

Текст сетевой дискуссии как источник сценарной информации¹

Кузнецова Ю.М.,

*кандидат психологических наук, старший научный сотрудник,
kuzjum@yandex.ru*

Осипов Г.С.,

доктор физико-математических наук, директор, gos@isa.ru

Смирнов И.В.,

кандидат физико-математических наук, заведующий отделом, ivs@isa.ru

Чудова Н.В.,

*кандидат психологических наук, старший научный сотрудник,
nchudova@gmail.com*

*Институт проблем искусственного интеллекта Федерального
исследовательского центра «Информатика и управление» РАН*

Аннотация

В работе рассматриваются возможности применения нового инструмента автоматического анализа текста — машины РСА — при исследовании сценариев поведения человека. С точки зрения реляционно-ситуативного анализа, реализованного в этом инструменте, сценарии рассматриваются как предикативно-синтаксемные структуры, определяющие состав действий, акторов и условия совершения действий в определенной ситуации. В качестве источника сценарной информации, циркулирующей на социальном уровне, предлагается использовать тексты дискуссий в социальных сетях.

Ключевые слова: поведение, сценарий, знаковая модель сознания, реляционно-ситуативный анализ, предикаты, синтаксемы

ВВЕДЕНИЕ

В настоящее время благодаря распространению электронных средств массовой коммуникации обмен и накопление вариантов поведенческих сценариев совершаются в масштабах поистине гигантских. Озвучивая свой вариант сценария поведения в определенной ситуации, человек решает важную психологическую задачу объективации своего знания, введения своего опыта в сферу социально значимых знаний; он получает возможность получить обратную связь (например, критику или благодарность) и содержательный разбор сообщаемой информации, найти последователей или сочувствующих и т.д. Исследователям сценарные тексты и комментарии

¹ Работа выполнена при поддержке РФФИ (№ 17-07-00651)

к ним дают возможность получить представление о функционирующих в социуме когнитивных репрезентациях проблемных ситуаций и проблеморазрешающего поведения. Возникающие в СМК обобщенные сценарные пространства предоставляют богатый материал для проведения исследований в интересах множества когнитивных дисциплин (лингвистики, психологии, социологии, компьютерных наук и т.д.).

Представление о сценарии в рамках модели знаковой картины мира

В [1] предложена модель знаковой картины мира субъекта деятельности. Данный подход базируется на представлениях отечественной школы психологии о роли знака в работе сознания и осуществлении целенаправленной деятельности.

Согласно Л.С. Выготскому [2] специфика человеческой психики связана с процессом опосредствования — между потребностями человека и его активностью по их удовлетворению «вклиниваются» знаки, в которых аккумулированы те значения предметов и действий, которые выработаны в данной культуре и восприняты («присвоены») человеком в ходе обучения, воспитания и развития. Именно этот опосредствованный культурными значениями характер человеческого знания делает человека хозяином своего поведения — произвольность в протекании психических процессов, позволяющая человеку удерживать внимание на некотором объекте, запоминать что-либо, избранное в соответствии со своими целями, размышлять об отвлечённых предметах, рефлексировать собственное эмоциональное состояние и управлять своими желаниями, то откладывая их удовлетворение, то и вовсе отказываясь от некоторых из них — вся эта произвольность, отличающая поведение взрослого человека от поведения животного или младенца, обеспечена использованием знаков (в первую очередь, естественного языка) как психологических орудий. А.Н. Леонтьев [3] развил это центральное положение культурно-исторического подхода Выготского, описав сознание как такую форму психического отражения, в которой образ действительности отделён от переживаний субъекта. Выражаясь языком искусственного интеллекта, сознание — это способ представления знаний, в котором элементом является знак. Знак как элемент сознания содержит в себе знания трёх типов: знание о предназначении предмета, в котором аккумулирован общественно-исторический опыт его использования; знание о свойствах этого предмета, возникающее либо в акте восприятия, опосредствованного такими категориями как форма, цвет, вес и проч., либо в акте мышления, опосредствованного абстрактными понятиями; наконец, знание о смысле этого предмета в жизни субъекта, в котором в таком значении-для-себя аккумулируется личный опыт взаимодействия с предметом. А.Р. Лурия уточнил представление о знаковом опосредствовании, описав развитие имени знака как элемента языковой системы в детском возрасте и его разрушение при мозговой патологии [4]. В процессе развития индивидуального сознания имя знака видоизменяется, приобретая морфологические признаки и синтаксические свойства, отвечающие тому значению, тому образу и тому смыслу, которые обозначаемый объект или действие приобрели для ребёнка в ходе

его освоения предметной действительности. При различных нарушениях мозгового функционирования процесс разрушения имени идёт по «швам» знака, отражая трёхаспектное его строение: утрачивается то способность связывать имя с образом, то способность использовать предмет по его назначению, то утрачивается память о личном опыте взаимодействия с предметом.

