждого актанта приписаны семантической валентности: *агенса, пациенс, инструмент* и т. д. Такое семантическое представление является достаточным для общей классификации фактов, содержащихся в тексте – при сохранении фактов в базу данных пользователю становится доступен поиск по структуре ситуации. Так, по запросу ‘ситуация, в которой некто живой что-то ест’ пользователь получит и предложение *кошка ест рыбу, и Вася спал все котлеты*. Семантические представления также используются роботом для выбора в компоненте сценариев коммуникативной реакции на поступившее сообщение.

Компонент сценариев

Семантическое представление сравнивается с набором сценариев – *продукций* или правил типа *если-то*. Для выбора роботом эмоциональной реакции используется инвентарь *д-сценариев* [4, 5]. Сценарии позволяют роботу рассматривать входящее событие как ситуацию ‘меня никто не любит’ (д-сценарий НЕНУЖН), ‘меня все обманывают’ (ОБМАН), ‘они не соображают, что делают’ (НЕАДЕКВ) или ‘на меня все обращают внимание’ (ВНИМАНИЕ) и т. д. Активизация сценария зависит как от входящего события, так и от «темперамента» или текущего «настроения» робота.

Сценарий связан с шаблонами высказываний, жестами и элементами мимики, описанными в составе поведенческого блока на языке BML – Behavior Markup Language [6]. При активизации сценария связанные с ним блоки BML передаются в компонент управления роботом для исполнения.

Компонент управления роботом

Компонент управления роботом синтезирует поступающие высказывания (используется Yandex Speech API), а также выполняет поступившие инструкции BML с помощью исполнительных элементов робота: рук, головы и элементов лица. Поступившие блоки BML упорядочиваются по снижению активизации породившего их сценария (неотложные реакции перемещаются вперёд) и выполняются по мере освобождения исполнительных элементов робота.

Робот получает доступ к библиотеке жестовых и мимических паттернов, сформированной на основе Русского эмоционального корпуса REC [7]. Корпус содержит видеозаписи реального общения людей в эмоциональных ситуациях: на университетских экзаменах, на приёме населения в муниципальной службе, а также на специально организованных интервью. Разметка корпуса выделяет жесты и элементы мимики, характерные для определённых коммуникативных реакций. Эти элементы зарисовываются с помощью программы Blender и сохраняются в базу данных, доступную роботу при выполнении пакетов BML.

Предлагаемая архитектура является «поверхностной» моделью когнитивных функций человека, которая, тем не менее, позволяет строить семантическое представление высказываний и выбирать поведенческий компонент для реакции робота.

Литература

1. Сирл Д. Разум, мозг и программы // Глаз разума / Ред. Д. Хофштадтер и Д. Деннет. Самара: Барак-М, 2003. С. 315-331.
2. Bocharov V. V., Alexeeva S. V., Granovsky D. V., Protopopova E. V., Stepanova M. E., Surikov A. V. Crowdsourcing morphological annotation // Компьютерная лингвистика и интеллектуальные технологии. Вып. 12 (19). Т. 1. М.: РГГУ, 2013. Р. 109-114.
3. Kotov A., Zinina A., Filatov A. Semantic Parser for Sentiment Analysis and the Emotional Computer Agents // Proceedings of the AINL-ISMW FRUCT 2015. 2015. P. 167-170.
4. Котов А. А. Механизмы речевого воздействия в публицистических текстах СМИ: Дис. ... канд. филол. наук; 10.02.19; - Защищена 23.06.03. М., 2003.
5. Котов А. А. «Машина Оруэлла»: подходы к автоматическому созданию воздействующих текстов // Понимание в коммуникации: Человек в информационном пространстве. Ярославль: ЯГПУ, 2012.
6. Kopp S., Krenn B., Marsella S., Marshall A., Pelachaud C., Pirker H., Thórisson K., Vilhjálmsson H. Towards a Common Framework for Multimodal Generation: The Behavior Markup Language // Intelligent Virtual Agents. 2006. P. 205-217.
7. Kotov A., Budyanskaya E. The Russian Emotional Corpus: Communication in Natural Emotional Situations // Компьютерная лингвистика и интеллектуальные технологии. Вып. 11 (18). Т. 1. М.: РГГУ, 2012. Р. 296-306.