Семантика первопорядковой динамической логики знания[1]

Е. Е. Ледников

ABSTRACT. In the paper the semantics for DK_{pr}-logic (first order dynamic logic of knowledge) are proposed. In such semantics some intuitive properties of possible worlds semantics are used.

Ключевые слова: семантика, модальность, интерпретация, модель, знание, убежденность, доказательство, вера, мнение, сомнение, опровержение.

В работах [1, 2] была сформулирована в аксиоматической форме и в виде аналитических таблиц первопорядковая динамическая логика знания (далее будем ее обозначать как DK_{pr}-логику). Сформулирована она в первопорядковом языке PL_d, содержащем счетное множество индивидных и предикатных переменных и исходные логические символы $\{\sim, \&, \vee, \supset, \equiv, \forall, \exists, =, K_\varphi, C_\varphi, G_\varphi, T_\varphi, B_\varphi, D_\varphi, R_\varphi\}$. Напомним, что модальные операторы языка PL_d, характеризующие ментальные состояния субъекта φ в процессе познавательной деятельности, означают, соответственно, «субъект φ знает, что...», «субъект φ убежден в том, что...», «субъект φ доказывает, что...», «субъект φ верит, что...», «субъект φ полагает, что...», «субъект φ сомневается в том, что...», «субъект φ опровергает, что...». Все формулы вида $K_\varphi A, C_\varphi A, G_\varphi A, T_\varphi A, B_\varphi A, D_\varphi A, R_\varphi A$ или их отрицания, где A — формула классической первопорядковой логики, являются формулами DK_{pr}-логики. Также формулами DK_{pr}-логики являются все формулы вида $(\forall x)B$ и $(\exists x)B$ и их отрицания, где B — формула вида $\nabla_\varphi A$, причем A — формула классической первопорядковой логики, ∇_φ — один из модальных операторов DK_{pr}-логики. Также формулами DK_{pr}-логики являются

[1]Исследование поддержано РГНФ, проект № 07-03-00335a.

все формулы вида $K_\varphi \nabla_\varphi A$, поскольку субъекту φ должно быть позволено осознавать все свои ментальные состояния.

Естественно возникает вопрос, какой может быть семантика подобной модальной логики. Воспользуемся хорошо известной идеей семантики возможных миров. Нам потребуются для каждого модального оператора наборы возможных миров с отношением достижимости на них. Что должны представлять собой подобные миры? Возьмем, например, оператор личностного знания K_φ. В этом случае возможные миры — это, говоря словами Я. Хинтикки, «эпистемические альтернативы» знаниям субъекта φ в «действительном» мире [3]. Но знания выражаются в высказываниях. Значит, все миры должны содержать наборы высказываний — от простых, т. е. пропозициональных переменных, до сложных, образованных с помощью логических связок и кванторов. Думается, что такую степень логической образованности для субъекта φ мы вправе допустить. В противном случае, какой вообще смысл гворить о логике рассуждений, основанной на его знаниях? Далее, запас знаний субъекта всегда конечен, хотя, с известной долей идеализации, можно допустить, что он может быть сколь угодно большим. Эти запасы знаний базируются на описаниях состояния [4], представляющих собой множества атомарных высказываний или их отрицаний. Но, в отличие от Р. Карнапа, исследовавшего алетические модальности и поэтому говорившего только о полных описаниях состояния, мы будем допускать в качестве эпистемически возможных миров и неполные описания состояния, т. е. такие, в которых отсутствуют как некоторые атомарные высказывания, так и их отрицания, а, значит, отсутствуют и сложные высказывания, которые можно было бы построить из отсутствующих атомарных. Данное допущение выглядит вполне уместным, поскольку реальный носитель знаний никогда не обладает полными, исчерпывающими знаниями. По-видимому, подобные эпистемические альтернативы должны удовлетворять условию непротиворечивости, т. е. ни одно атомарное высказывание не должно входить ни в один эпистемически возможный мир наряду со своим отрицанием.

Все сказанное относится и к другим наборам альтернатив, за исключением условия непротиворечивости. Например, в описания состояния, лежащие в основе доксатических альтернатив, вполне могут входить некоторые атомарные высказывания одновременно с их отрицанием. В самом деле, вполне мыслима, например, ситуация, когда субъект мнения полагает, что предсказывать судьбу невозможно, и при этом в сложных жизненных обстоятельствах он без колебаний обращается к гадалкам. Но подобная непоследовательность может быть характерной и для других умственных состояний субъекта φ. Скажем, субъекту позволительно верить во взаимоисключающие, несовместимые вещи — разве не с этим феноменом мы сталкиваемся в случае религиозной веры некоторых ученых?