Сама необходимость в знаковом опосредствовании и в сознательных представлениях возникает, согласно культурно-историческому подходу Выготского-Леонтьева-Лурии, в ситуации совместной деятельности. Именно необходимость разделения ролей ради достижения ценного для всех результата вынуждает человека пользоваться языком (шире — знаковой системой) как средством, позволяющим построить «вторую реальность» (термин А.Р. Лурии) воображаемого итога совместных усилий. На этой «сценической площадке» (термин А.Н. Леонтьева) для субъекта деятельности разыгрывается желаемая ситуация: участники совместной деятельности, представленные именами ролей (не реальные Вася и Петя, а «загонщик» и «сидящий в засаде»), совершают действия, совместные результаты которых приводят к тому, что и потребность самого субъекта и потребности его партнёров удовлетворяются (загонщик загоняет дичь к месту, где в засаде сидит стрелок, в результате чего добытая дичь достаётся обоим). В современных когнитивных исследованиях, выполненных на шимпанзе и детях полутора-двух лет, эти представления получили экспериментальное подтверждение. Показано, что совместная деятельность ребёнка и взрослого основана на присущей только человеку способности к разделению ролей и умению воспринимать действия в структуре деятельности как выполняемые независимо от актуальной потребности участника событий. «Разделённая намеренность» в терминологии М. Томаселло — это специфичная для людей готовность создавать общее для участников ситуации представление о желаемом исходе совместных усилий. Этой форме намерений соответствуют особые когнитивные средства: «диалогические когнитивные репрезентации — это онтогенетический продукт, возникающий в результате уникальной, присущей только человеку биологической адаптации к основанным на разделённой намеренности социальным взаимодействием» (цит. по [5]). Знаковое представление как средство абстрагирования позволяет создавать сценарий полезной для любых конкретных участников совместной деятельности; именно знаковое опосредствование позволяет уже при реализации сценария каждому выполнять свою роль, а не реагировать на происходящее в соответствии с актуальной потребностью (исполняющий роль загонщика отгоняет от себя дичь, а не преследует её; участник охоты, сидящий в засаде, терпеливо её дожидается, а не ищет по лесу).

Построенная на этих теоретических положениях модель знака включает в себя представление о существовании у любого не вырожденного знака четырёх компонентов — имени, значения, образа, смысла. Соответственно, представление знаний в знаковой картине мира интеллектуального агента реализуется как система четырёх типов неоднородной семантической сети, в которой работа трёх сетей — сети значений, сети образов и сети смыслов — опосредствована работой на сети имён [6]. Таким образом, умение оперировать именами и анализировать текст оказывается необходимой функцией модели знаковой картины мира. Именно в этом контексте в данной работе нас будет интересовать проблема извлечения сценариев из текстов.

В картине мира конкретного человека, как правило, один из компонентов знания имеет преимущество и является источником так называемых фиксированных

познавательных гипотез [7]. Проще говоря, одни склонны ориентироваться на житейское знание с его культом здравого смысла и умением опираться на норму, традицию, культуру, другие больше доверяют научному знанию, где процедуры описания и сравнения свойств предметов отрефлексированы, стандартизированы и претендуют на выявление закономерностей объективного характера, третьи всегда задаются вопросом о смысле происходящего и видят мир как бы глазами действующих субъектов. При моделировании знаковой картины мира эти различия требуют и использования разных методов. Психологический анализ позволяет утверждать, что значение предмета, презентующее обобщённый опыт его использования в данной культурно-бытовой традиции, при развёртке в сознании человека представляется в форме такой структуры как сценарий, образ предмета — как способ решения интеллектуальной задачи, смысл — как нарратив, в котором главным героем выступает субъект деятельности [8]. Таким образом, любое представление знаний в знаковой картине мира имеет форму текста, но сами типы текстов оказываются разными.

В данной работе нас будет интересовать такой тип текста, в котором эксплицирован локализованный в сознательных представлениях отдельных людей общественно-исторический опыт совместной деятельности, т.е. такая модель реальности, которая фиксирует схему совместной деятельности с присутствующими ей ролями исполнителей и ролями предметов, которыми они оперируют. Другими словами, мы хотим иметь возможность реализовывать процедуру планирования сценария на сети значений и для этого хотим иметь возможность извлекать сценарии из текстов в такой форме, которая допускает последующее использование извлечённой схемы на сети имён.

Уточним понятия.

Под сценарием в концепции знаковой картины мира мы понимаем схему полимотивированной групповой деятельности, осуществляемой в трёх сферах: физической, коммуникативной, ментальной.

