

Моделирование естественно-языкового вывода для робота-компаньона Ф-2

А. А. Котов^{1,2,3}, **А. А. Филатов**^{2,4}, **З. А. Носовец**¹

¹ НИЦ «Курчатовский институт», Москва

² Российский государственный гуманитарный университет,
Москва

³ Московский государственный лингвистический университет,
Москва

⁴ ООО «Яндекс.Технологии», Москва
e-mail: kotov@harpia.ru

Аннотация. В рамках проекта робота Ф-2 мы разрабатываем механизм естественно-языкового вывода. Он позволяет роботу по входящей ситуации, например, по смыслу высказывания, реконструировать возможные причины и следствия, которые далее могут быть оценены по их эмоциональности. Этот механизм обеспечивается работой 4500 рациональных сценариев (р-сценариев), которые получены с помощью кластеризации семантики текстов и с помощью опроса информантов.

Ключевые слова: когнитивная модель, роботы-компаньоны, понимание текста.

Традиционные модели концептуальной обработки информации основывались на механизме *сценариев*, которые позволяли построить для описываемого события его возможные причины и следствия (Schank, 1975). Подобные методы применялись и в классических моделях решения задач (Newell, Simon, 1972): процесс решения был представлен как поиск в проблемном пространстве, где каждое применение операции меняло репрезентацию ситуации, и таким образом от исходной ситуации задачи к её целевому состоянию строился путь, являвшийся решением. В настоящее время для систем ответов на вопросы наиболее попу-

Исследование выполнено за счет гранта Российского научного фонда № 19-18-00547, <https://rscf.ru/project/19-18-00547/>.

© Котов А.А., Филатов А.А., Носовец З.А., 2024

лярны нейросетевые методы, основанные на предварительной тренировке нейросети на больших массивах текстовых данных. Вместе с тем, классические подходы, основанные на сценариях, могут обладать потенциальными преимуществами и лучше моделировать когнитивные функции человека при размышлении или решении задач.

Мы разрабатываем когнитивную модель для робота-компаньона Ф-2: эта модель должна обеспечивать естественную коммуникацию робота с человеком с помощью речи, мимики и жестов, а также воспроизводить процессы, связанные с эмоциональным реагированием и пониманием речи. Существенным аспектом понимания для нас является возможность робота соотносить входящие события с известными роботу эмоциогенными паттернами – д-сценариями (Котов, 2021). Если некоторое входящее событие совпадает с посылкой д-сценария, этот д-сценарий активизируется и выполняет на работе некоторую поведенческую (эмоциональную) реакцию. Такое входящее событие может быть смыслом текста (обращённое к роботу высказывание вызывают негативные или позитивные реакции робота), визуальным событием (человек смотрит на робота – это приятно роботу или вызывает беспокойство) и тактильным событием (человек трогает робота – робот воспринимает это как приятное поглаживание или как беспокоящее вторжение в личную сферу). Контраст между негативными и позитивными оценками одного входящего события, как мы считаем, может быть ключевым методом создания сложной когнитивной модели робота, методом иронии и моделирования психического (theory of mind).

Другой важный аспект понимания – это возможность конструирования возможных причин и следствий наблюдаемого события, возможность целеориентированного действия, когда субъект перебирает различные методы для достижения цели. Именно это направление считалось центральным для классических моделей концептуальной обработки информации (модели Шенка, Ньюэлла и Саймона), но в нашей модели эта способность важна для робота, поскольку позволяет соотнести рациональный вывод с д-сценариями: оценить возможные негативные и позитивные причины и следствия наблюдаемого события. В противопоставлении эмоциональных процессов (д-сценариев) и механизма рационального вывода (р-сценариев) мы следуем модели

CogAff (Sloman, 2001), в которой эти процессы разнесены на разные когнитивные уровни. Они конкурируют при обработке каждого стимула, но при этом взаимодействуют при сложных (*вторичных* и *третичных*) эмоциях. Например, при *вторичной эмоции* человек может в ходе рационального вывода прийти к умозаключению, которое вызовет эмоциональную активацию.

Модель естественно-языкового вывода робота Ф-2 состоит из 4500 р-сценариев. У каждого р-сценария есть посылка (распознаваемый фрейм) и могут присутствовать несколько следствий – связанных фреймов: предшествующие ситуации, последующие ситуации и интерпретации (ситуация, по-другому представляющая ситуацию посылки). При поступлении стимула все р-сценарии упорядочиваются по близости их посылки к стимулу. Победившие р-сценарии строят свои связанные фреймы, чтобы сравнить их с д-сценариями – так может быть оценена эмоциональность ситуации. К каждому фрейму сценария может быть приписано действие: в текущем варианте все р-сценарии настроены сообщать в речи прототипическую ситуацию для каждого связанного фрейма, хотя в реальности такое речевое поведение и кажется избыточным. В данном варианте синтеза р-сценарий использует названия прототипических участников ситуации (например, *крыса бегала*); хотя он и сохраняет референт из входящего события (*курица бегала*), но пока не может подставить его в нужную синтаксическую позицию и обеспечить грамматическое согласование (**мама менял трубу*). В следующем диалоге мы приводим ответы робота (Р) на высказывания человека (Ч), в квадратных скобках указывая сохранённые роботом референты.

