

с нормативными решениями: «в случае p , Oa », где a – действие¹³⁹.

Интересно также, что Г. Х. фон Вригт в одной из более поздних работ решительно порывает со своим влиятельным, ориентированным на модальную логику подходом 1951 г., и заявляет, что «формальная логика, как мы ее знаем сегодня, по сути есть логика статичного мира. ...Для изменений нет места в этом мире. ...Действия же существенным образом связаны с изменениями»¹⁴⁰. Переход к деонтической логике, в которой нормы налагаются на действия, изменяющие (сохраняющие) положения дел¹⁴¹, требует новых семантических подходов, определяющих норму, например, на основании аксиологического сравнения (предпочтения) ситуаций¹⁴²: Oa – действие a обязательно, если только его выполнением можно избежать вреда, если, таким образом, воздержание от него ведет к ухудшению ситуации; Fa – действие a запрещено, если его выполнение вызывает ухудшение ситуации; Pa – действие a позволено, если оно своим выполнением не ухудшает ситуацию, стало быть не следует опасаться его негативных последствий; Ia – действие a нормально безразлично, если ни его выполнение, ни его невыполнение не изменяют оценку ситуации.

НОРМЫ И ЦЕННОСТИ В КОГНИТИВНОЙ СТРУКТУРЕ АГЕНТА

А. Ю. Комаров

*аспирант 1 курса по направлению «Математика и механика»
Института математики и компьютерных наук
Уральского федерального университета
имени первого Президента России Б. Н. Ельцина,
г. Екатеринбург*

Специалисты в области искусственного интеллекта в последние годы уделяют значительное внимание понятию агента,

¹³⁹ Альчуррон К. Э., Булыгин Е. В. Нормативные системы // «Нормативные системы» и другие работы по философии права и логике норм. СПб.: Изд-во С.-Петербург. ун-та, 2013. С. 53–98; Frassen B. C. van. The Logic of Conditional Obligation // J. of Philosophical Logic. 1972. № 1. P. 417–438; Hansson B. An Analysis of Some Deontic Logics // Deontic Logic. Dordrecht; Boston; L., 1971. P. 121–147.

¹⁴⁰ Wright G. H. von. Norm and Action, A Logical Enquiry. London, 1963. P. 15.

¹⁴¹ Meyer J.-J. Ch. A Different Approach to Deontic Logic: Deontic Logic Viewed as a Variant of Dynamic Logic // Notre Dame Journal of Formal Logic. 1988, Vol. 29. № 1. P. 109–136; Кислов А. Г. Динамическая логика и деонтические операторы «в строгом смысле» // Философия науки. 2012. № 3(54). С. 65–80.

¹⁴² Кислов А. Г. Динамический подход к деонтической логике: семантика нормативных операторов // Научный ежегодник Института философии и права Уральского отделения Российской академии наук. 2013. Т. 13. № 3. С. 20–35.

и особое – понятию когнитивного агента¹⁴³. Под *агентом*, как правило, понимается искусственная система, осуществляющая некоего рода деятельность, а под *когнитивным агентом* – агент, деятельность которого осуществляется с опорой на познавательные процессы. Аксиологическая проблематика в применении к искусственным агентам рассматривается чаще в прикладном, чем в теоретическом аспекте. К немногим исключениям можно отнести усилия по созданию общих кодексов поведения роботов, начиная с «законов робототехники» А. Азимова и заканчивая попыткой формулировки норм поведения боевых роботов, предпринятой учеными Калифорнийского университета несколько лет назад¹⁴⁴. Однако, поскольку подобные кодексы общего плана неизбежно формулируются в понятиях, с трудом поддающихся формализации, то попытки использовать эти кодексы на практике терпят неудачу. Так, военные ведомства ведущих держав в последнее время заявляют, что вынуждены отказаться от разработок концепций применения автономных боевых роботов ввиду непреодолимых трудностей, возникающих при попытках формализации аксиологии этой предметной области¹⁴⁵.