Под деятельностью мы, вслед за А.Н. Леонтьевым, понимаем активность, направленную на удовлетворение потребности, представленной конкретным предметом. Предмет, удовлетворяющий потребность, в деятельностном подходе принято называть мотивом; для нас в данном случае важна именно полимотивированность деятельности, её способность удовлетворить сразу несколько потребностей и каждую — с помощью не одного, единого для всех участников, мотива, а с помощью различных, уникальных для каждого субъекта деятельности, предметов. Именно множественность мотивов в групповой деятельности даёт людям возможность выбирать себе разные роли в этой деятельности и «подгонять» исполнение роли под свои возможности и требования. И именно эта возможность исполнять разным людям одну и ту же роль, а одному человеку — разные роли в разное время порождает, на наш взгляд, сам концепт сценария. Употребление этого концепта по отношению к бытовым ситуациям основано на представлении об уникальных событиях жизни отдельного человека как о типовых, случающихся и в его жизни неоднократно, и регулярно происходящих с другими

людьми. Именно это ощущение закономерности в смене сцен-действий, обязательности определённого взаимодействия с «партнёрами по сцене», типичности некоторых переживаний и мыслей для исполнителей ролей, и отражается в понятии сценария.

Групповая деятельность, по мысли А.Н. Леонтьева, является генетически исходной для сознательной деятельности. Представление будущих промежуточных итогов активности в сознании, на сценической площадке воображения, необходимо в том случае, если достичь удовлетворения нельзя одним усилием, с помощью активности одного типа и одного направления. *Есть малину* в малиннике можно, не ставя себе целей и не планируя ничего заранее, а вот *собирать малину* «не приходя в сознание» уже невозможно — необходимо запастись корзинкой и нужно не отвлекаться на желание отправить ягоду в рот, а не в корзину. Вот эта форма активности, ориентированная не на удовлетворение потребности, а лишь на достижение промежуточных целей (*запастись корзиной, собирать в корзину, не есть ягоду сразу*) получила название действия. В антропогенезе и в онтогенезе действие как активность по достижению представленной в сознании цели сначала возникает в совместной деятельности (трудовой и детско-родительской, соответственно) и лишь затем этот механизм организации сложной многосоставной деятельности формируется уже как монособъектный, в котором все роли исполняются последовательно одним человеком.

Действия, таким образом, это та активность, которая имеет представленность в сознании (в виде цели действия), и, следовательно, действия — это то, что называется в текстах и обсуждается в дискуссиях (например, сетевых), когда обсуждаются сценарии как нормативные, «правильные» с точки зрения социума, схемы событий. Поэтому анализ текста, проводимый в интересах выявления сценария, очевидно связан с исследованием предикативной структуры, однако он не может этим исчерпываться. Любое действие осуществляется в конкретных условиях, и эти условия учитываются с помощью третьей формы активности, выделяемой деятельностным подходом — с помощью операций. Операция — это активность релевантная условиям, в которых дан мотив. Операциональный состав деятельности определяется как внешними условиями, так и внутриличностными. Именно набор операций, реализующих то или иное действие, есть реальность, которая дана человеку в его непосредственных ощущениях, и именно эта реальность обобщается им с помощью слова, обозначающего действие/цель. Таким образом, сценарный анализ текста должен включать в себя такой предикатный анализ, который представляет действие не в его абстрактной форме, а в виде веера возможных (возможных с точки зрения достижения цели) конкретных воплощений.

Формы выражения сценарной информации

В когнитивной лингвистике сценарий определяется как «концептуальная структура для процедурного представления знаний о типизированной ситуации. В отличие от схемы, содержащей, как правило, лишь общий алгоритм действий, в сценарии знание привязывается к конкретным участникам ситуации — актантам» [8]. В работе [9] проблема языкового выражения сценариев поведения рассматривается в контексте уже упоминавшейся знаковой модели сознания [1]. Были выделены следующие общие характеристики вербально выражаемых поведенческих сценариев: 1) культурное происхождение и связь с языковой картиной мира; 2) функция социального контроля и регуляции поведения; 3) обязательное наличие темпоральных,

акциональных и ролевых компонентов; 4) иерархичность структуры, в которой вышележащие компоненты отражают стабильные признаки типовой ситуации, а нижележащие заполняются в зависимости от конкретных условий реализации сценария; 5) сосуществование прототипического и отклоняющихся вариантов.