Ч: *Курица бегала по двору.*

Р: *Крыса[курица] бегала по кругу. Ранее крыса[курица] потерялась. Потом крыса [курица] выбралась на свободу. Это значит, крыса [курица] искала выход.*

Ч: *Люди ненавидят роботов.*

Р: *Человек ненавидел мать[роботов]. Ранее человек обиделся на мать[роботов]. Ранее мать[роботы] сделала плохо человеку. Ранее мать[роботы] обидела человека.*

Ч: *Мама мыла картошку.*

Р: *Хозяйка[мама] чистила картошку. Ранее хозяйка[мама] купила картошку. Потом хозяйка[мама] приготовила картошку. Потом гости ели картошку.*

Ч: *Мама мыла раму.*

Р: *Инженер[мама] менял трубу[раму]. Ранее труба[рама] сломалась. Потом труба[рама] работала.*

Множество р-сценариев получено нами с помощью кластеризации семантических представлений 10 млн. высказываний, разобранных нашим семантическим парсером. Для каждой валентности выбран наиболее частый (прототипический) актант. Чтобы получить для сценариев связанные фреймы мы попросили информантов описать связанные ситуации: *Что случилось до этой ситуации? Что было потом? Что может означать эта ситуация?* Далее мы разобрали парсером полученные ответы и на основе их смыслов сформировали связанные фреймы для каждого р-сценария. Из приведённого диалога видно, что робот пытается восстанавливать намерения действующих лиц на основании д-сценариев: 'курица бегала, потому что потерялась, но потом спаслась', 'люди ненавидят роботов, значит роботы ранее сделали плохо людям'. Для механизма сценариев важно, что эмоционально значимые следствия добавляют эмоциональности исходной ситуации: высказывание *Мама мыла картошку* позитивно, потому что потом 'гости едят картошку' (робот отождествляет себя с 'гостями') – это позитивное событие, перехватываемое д-сценарием ВКУС.

В приведённом примере для каждого входящего высказывания демонстрируется выход только из одного р-сценария. Вместе с тем, ключевым методом нашей модели является использование нескольких сценариев, что должно повышать когнитивную сложность робота. Высказывание *Мама мыла раму* наиболее близко к сценарию 'инженер менял трубу' (близость 0,174) – это не вполне релевантный сценарий, но сценарий для ситуаций 'X помыл Y' → 'Y чистый' в нашем инвентаре пока отсутствует (Отдельная задача состоит в оценке того, какой набор сценариев является достаточным для разных уровней понимания текста на естественном языке. Мы предполагаем, что существующий инвентарь из 4500 сценариев должен быть увеличен, но целевые значения пока неясны). Это же высказывание также обладает близостью (0,162) к сценарию 'человек мыл перчатки' → 'человек нервничал'. Использование этого сценария позволяет роботу предположить, что 'мама моет раму, потому что нервничает' – эта интерпретация увеличивает вариативность понимания и повы-

шает когнитивную сложность модели. Также высказывание *мама мыла раму* близко (0,16) к сценарию ‘хозяйка чистила картошку’ → ‘хозяйка приготовила картошку’. Как видно по величине близости, это менее релевантный сценарий для исходной ситуации, но он может использоваться для перспективного механизма юмора: ‘мама мыла раму, значит, мама почистит раму, и гости будут есть раму’.

Дополнение когнитивной модели робота системой естественно-языкового вывода позволяет роботу интерпретировать действия героев текста, обогащать механизм реакций на входящее событие, создавать альтернативные понимания и, в перспективе, использовать их для сложных коммуникативных реакций и юмора.

Список литературы

1. Котов А. А. Механизмы речевого воздействия. М.: РГГУ. 2021.
2. Newell A., Simon H. Human problem solving. Englewood Cliffs, NJ: Prentice-Hall. 1972.
3. Schank R. C. Conceptual information processing. Amsterdam: North Holland. 1975.
4. Sloman A. Beyond shallow models of emotion // Cognitive Processing. 2001. V. 2. N. 1. P. 177–198.

УДК 159.9.072.533

Распределение внимания в ходе категориального научения у детей с РАС: план исследования

А. А. Котов, К. Н. Лужнова

*Научно-учебная лаборатория нейробиологических основ когнитивного развития, НИУ ВШЭ, Москва
e-mail: knluzhnova@edu.hse.ru*

Аннотация: В данной статье описывается план исследования распределения внимания в ходе категориального обучения у детей с расстройством аутистического спектра (РАС) по сравне-