В некоторых прикладных задачах формализация, тем не менее, возможна. Например, при создании беспилотных транспортных средств используются нормы, базирующиеся на ПДД, формализованное представление которых не представляется чем-то недостижимым. Однако, и в этих случаях трудоёмкость зачастую настолько велика, что построение систем, пригодных для использования в реальных условиях, занимает порой десятилетия работы научных коллективов, включающих десятки и сотни сотрудников, как например, это имеет место при разработке беспилотных автомобилей, предназначенных для движения по дорогам общего пользования.

К принципиальным моментам подобных практических разработок следует отнести то, что такие системы всегда базируются на формализмах норм, и именно нормы, а не ценности, закладываются в управляющую систему в качестве базы рациональности агента. Аксиологическая теория, однако, утверждает, что нормы, вообще говоря, базируются на системе ценностей, являясь лишь ее выражением. Хорошо известно аксиологическое положение

¹⁴³ Рассел С., Норвиг П. Искусственный интеллект. Современный подход. М.: Вильямс, 2007.

¹⁴⁴ В США предложили написать кодекс поведения боевых роботов [Электронный ресурс]. URL: <http://lenta.ru/news/2009/02/16/robots/> (дата обращения: 10.09.2016).

¹⁴⁵ Пентагон отказался от автономных боевых роботов [Электронный ресурс]. URL: www.vz.ru/news/2016/9/15/832749.html (дата обращения: 15.09.2016).

о вариативности норм при неизменности ценностей. Часто многообразие норм вырастает из небольшого количества базовых ценностей. К примеру, те же «законы робототехники» Азимова со всеми их вариациями очевидным образом порождаются простой иерархией ценностей: жизнь и здоровье человека, императивы человека, целостность робота.

Перечисленные свойства ценностей – их фундаментальность, сравнительно небольшое количество и относительная простота их формулировки – вызывают желание при конструировании агентных систем в качестве основы, формирующей поведение агентов, использовать именно ценности, а не нормы. Однако, как было сказано, при таком подходе возникают существенные трудности, связанные с тем, что формализация фундаментальных аксиологических категорий каждый раз сталкивается со слабой методологической разработанностью, присущей данной проблеме. В настоящей работе предлагается подход, позволяющий проводить формализацию ценностной системы агента, избегая при этом как экспликации аксиологических категорий, так и трудоемкого процесса предварительного формулирования и верификации системы норм и правил, определяющих поведение агента.

Пусть агент формально задан как тройка $\langle S, A, U \rangle$, где S – некоторое разбиение множества всевозможных наборов сенсорных данных, которое уместно назвать множеством *элементарных ситуаций*, A – множество доступных агенту действий, включающее нулевое действие, и U – алгоритм управления, отвечающий за выбор действия из множества A в конкретный момент времени. Тогда поведение агента определяется как последовательность $s_1 a_1 s_2 a_2 s_3 a_3 \dots$, где $s_i \in S$, а $a_i \in A$.

Развитие класса методов машинного обучения, известного как *обучение с подкреплением*¹⁴⁶, показывает, что при наличии механизма *оценки* осуществляемых агентом действий (вознаграждения действий) задача нахождения оптимального управления U , максимизирующего суммарное вознаграждение в некоторой более или менее длительной перспективе, имеет удовлетворительные с практической точки зрения решения. При этом необходимый оценочный механизм закладывается в конструкцию агента разработчиком, и агент не имеет возможности как-то на него влиять. К примеру, в задаче обучения агента локомоции в качестве вознаграждения (подкрепляющего воздействия) на каждом шаге может быть взято продвижение агента в требуемом направлении; можно сказать, что такое продвижение становится для агента безусловной *ценностью*.

¹⁴⁶ Саттон Р., Барто Э. Обучение с подкреплением. М.: Бином. Лаборатория знаний, 2012.