Конкретные речевые формы сценарной информации могут быть разными. Так, обязательные для исполнения *сценарии-инструкции* и *сценарии-ритуалы* в типичном случае включают в себя признаки ситуации, в которых их следует исполнять, и перечисление предписаний, требований и запретов. Обязательность исполнения такого рода сценариев определяется их прагматической (инструкции) или социальной (ритуалы) целесообразностью, нарушение которой ведет либо к недостижению желательного результата, либо к социальному порицанию. Жесткость предписаний может варьировать: в предельном варианте выполнение сценария строго требуется (*Не стой под стрелой!*), а в других случаях определение ситуации как соответствующей определенному сценарию и выбор конкретных средств и способов его реализации принадлежит самому человеку (*Вы можете сами определить последовательность ремонта квартиры, руководствуясь целесообразностью, планировкой, сроками и финансовыми возможностями. Главное условие: последующий шаг не должен приводить к исправлениям уже того, что было сделано на предыдущем этапе* — из «Советов по ремонту квартир»). Свобода выбора в последнем случае отражает не критичность ситуации, не требующей быстрой и однозначной реакции, и, наоборот, допускающей не только пробующее поведение с перебором разных вариантов, но и поведение поисковое, связанное с выработкой совершенно нового для субъекта способа поведения. Речевым выражением такого рода необязательных сценариев могут выступать *советы, предложения, описания собственного опыта* и вообще «*случаев из жизни*».

Компонентом высшего уровня развернутого сценария можно считать его наименование, которое отражает его сущность, служит средством идентификации сценария и, тем самым, открывает доступ к информации о его содержании. Например, глагол *упрашивать* можно использовать для названия сценария, где один из участников общения старается уговорить слушателя сделать что-либо против своей воли, а сценарные условия таковы, что тактика разговора с позиции силы исключается по ряду причин [10]. В конфликтной ситуации возможна реализация таких сценариев, как: 1) *подыгрывание партнеру* с целью не обострять отношения, не выносить имеющиеся разногласия или противоречия на открытое обсуждение, не выяснять отношения; здесь применяются тактики согласия, уступки, одобрения, похвалы, обещаний и т.п.; 2) *игнорирование проблемы*, когда коммуникант применяет тактики умолчания, ухода от темы или смены сценария; 3) *интересы дела прежде всего*, при котором коммуниканты стремятся к выработке взаимоприемлемого решения, пониманию и компромиссу и используют тактики переговоров, уступок, советов, согласия, просьбы и т.п. [11].

Вне зависимости от характера сценария, в его речевой организации особую роль играют предикативные структуры, определяющие характер

действий, указывающие на участников и условия осуществления действий. В семантике глагола, — пишет Е.С. Кубрякова, — имплицированы такие компоненты событий, как причина и результат, цель и способ совершения, участники (тот, кто выполняет действие или тот / то, на кого / что оно переходит) и т.п.; значение глагола может быть понято как инструкция к осуществлению особого действия. По словам Л. Теньера, «глагольный узел», являясь центром предложения, выражает стоящую за этим предложением «маленькую драму», ибо в нем обязательно имеется действие, а чаще всего также действующие лица и обстоятельства [12]. Таким образом, анализ предикатов и связанных с ними в конкретном высказывании вербальных единиц позволяет восстанавливать основные компоненты поведенческого сценария, определяемые в терминах функциональной типологии синтаксиса А.Г. Золотовой [13]. Материалом для такого рода анализа в первую очередь могут быть тексты, созданные со специальной целью передачи сценарной информации.

Нужно отметить, что в наше время благодаря распространению электронных средств массовой коммуникации обмен и накопление вариантов поведенческих сценариев совершаются в масштабах поистине гигантских. Озвучивая свой вариант сценария поведения в определенной ситуации, человек решает важную психологическую задачу объективации своего знания, введения своего опыта в сферу социально значимых знаний; он получает возможность получить обратную связь (например, критику или благодарность) и содержательный разбор сообщаемой информации, найти последователей или сочувствующих и т.д. Исследователям сценарные тексты и комментарии к ним дают возможность получить представление о функционирующих в социуме когнитивных репрезентациях проблемных ситуаций и проблеморазрешающего поведения. Возникающие в СМК обобщенные сценарные пространства предоставляют богатый материал для проведения исследований в интересах множества когнитивных дисциплин (лингвистики, психологии, социологии, компьютерных наук и т.д.).

Пример анализа сценарного пространства сетевой дискуссии

В качестве примера текста, содержащего описание житейского нежесткого сценария, рассмотрим размещенный на портале picasu.ru пост под названием «Счастливым день», который вызвал оживленное обсуждение (за сутки появилось более 600 комментариев). Ключевым предикатом в тексте является глагол *разбогател*, появившийся в первой же фразе: *Однажды в детстве я сказочно разбогател*. Этот глагол может быть использован в качестве названия для сценария в целом. Далее описываются конкретная история «разбогатения»: *Пошел звонить в автомат, он мои две копейки съел, не позвонил, а монетку не возвращает. Присмотрелся, а там двушка заклинила язычок, который закрывает окошко возврата монет. Стал я её ногтем выковыривать, куда там! Видно, что до меня кто-то долго монету ножом пытался выскоблить, так, что на той части которая была видна живого места не осталось. Плюнул я и со словами "Так не доставайся же ты никому!" втокнул монетку внутрь. Язычок тут же встал на место, а когда я на него чуть нажал, оттуда двушки посыпались. Целая река! Я ладошки подставил, так в них едва эта гора монет поместилась. Пошел накупил себе всяких вкусняшек. День удался!*