Перейдем теперь к построению семантики DK_{pr}-логики. Моделью M будет $(U, H, W^k, W^c, W^g, W^t, W^b, W^d, W^r, R^k, R^\nabla, Val)$, где U — предметная область индивидов $a, b, c, d \ldots$ — универсум рассуждения, H — выделенный («реальный») мир, W^k, W^c, W^g, W^t, W^b, W^d, W^r — множества миров (альтернатив) соответствующей индексу динамической модальности, R^k — отношение достижимости на W^k (на эпистемических альтернативах), R^∇ — отношение достижимости на альтернативах остальных видов модальностей, Val — функция приписывания значений выражениям языка PL_d — функция означивания. Отношение R^k обладает свойствами рефлексивности и транзитивности, на отношение R^∇ не наложено никаких ограничений.

Для каждой пары множеств миров W^k, W^∇ существует множество S такое, что $\{S\} = \{p, q, r, \ldots A_1, \ldots A_m\} = \{W_1^k\} \cap \{W_2^k\} \cap \ldots \{W_j^k\}$ (где j — число элементов в W^k), причем $\{S\} \subseteq \{W_i^\nabla\}$ для произвольного возможного мира W_i^∇. Другими словами, множество S — это общая часть эпистемических альтернатив, которая является подмножеством множества высказываний произвольной альтернативы W_i^∇.

Охарактеризуем теперь с помощью определений функцию означивания Val:

ОПРЕДЕЛЕНИЕ 1. $Val(x_i) = a \in U$, т. е. любой предметной переменной x_i эта функция сопоставляет некоторый индивид a из предметной области U.

ОПРЕДЕЛЕНИЕ 2. $Val(P_i^n) = V(P_i^n) = \{\langle a_1, a_2, \ldots a_n\rangle, \langle b_1, b_2, \ldots b_n\rangle, \ldots\}$, т. е. любой n-местной предикатной переменной P_i^n

функция Val сопоставляет объем V, состоящий из множеств упорядоченных n-ок индивидов из U.

ОПРЕДЕЛЕНИЕ 3. $Val(A) =$ *И* или $Val(A) =$ *Л*. Другими словами, любой формуле A языка PL_d функция Val сопоставляет семантический объект «истина» или семантический объект «ложь». Эту же мысль можно выразить как $M, W_i \models A$ или $M, W_i \not\models A$ — как выполнимость (соответственно, невыполнимость) формулы A в возможном мире W_i модели M (разумеется, при функции означивания Val, входящей в определение модели M). Отношение выполнимости $\models$ определяется следующим образом:

1) для элементарного высказывания p или его отрицания $\sim p$: $M,\ W_i \models p\ (\sim p)$, если $p(\sim p) \in W_i$;

2) для отрицания произвольного неэлементарного высказывания A, содержащего только те элементарные высказывания, которые принадлежат к $W_i : M, W_i \models\sim A$, если, $M, W_i \not\models A$; $M, W_i \not\models\sim A$, если $M, W_i \models A$;

3) для конъюнкции $A_1\&A_2 : M, W_i \models A_1\&A_2$, если $M, W_i \models A_1$ и $M, W_i \models A_2$; $M, W_i \not\models A_1\&A_2$, если $M, W_i \not\models A_1$ или $M, W_i \not\models A_2$;

4) для дизъюнкции $A_1 \vee A_2$: $M, W_i \models A_1 \vee A_2$, если $M, W_i \models A_1$ или $M, W_i \models A_2$; $M, W_i \not\models A_1 \vee A2$, если $M, W_i \not\models A_1$ и $M, W_i \not\models A_2$;

5) для импликации $A_1 \supset A_2$: $M, W_i \models A_1 \supset A_2$, если $M, W_i \not\models A_1$ или $M, W_i \models A_2$; $M, W_i \not\models A_1 \supset A_2$, если $M, W_i \models A_1$ и $M, W_i \not\models A_2$;

6) для высказывания о знании $K_\varphi A$: $M, W_i^k \models K_\varphi A$, если $M, W_j^k \models A$ в каждом W_j^k таком, что $W_i^k R^k W_j^k$, иначе $M, W_i^k \not\models K_\varphi A$;

