И.Б.Микиртумов

СЕМАНТИКА АНАФОРЫ КАК ОПЕРАТОРА ДООПРЕДЕЛЕНИЯ

Abstract. *This paper deals with logic of anaphoric update operator. It is well known that the diverse variants of dynamic semantics are efficient way to interpretation of anaphora. At the same time an anaphora itself may be considered as update operator with property of eliminating. This operator acts within the limits of the epistemic attitudes of the recipient of phrase and determines the meaning of anaphoricaly defined names. It is possible to express the static meaning of anaphor's operator by means of epistemic logic by use the methods of van Eijck and de Vries's update logic.*

Первая версия динамической семантики для языка логики предикатов была построена Гренендийком и Штокхофом [1] с целью интерпретации межсентенциальной или внешней анафоры, присутствующей, например, в последовательности высказываний

У фермера есть осел. Он его бьет (1)

и анафоры внутренней, например, во фразе

Если у фермера есть осел, то он его бьет (2)

В (1) и (2) мы подразумеваем кореференциальность или анафорическую связь определенных имен «он», «его» с неопределенными именами «фермер» и «осел».

Подразумеваемые условия истинности для (1) требуют от любого означивания *i*, верифицирующего конъюнкцию образующих его предложений, верифицировать и равенства «фермер»=«он» и «осел»=«его». При разных семантических подходах это условие выполняется почти одинаково. Например, в семантике ящиков Хейм [2] анафора предполагается как уже заданная интерпретатором на подготовительной стадии анализа. В логике динамического означивания (ЛДО) ван Эйка и де Вриза [3], которая модифицирует подход Гренендийка и Штокхофа, значения переменных зафиксированы операторами динамического означивания, а анафорически определенным именам сопоставлены те же переменные, что и антецедентам анафоры. Такие решения трудно назвать естественными, если речь идет о логической природе самой анафоры, а не о том, как интерпретировать уже отождествленные имена. Мы попытаемся раскрыть логические свойства анафоры, оставаясь в границах динамического анализа значения.

Язык динамической логики с эпистемическим модальностями $L_{ДЭ}$ содержит два уровня. Первый уровень образован языком ЛДО ван Эйка и де Вриза, к которому добавлены эпистемические динамические модальности и оператор анафорического доопределения – $L_{ЛДОА}$. Этот язык, в свою очередь, в качестве внутреннего содержит язык логики предикатов первого порядка с равенством, но без индивидных и функциональных констант – $L_{ЛП}$.

Метапеременные по формулам $L_{ЛП}$: A, B, C, ... ; по формулам $L_{ЛДОА}$: π_1, π_2, π_3, ...; по формулам $L_{ДЭ}$: φ, ψ, ζ.

Правильно построенные выражения $L_{ЛДОА}$ и $L_{ДЭ}$ называются программами.

Определение программы $L_{ЛДОА}$.

1. Формулы $L_{ЛП}$ являются программами $L_{ЛДОА}$.

2. Если π, π_1, π_2 – программы, $B \in L_{ЛП}$, то $(\pi_1 ; \pi_2)$, $(\pi_1 \Rightarrow \pi_2)$, $\neg\pi$, $\exists x\pi$, $(\eta x{:}\ \pi)$, $\mathrm{P}_a\pi$, $\pi_1{\uparrow}(x{\leftarrow}y)$ – программы (которые обозначают соответственно последовательное выполнение π_1 и π_2, динамическую импликацию, динамическое отрицание (совпадает с обычным, если $\pi \in L_{ЛП}$), квантор существования, динамическое означивание переменной x, эпистемический оператор динамической возможности, анафорическое доопределение программы π_1 по переменной x).

Определения: $\forall x\pi =_{\mathrm{def}} \neg\exists x\neg\pi$; $\mathrm{N}_a\pi =_{\mathrm{def}} \neg \mathrm{P}_a\neg\pi$; $\top$ есть тавтология, а $\bot =_{\mathrm{def}} \neg\top$.

Определение программы $L_{ДЭ}$.

1. Если $\pi \in L_{ЛДОА}$, то $\pi \in L_{ДЭ}$.

2. Если $\pi \in L_{ЛДОА}$, $A \in L_{ЛП}$, то $\langle\pi\rangle A \in L_{ДЭ}$.

3. Если $\varphi \in L_{ДЭ}$, то, $\otimes_a\varphi$, $\Diamond_{\mathrm{a}}\varphi \in L_{ДЭ}$.

4. Если φ, $\psi \in L_{ДЭ}$, то $\neg\varphi$, $\varphi\wedge\psi$, $\varphi\rightarrow\psi \in L_{ДЭ}$.