В тексте представлено описание эмоционально насыщенного события, разворачивающегося как сценарий, то есть, последовательных изменений ситуации, частично

определяемых действиями автора, частично — случайностью. Эмоциональная значимость описываемого события выражается с помощью таких средств, как интонация (*куда там! Целая река! День удался*), эмотивная и оценочная лексика (*сказочно, живого места не осталось, плюнул, река, гора, вкусняшки, удался*), а также обращение к прецедентным текстам (*Так не доставайся же ты никому!*). Выделение предикатных структур, относящихся к действиям автора-субъекта, дает следующие конструкции: *пошел звонить в автомат, стал выковыривать (двушку) ногтем, втолкнул монетку внутрь, нажал (на язычок), подставил ладошки, пошел накупил вкусняшек*. Другими важными «участниками» ситуации явились: *автомат (съел две копейки), двушка (заклинила), двушки-монеты (посыпались, едва поместились в ладошках)*.

В комментариях к инициативному тексту представлены альтернативные сценарии. Их можно выделять из общей массы комментариев (довольно часто отклоняющихся от заданной инициативным текстом темы) по тому, что в них имеются вербальные единицы, либо соответствующие базовыми компонентам исходного текста (*я — я, купил — купил, автомат (телефонный) — автомат (игровой)*), либо находящиеся с компонентами исходного текста в парадигмальных отношениях (близость значения: *река — поток, купил — потратил; вкусняшки — колбаса или антонимичность: разбогател — проиграл, остался ни с чем, долг, обокрал, воровство; вкусняшки — голод*).

При сохранении общего смысла и эмоциональной оценки происходит:

- замена «источника обогащения», в качестве которого в комментариях выступает игровой автомат: *...И из него посыпался джекпот... реально прям поток, аж на пол вываливалось, не успевал выгрести в карманы [K1], ...выпало 777 вроде, деньги рекой посыпались, пакет пришлось подставить [K16]* т.п.;
- замена предмета, на который было потрачено «богатство»: *Купил еды на неделю и даже на 1,5 л Жигулевского хватило [K2], ...на все деньги купил сувенирного морского ежа [K16], ...накупила всякой сырокопченной колбасы [K215]* и т.п.;
- меняется эмоциональный знак события: *А дома меня ждал серьезный разговор на тему азартных игр [K3], ...потом неделю ждал, что за мной менты придут [K334]*;
- меняется исход: *выходя из магазина с полными карманами пятак, мне в голову пришла гениальная мысль, мол, имея столько денег, я могу выиграть еще больше. Через 10 минут я покинул магазин ни с чем [K78], А я вот подсел, столько бабла проиграл [K64], кинул туда деньги, ходил потом весь день голодный (109), ...никаких выигрышей, лишь кредит на 500 тысяч, чтобы закрыть то, что набрал в долг для таких вот «игрулек» [K210]* и т.п.;
- меняется исход и состав активных участников событий: *тетка выбежала, начала орать и отбирать деньги. Сзади подошел взрослый высокий*

мужик, взял ее за шкурку, кивнул нам, мы набрали сколько успели и убежали (K22), ...как только я потянул ручки к богатству, начала на меня орать, а потом вышвырнула пинками из магазина. Остальные взрослые просто стояли и смотрели, никто не попытался помочь (K152) и т.п.;

- *меняется общая оценка ситуации: Вот да, обокрал телефонного оператора и даже не задумался о том, что это плохо. Еще и хвастается этим (K223), А чего все радуются и плюсы ставят? Ведь это же чистой воды воровство (K357) и т.п.*

Таким образом, данное сетевое обсуждение содержит варианты сценария «Разбогател», в котором в качестве субъекта выступает персонаж Я (авторы инициативного текста и комментариев), в качестве источника «богатства» — автомат (часть авторов комментариев, не знакомых с устройствами старых таксофонов, посчитали, что речь идет об игровом автомате), а в качестве его применения — покупка каких-то привлекательных предметов. В типичном случае события происходили в детстве авторов и переживались как приятный подарок судьбы. Альтернативными к исходному сценарию в основном являются варианты, в которых появляются либо дополнительные участники, которые могут помогать или мешать получению «богатства», либо инвертируются исход (потеря вместо неожиданного приобретения) или оценка исхода (социальное порицание вместо эгоцентрического удовольствия). Встречается также перенос времени события из детства в юношеский возраст или взрослость.