7) для высказывания о любом ином ментальном состоянии субъекта φ: $M, W_i^\nabla \models \nabla_\varphi A$, если $M, W_j^\nabla \models A$ в каждом W_j^∇ таком, что $W_i^\nabla R^\nabla W_j^\nabla$, иначе $M, W_i^\nabla, \not\models \nabla_\varphi A$;

8) для элементарной формулы исчисления предикатов $P^n(x_{1...}x_n)$: $M, W_i \models P^n(x_{1...}x_n)$, если $\langle Val(x_1), \ldots Val(x_n)\rangle \in V(P^n)$, в противном случае $M, W_i \not\models P^n(x_{1...}x_n)$;

9) для формулы с квантором общности $(\forall x)A(x)$: $M, W_i \models (\forall x)A(x)$, если для каждой функции Val', отличающейся от Val самое большее приписыванием значения для x в $A(x)$, $M, W_i, Val' \models A(x)$. В противном случае, $M, W_i \not\models (\forall x)A(x)$;

10) для формулы с квантором существования $(\exists x)A(x)$: $M, W_i \models (\forall x)A(x)$, если существует хотя бы одна функция Val', отличающаяся от Val самое большее приписыванием значения для x в $A(x)$, такая, что $M, W_i, Val' \models A(x)$. В противном случае, $M, W_i \not\models (\exists x)A(x)$.

Высказывание A общезначимо ($\models A$), если $M, H \models A$ во всех своих моделях M.

Теперь следует остановиться на особенностях квантификации в модальных контекстах, или, как раньше было принято говорить, на *de re* модальностях. Как избежать известных модальных парадоксов, связанных с подставимостью тождественного, $\forall$-введением и $\exists$-удалением? Ранее нами указывалось [5], что значениями квантифицируемых переменных в эпистемических контекстах должны быть «известные» индивиды, а в доксатических контекстах (контекстах мнения) — так сказать, «полагаемые» индивиды. Обобщая данную идею, нам следует сделать вывод, что при построении DK_{pr}-логики потребуются также индивиды, в существовани которых субъект φ убежден, существование которых он доказал или подтвердил эмпирическими исследованиями, в существование которых он верит, в существовании которых он сомневается и существование которых он опровергает. Весь спектр подобных индивидов можно задать в языке PL_d посредством соответствующих индивидных дескрипций. Их можно считать, в духе идей Б. Рассела, «неполными» символами. Но вводиться в язык они должны контекстуальными определениями, отличающимися от расселовских. С учетом специфики возникающих в DK_{pr}-логике модальных контекстов, каждый объект квантификации должен удовлетворять условию, так сказать, «модального» существования и единственности — он должен существовать в единственном экземпляре не только в выделенном «реальном мире», но и в соответствующих наборах возможных миров.

А именно каждый известный субъекту φ индивид должен характеризоваться индивидной дескрипцией вида:

ОПРЕДЕЛЕНИЕ 4. $E_k!(\iota x)A =_{df} (\exists y)[K_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \mathbf{.}\, z = y))]$, контекстуально элиминируемой с помощью определения:

ОПРЕДЕЛЕНИЕ 5. $[(\iota x)A]B(\iota x)A =_{df} (\exists y)[K_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \mathbf{.}\, z = y)) \,\&\, B(y)]$. Индивид, в существовании которого субъект φ убежден, будет характеризоваться индивидной дескрипцией вида:

ОПРЕДЕЛЕНИЕ 6. $E_c!(\iota x)A =_{df} (\exists y)[C_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \mathbf{.}\, z = y))]$, контекстуально элиминируемой с помощью определения:

ОПРЕДЕЛЕНИЕ 7. $[(\iota x)A]B(\iota x)A =_{df} (\exists y)[C_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \mathbf{.}\, z = y)) \,\&\, B(y)]$.

Данная схема контекстуальных определений и элиминации справедлива для контекстов любых динамических модальностей:

ОПРЕДЕЛЕНИЕ 8. $E_\nabla!(\iota x)A =_{df} (\exists y)[\nabla_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \mathbf{.}\, z = y))]$,

ОПРЕДЕЛЕНИЕ 9. $[(\iota x)A]B(\iota x)A =_{df} (\exists y)[\nabla_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \mathbf{.}\, z = y)) \,\&\, B(y))]$.