Рассмотрим модификацию метода обучения с подкреплением, которую можно назвать *обучение с оценкой ситуации*. Пусть агент располагает системой ценностей, позволяющей оценивать его поведение, причем оцениваются не действия агента, а ситуации, и пусть формальным выражением системы ценностей агента будет оценочная функция $\xi: S \rightarrow R$, где R – некоторое подмножество R . Уместно назвать R *оценочным множеством*. В простейшем случае $R = \{0, 1\}$, где 0 имеет значение «плохо», а 1 – «хорошо». Оценочная функция позволяет оценить любую ситуацию из множества возможных. Преимущество предлагаемой формализации состоит, прежде всего, в том, что оценочная функция, являясь лишь формальным выражением системы ценностей агента, вообще говоря, не требует формулировок этих ценностей, их семантического анализа и экспликации. Заметим, что подход, при котором оцениваются ситуации, а не действия, располагает очевидной производной оценкой ξ' и для действий агента:

$$\xi'(a_i) = \xi(s_{i+1}) - \xi(s_i)$$

то есть агент может естественным образом оценить также и свои действия с точки зрения улучшения или ухудшения текущей ситуации. Представляется особенно важным, что при таком подходе отсутствует семантика вознаграждения, подкрепления, дрессировки и тому подобного, и из категории *обучения с учителем* метод переходит в разряд методов *самообучения*, не теряя при этом возможность пользоваться результатами, получаемыми при разработке методов класса обучение с подкреплением. Можно сказать, что при таком подходе мы приобретаем многое, не теряя ничего. В итоге же система норм и правил поведения оказывается сформированной агентом самостоятельно, без участия разработчика, причем оптимальность сформированной системы гарантирована методом.

В качестве примера можно рассмотреть упомянутую выше задачу обучения беспилотных транспортных средств движению по дорогам общего пользования. При описываемом подходе нам уже не потребуется переводить ПДД на язык, понятный агенту, не потребуется проводить формализацию неписаных правил поведения на дорогах и всего прочего, что имеет в виду водитель-человек. Нам потребуется лишь представить систему ценностей в виде оценочной функции, положив, к примеру, что плохими являются ситуации аварии, то есть причинения ущерба участникам дорожного движения (включая, разумеется, и пешеходов),

а также ситуации получения взысканий от регулирующей процесс дорожного движения структуры. После того, как оценочная функция определена, можно запускать процесс обучения, руководствуясь тем или иным методом класса обучения с подкреплением. Разумеется, поскольку в основе таких методов всегда лежит вариант метода *проб и ошибок*, то большая часть обучения для этой задачи должна проводиться в эмуляторе, однако, можно полагать, что затраты на создание эмулятора окажутся существенно ниже затрат, возникающих при традиционном подходе. Заметим, что метод позволяет легко корректировать базовые ценности и экспериментировать, добиваясь улучшения результата. Немаловажно и то, что для каждой выбранной системы ценностей метод гарантирует оптимальность поведения, тогда как при традиционном подходе результат не может быть гарантирован в принципе.

Значительный интерес представляет вопрос, способны ли агенты формировать систему ценностей самостоятельно, без вмешательства разработчика. Формальным выражением системы ценностей агентов в рамках описываемого подхода является оценочная функция $\xi = \xi(s)$. Пусть $|S| = n$, $|A| = m$, $|R| = k$. Тогда общее число всевозможных оценочных функций $N_\xi = k^n$, в случае двухэлементного оценочного множества $R = \{0, 1\}$, имеем $N_\xi = 2^n$, что значит, что каждые десять дополнительных ситуаций увеличивают количество оценочных функций на три порядка. При $n < 30$ для нахождения наилучшей оценочной функции можно использовать, например, методы класса *генетических алгоритмов*. Если же количество возможных ситуаций велико, то для его уменьшения можно применить, например, методы, связанные с использованием *искусственных нейронных сетей*. Заметим, что рост мощности оценочного множества может быть постепенным и для его обеспечения уместно воспользоваться *эволюционными стратегиями*.

В заключение отметим, что стартовое число $N_\xi = 2^n$ существенно меньше, чем общее минимальное число $N = m^n$ поведенческих стратегий, которые пришлось бы перебирать при отсутствии оценочной функции. Очевидно, что такой перебор оказывается практически неосуществим даже при сравнительно небольших значениях мощностей множеств S и A . Так, при $n = m = 10$ имеется $N = 10^{10} = 10\,000\,000\,000$ возможных поведения агента. Предлагаемый метод, использующий оценочную функцию $\xi = \xi(s)$, как считает автор, позволит подойти к решению задач, традиционно относимых в области искусственного интеллекта к категории сложных.