В ходе обсуждения усилиями коммуникантов выстраивается сценарное пространство, в котором отражаются типовые сценарные ходы, связанные с обсуждаемой ситуацией, и, благодаря этому, позволяющее судить о групповых инвариантах поведения и эмоциональных реакций на эту ситуацию. Предварительный анализ показывает, что в достаточно объемном сетевом обсуждении на определенном этапе перестают появляться новые варианты сценариев, и вновь вступающие в общение коммуниканты либо повторяют с незначительными изменениями уже имеющиеся сценарии, либо организуют отдельные ветви, где обсуждаются нерелевантные исходному тексту вопросы. По-видимому, наличие такого порога может служить признаком определенной полноты перечня представленных в дискуссии сценарных вариантов, то есть, того, что исследователи располагают исчерпывающей информацией о групповой специфике процессов идентификации ситуации, выделения ее компонентов, планирования и оценки результатов действий.

Технически анализ сетевых дискуссий, направленный на выявление сценарных пространств, может строиться на основе выделения в исходном сценарном тексте таких предикативных структур, которые относятся к субъектам подразумеваемых сценарием действий, и затем поиска в текстах дискуссии вербальных единиц, связанных с компонентами исходного сценария парадигмальными отношениями. Выявленный таким образом комплекс взаимодополняющих предикатно-синтаксемных структур и будет отражать круг вариантов сценариев поведения в определенной ситуации, характерных для группы участников обсуждения.

Реляционно-ситуационный анализ как инструмент сценарного анализа

Для моделирования сценариев при автоматическом сценарном анализе целесообразно использовать предикатно-аргументную семантику, которая представляет

высказывание на естественном языке в виде действия, события или ситуации, которые выражаются предикатами, и множеством участников этих событий или ситуаций, каждый из которых играет свою роль. Такая модель в принципе не зависит от предметной области и способна описать почти любое высказывание на естественном языке.

В [14] была предложена реляционно-ситуационная модель текста, на основе которой был создан метод реляционно-ситуационного анализа (РСА) [15]. Эта модель основана на коммуникативной грамматике [16], описывающей синтаксические средства выражения в языке ситуаций и ролей участников, и понятии синтаксемы. Синтаксемой называется минимальная синтактико-семантическая единица языка, несущая обобщенный категориальный смысл и характеризующаяся взаимодействием морфологических, семантических и функциональных признаков. В разных языках синтаксема выражается по-разному, в русском она задается предлогом, падежом и категориальным классом существительного, например, синтаксема «с помощью метода» определяется предлогом «с помощью» с родительным падежом и означает инструмент, которым выполняют какое-либо действие. Но при восприятии человеком текста ему недостаточно лишь понимания значений синтаксем — он опирается на интуитивные представления об их сочетаемости друг с другом. В РСА эта сочетаемость определяется семейством бинарных отношений на множестве синтаксем, так что семантические отношения понимаются как логические связи между денотатами тех имён, которые содержатся в тексте. Например, каузальная связь, один компонент которой обозначает причину проявления другого компонента спустя какое-то время (*Казнокрадство приводит к обнищанию населения; Длительные периоды гипертензии приводят к нарушениям диастолической и систолической функции миокарда*) распознаётся читателем в тексте за счёт того, что существующая у него когнитивная схема каузации *X приводит к У* в тексте выражена с помощью определённых семантических ролей конкретных имён и логических отношений между ними. Собственно говоря, именно опора на семантически определённые связи между упомянутыми в тексте объектами и позволяют человеку при чтении воспринимать не текст, а реальность, которую тот описывает. Так что реляционно-ситуационная модель представляет собой модель понимания человеком текста и позволяет представить текст не как «вещь в себе», а как средство описания реальности. В РСА семантика высказываний и текста, состоящего из высказываний, представляется в виде семантической сети, в узлах которой находятся предикатные слова и синтаксеммы, а в качестве ребер выступают синтаксические и семантические связи [14], [17]. Таким образом, с помощью РСА мы можем получить математическую (графовую) модель сценария поведения, где элементарные акты поведения представлены предикатно-синтаксемными структурами, а их сочетание образует последовательность сцен.

Реляционно-ситуационная модель текста строится в результате реляционно-ситуационного анализа текста, включающего морфологический и семантико-синтаксический анализ. Последний включает в себя выделение в тексте предикатно-аргументных структур и установление

семантических ролей аргументов. Эта задача известна как установление семантических ролей (semantic role labeling).

Основным источником информации для семантического анализа в методе реляционно-ситуационного анализа является словарь предикатных слов. Он содержит словарные статьи, каждая из которых соответствует некоторой ситуации. В словарной статье содержатся предикатные слова (глаголы, причастия, деепричастия, девербативы), которые указывают какую ситуацию выражает данное предложение. Помимо этого, словарная статья содержит семантические значения, которые могут быть назначены синтаксемам в заданной ситуации, а также признаки, которыми должны обладать синтаксеммы, чтобы получить семантические значения. В качестве основных признаков используются падеж, предлог, а также семантический класс существительного.