Во всех приведенных определениях выражение $[(\iota x)A]B(\iota x)A$ обозначает модализированную формулу B, содержащую на одном из аргументных мест индивидную дескрипцию $[(\iota x)A]$, причем область действия этой дескрипции максимальна. Выражения $E_k!(\iota x)A, E_c!(\iota x)A$ и $E_\nabla!(\iota x)A$ — соответственно утверждения об известном субъекту φ существовании и единственности объекта $(\iota x)A$, о существовании и единственности объекта $(\iota x)A$, в котором субъект φ убежден, о существовании и единственности объекта $(\iota x)A$, на который направлено одно из остальных ментальных состояний субъекта φ.

Исходя из сказанного, правила для кванторов в модальных контекстах DK_{pr}-логики уместно, в отличие от классической логики предикатов, переформулировать следующим образом:

($\forall$-удаление). Если $\vdash (\exists y)[\nabla_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \centerdot\, z = y))]$, то $\vdash (\forall w)B_1(w) \supset B_2(\iota x)A$. Здесь ∇_φ — тот из личностных модальных операторов DK_{pr}-логики, в области действия которого находится индивидная дескрипция $(\iota x)A$ в формуле B_2. При этом модализированная формула B_2 отличается от модализированной формулы B_1 только вхождением индивидной дескрипции $(\iota x)A$ на место свободного вхождения индивидной переменной w.

($\exists$-введение). $\vdash B_2(\iota x)A \supset (\exists w)B_1(w)$. Опять модализированная формула B_2 отличается от модализированной формулы B_1 только вхождением индивидной дескрипции $(\iota x)A$ на место свободного вхождения индивидной переменной w. Следует обратить внимание на то обстоятельство, что контекстуальные определения формулы $B_2(\iota x)A$ (определения 5, 7, 9) предполагают ее истинность только при соответствующем модальном существовании и единственности дескрипции $(\iota x)A$. Поэтому дополнительная посылка такого рода была бы излишней.

Еще одна проблема квантификации в DK_{pr}-логике связана с нарушением в модальных контекстах правила подставимости тождественного (хорошо известный парадокс «Утренней Звезды» и «Вечерней Звезды» связан как раз с этим нарушением). Чтобы подобного парадокса не возникало, достаточно принять следующие два правила:

a) (=-подставимость индивидной дескрипции на место свободного вхождения индивидной переменной). Если $\vdash (\exists y)[\nabla_\varphi(A(y) \,\&\, (\forall z)(A(z) \equiv \centerdot\, z = y))]$, то $\vdash (\iota x)A = w \supset \centerdot\, B_1(w) \equiv B_2(\iota x)A$. Модализированная формула B_2 отличается от модализированной формулы B_1 только вхождением индивидной дескрипции $(\iota x)A$ на место свободного вхождения индивидной переменной w;

b) (=-подставимость одной индивидной дескрипции на место другой). Если $\vdash (\exists z_1)[\nabla_\varphi(A_1(z_1) \,\&\, (\forall y_1)(A_1(y_1) \equiv \centerdot\, y_1 = z_1))]$, $\vdash (\exists z_2)[\nabla_\varphi(A_2(z_2) \,\&\, (\forall y_2)(A_2(y_2) \equiv \centerdot\, y_2 = z_2))]$, то $\vdash (\iota x_1)A_1 = (\iota x_2)A_2 \supset \centerdot\, B_1(\iota x_1)A_1 \equiv B_2(\iota x_2)A_2$. Модализированная формула B_2 отличается от модализированной формулы B_1 только вхождением индивидной дескрипции $(\iota x_2)A_2$ на место вхождения индивидной дескрипции $(\iota x_1)A_1$.

Может показаться, что построение DK_{pr}-логики ведет к существенному усложнению синтаксических правил языка. Однако приведенные правила детерминируются предложеной семантикой, не оставляющей нам свободы в их формулировках.

Литература

[1] *Ледников Е.Е.* Об одном варианте динамической логики знания // Логические исследования. Вып. 14. М., 2007. С. 218–223.

[2] *Ледников Е.Е.* Набросок первопорядковой кванторной динамической логики знания // Современная логика: проблемы теории, истории и применения в науке. Материалы X Общероссийской научной конференции 26–28 июня 2008 г. Санкт-Петербург, 2008. С. 287– 289.

[3] *Хинтикка Я.* Семантика пропозициональных установок // Логико-эпистемологические исследования. М., 1980. С. 68–101.

[4] *Карнап Р.* Значение и необходимость. М., 2000. С. 38–39.

[5] *Ледников Е.Е.* Некоторые особенности первопорядковой кванторной логики знания и мнения // Логические исследования. Вып. 12. М., 2005. С. 207–210.