Процедуру семантического анализа предложения можно условно разделить на четыре этапа: поиск предикатных слов и соответствующих им словарных статей, поиск потенциальных аргументов, установление аргументам семантических значений, выбранных из каждой словарной статьи, выбор наилучшей словарной статьи и семантических значений. Возможно и применение методов машинного обучения для установления семантических ролей, описанное подробно в работе [18].

В завершение соотнесём существующие возможности метода РСА и потребности сценарного анализа.

Именно РСА позволяет выделять операциональный состав деятельности: предикаты с заполненными семантическими ролями это и есть операции в психологическом смысле слова. Соединение данных РСА со слотами сценария в смысле Р. Шенка [19] и Д.А. Пospelова [20] в форме заполнения слотов сценария данными РСА и будет представлением сценарной информации — описанием действий исполнителей, их мотивов и условий исполнения действий.

Для реализации процедуры заполнения слотов сценария данными РСА необходимо привлечение понятия коммуникативного регистра Г.А. Золотовой [21], поскольку в разных регистрах заполнение слотов сценария предикатами и синтаксемами происходит по-разному.

Для текста в репродуктивном регистре, где событие описывается из позиции включённого наблюдателя, имена конкретны (не роль загонщика, а поведение Пети, не роль дичи, а описание конкретного зверя) и обстоятельства времени и действия — конкретны, нужна процедура абстрагирования. Именно эту функцию здесь выполняет РСА. Но дальнейшее обобщение данных РСА требует уже либо применения какой-либо классификации предикатов (см., например, [14], [22]), либо классификации объектов, выступающих в роли предмета внимания наблюдателя. Так что выделение сценария из текста репродуктивного регистра будет определяться задачами исследователя — то, сценарий какой именно деятельности его интересует, будет определять классификацию предикатов и/или объектов.

Текст в информативном регистре как сообщении об осмысляемом сторонним наблюдателем событии уже содержит абстрагирование и обобщение. В случае дискуссии при наличии нескольких текстов в информативном регистре, РСА даёт информацию о том, как одно и то же событие обобщается разными наблюдателями. Это

обеспечивает возможность выявлять общее в сценарии (одни и те же предикаты, одинаковое заполнение семантических ролей) и различное.

В генеративном регистре, в котором человек выдаёт сентенции максимально обобщённого характера («афоризмы, пословицы, общечеловеческие истины»), извлечение сценария возможно только по отношению к типам деятельности (*работа не волк, в лес не убежит, терпение и труд всё перетрут, возлюби ближнего как самого себя, единственная настоящая роскошь — это роскошь человеческого общения, никто не забыт и ничто не забыто* и т.п.).

Тексты в волюнтивном регистре, в котором человек формулирует пожелание/требование к адресату, могут содержать элементы сценария (*вот сами и подавайте в суд на мэра, а я на Вас посмотрю! ; подай мне корзину, за малиной схожу* и т.п.). Использование текстов этого регистра для извлечения сценария требует дополнительной работы по автоматизации разделения коммуникативных действий и собственно описания/называния сценария.

Реактивный регистр выражает реакцию на коммуникативную ситуацию. Этот тип текстов для сценарного анализа интересен только в плане выделения сценариев одного типа — сценариев коммуникации.

Заключение

Анализ поведения человека с позиций деятельностного подхода, предложенного А.Н. Леонтьевым, позволяет выделить в качестве психологической основы деятельности культурно и ситуативно опосредованные ментальные схемы, или сценарии. Культурные по своей природе, сценарии имеют вербальную форму существования, что позволяет применять для их исследования методы анализа текстов. Особое значение при этом имеют тексты сценарного содержания, специально создаваемые в ходе общения в социальных сетях, и благодаря этому доступные для сбора и анализа. В настоящей работе предлагается принципиальная схема выявления сценарной информации, основанная на применении метода реляционно-ситуативного анализа (РСА).

В качестве компонентов сценария предлагается рассматривать ключевые предикатно-синтаксемные структуры, автоматически выявляемые с помощью машины РСА в текстах сетевых дискуссий и отражающие представления о действиях, действующих лицах и важнейших условиях ситуативного поведения. Совокупность предложенных в дискуссии, посвященной поведению в определенной ситуации, сценариев предлагается рассматривать как сценарное пространство, которое на групповом уровне объединяет социальные представления о правильном и отклоняющемся поведении, а для отдельного человека представляет собой информационный ресурс в случае отсутствия собственного опыта.

Литература

1. *Осипов Г.С., Панов А.И., Чудова Н.В., Кузнецова Ю.М.* Знаковая картина мира субъекта поведения. М.: Физматлит, 2018
2. *Выготский Л.С.* Мышление и речь. — Изд. 5, испр. — М.: Лабиринт, 1999
3. *Леонтьев А.Н.* Проблемы развития психики. — 4-е издание. — Москва : Издательство Московского университета, 1981. — 584 с.
4. *Лурия А.Р.* Язык и сознание. / Под редакцией Е. Д. Хомской. — М: Изд-во Моск. ун-та, 1979 — 320 с.
5. Когнитивная психология. История и современность. Хрестоматия.: пер. с англ. Под ред. М. Фаликман и В. Спиридонова. — М.: Ломоносовъ, 2011. — 384с, с.90
6. *Брунер Дж.* Психология познания. За пределами непосредственной информации. М.: Прогресс, 1977
7. *Ананьева М.И., Кузнецова Ю.М., Чудова Н.В.* Сценарий как форма репрезентации события в знаковой картине мира // Труды ИСА РАН, 2018. №4
8. *Баранов А.Н., Добровольский Д.О.* Понятие речевой формулы: определение и типология // Русский язык сегодня. Сб. статей. Вып. 2 /Отв. ред. Л.П. Крысин. М.: Азбуковник, 2003
9. *Кузнецова Ю.М.* Сценарный подход к анализу текстов // Труды ИСА РАН. 2018. Т. 68. Вып. 1. С. 31–41.
10. *Шляхов В.И.* Сценарии русского речевого взаимодействия // Русский язык за рубежом. 2007. № 2. С. 26–34
11. *Третьякова В.С.* Конфликт в лингвистических категориях // Юрислингвистика. 2010. № 10. С. 141–149
12. *Кубрякова Е.С.* Язык и знание: На пути получения знаний о языке: Части речи с когнитивной точки зрения. Роль языка в познании мира. М.: Языки славянской культуры, 2004. С. 258, 265–266
13. *Золотова Г.А.* Синтаксический словарь: Репертуар единиц русского синтаксиса. М.: Едиториал УРСС, 2006
14. *Осипов Г.С.* Приобретение знаний интеллектуальными системами: Основы теории и технологии. М.: Наука, Физматлит, 1997
15. *Osipov G.S., Smirnov I.V. and Tikhomirov I.A.* Relational-situational method for text search and analysis and its applications // Scientific and Technical Information Processing, Springer-Verlag New York, Inc. Secaucus, NJ, USA, Dec 2010. — Volume 37, Number 6, 432–437
16. *Золотова Г.А., Онипенко Н.К., Сидорова М.Ю.* Коммуникативная грамматика русского языка. М.: Ин-т рус. яз. РАН им. В.В. Виноградова, 2004
17. *Smirnov Ivan, Tikhomirov Ilya.* Heterogeneous semantic networks for text representation in intelligent search engine EXACTUS. //Proceedings of workshop SENSE'09 — conceptual Structures for Extracting Natural language SEmantics, The 17th International Conference on Conceptual Structures (ICCS'09), July 2009, Moscow, Russia. P.1–9.
18. *Shelmanov A.O., Smirnov I.V.,* Methods for Semantic Role Labeling of Russian Texts // Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue" (2014). Issue 13 (20). — 2014. — pp. 580–592.
19. *Шенк Р.* Обработка концептуальной информации. Пер. с англ. — М.: Энергия, 1980. — 360 с
20. *Поспелов Д.А.* Представление знаний. Опыт системного анализа //Системные исследования. Методологические проблемы. Ежегодник / Под ред. Д.М. Гвишиани, В.Н. Садовского. №17, М.: Наука, 1985, с.83–102
21. *Золотова Г.А.* Роль грамматики в композиции текста // Язык. Культура. Гуманитарное знание: Научное наследие Г.О. Винокура и современность. М.: Научный мир, 1999
22. *Кузнецова Ю.М., Смирнов И.В., Исаков В.А., Станкевич М.А., Чудова Н.В.* Создание инструмента автоматического анализа текста в интересах социо-гуманитарных исследований. Ч.2. Машина РСА и опыт её использования // Искусственный интеллект и принятие решений (в печати)

THE NETWORK DISCUSSION TEXT AS A SOURCE OF SCENARIO INFORMATION

Kuznetsova Y.M.,

PhD in Psychological Sciences, Senior Researcher,

Osipov G.S.,

PhD in Mathematics, Head of IPAI

Smirnov I.V.,

PhD in Mathematics, Head of Department

Chudova N.V.,

PhD in Psychological Sciences, Senior Researcher,

Institute for Problems of Artificial Intelligence, Federal Research Center

“Computer Science and Control” of Russian Academy of Sciences,

Moscow

Abstract

The paper discusses the suitability of a text relational-situational analysis (RSA) for study the behavioral scenarios. In the relational-situational text model scenarios are considered as predicative-syntaxeme structures that combine the situational composition of actions, actors and conditions. Texts of discussions in social networks are proposed to be used as a source of information about group shared scenarios.

Keywords: behavior, the script, sign model of consciousness, relational-situational analysis, predicates, syntaxeme