Определения: $[\pi]A =_{\mathrm{def}} \neg\langle\neg\pi\rangle A$ и $\Box_{\mathrm{a}}\varphi =_{\mathrm{def}} \neg\Diamond_{\mathrm{a}}\neg\varphi$.

Семантика $L_{ДЭ}$ задается на модели $\mathbf{M}=\langle\mathbf{W}, \mathbf{D}, \mathbf{I}, \mathbf{S}, \mathbf{R}_a, \mathbf{R}_a{}^{\delta}, \mathbf{Val}, \ast\rangle$, где $\mathbf{W}$ – множество миров, $\mathbf{D}$ – область индивидов, $\mathbf{I}$ – функция интерпретации констант, $\mathbf{S}$ – множество носителей эпистемических установок, $\mathbf{R}_a$ и $\mathbf{R}_a{}^{\delta}$ – отношения эпистемической и динамической достижимости, которыми располагает носитель пропозициональной установки a, $\mathbf{Val}=\{i : \mathrm{Var}\rightarrow\mathbf{D}\}$ – множество означиваний переменных, $\ast$ – функция, порождающая «мир актуального знания a» для каждого мира и означивания. Значение программ $L_{ЛДОА}$ определяется относительно означиваний на элементах $\mathbf{W}$ функцией $|\pi|_{\mathrm{M}}$: $\mathbf{Val}\times\mathbf{W}\rightarrow\mathcal{P}(\mathbf{Val}\times\mathbf{W})$. (Индексы модели опускаем.)

1. $(i, w)|\bot| = \varnothing$
2. $(i, w)|P^n(t_1, ..., t_n)| = \{i, w\}$, если $<i(t_1), ..., i(t_n)> \in \mathbf{I}^w(P^n)$

= $\varnothing$ в противоположном случае

3. $(i, w)|t_m=t_k|$ = $\{i, w\}$, если $<i(t_m), i(t_k)> \in \mathbf{I}^w(=)$

= $\varnothing$ в противоположном случае

4. $(i, w)|A \wedge B|$ = $\{i, w\}$, если $(i, w)|A| \cap (i, w)|B| \neq \varnothing$

= $\varnothing$ в противоположном случае

5. $(i, w)|\pi_1 ; \pi_2|$ = $\cup\{(j, w)|\pi_2| : (j, w) \in (i, w)|\pi_1|\}$

6. $(i, w)|\pi_1 \Rightarrow \pi_2|$ = $\{i, w\}$, если для любого $(j, w) \in (i, w)|\pi_1|$ верно, что $(j, w)|\pi_2| \neq \varnothing$

= $\varnothing$ в противоположном случае

7. $(i, w)|\neg\pi|$ = $\{i, w\}$, если $(i, w)|\pi_1| = \varnothing$

= $\varnothing$ в противоположном случае

8. $(i, w)|\eta x{:}\ \pi|$ = $\cup\{(i(x/\mathbf{d}), w)|\pi|\}$, где $\mathbf{d} \in \mathbf{D}$

9. $(i, w)|\exists x\pi|$ = $\{i, w\}$, если $\{\mathbf{d} \in \mathbf{D}{:}\ (i(x/\mathbf{d}), w)|\pi| \neq \varnothing\} \neq \varnothing$

= $\varnothing$ в противоположном случае;

10. $(i, w)|\mathrm{P}_a\pi|$ = $\{i, w\}$, если существует j, такое, что $i\mathbf{R}_a{}^{\delta}j$

=$\varnothing$ в противоположном случае

11. $(i, w)|\pi\uparrow(x \leftarrow y)|$ = $(w, i)|\pi'|$, где π' получена из π вставкой перед подпрограммой π_n, в которой впервые встречается y, записи $\lfloor(\eta y{:}\ x=y) ; \rfloor$.

Отношение динамической достижимости $\mathbf{R}_a{}^{\delta}$, релятивизованное к носителю эпистемической установки a, связывает означивания i и j в том случае, если вычисление значения программы π в i в контексте знания a о положении дел в мире w приводит к означиванию j.

Определение истинности (выполнимости) для $L_{ДЭ}$:

i, wё φ $\Leftrightarrow (i, w)|\varphi| \neq \varnothing$

i, wё $\langle\pi\rangle A$ $\Leftrightarrow (i, w)|\pi ; A| \neq \varnothing$

i, wё $\otimes_a\varphi$ $\Leftrightarrow i, w_a$*¦ -φ.

i, wё $\Diamond_a\varphi$ $\Leftrightarrow$ существует v, такой, что $w\mathbf{R}_a v$ и i, vё -φ.

Отношение $\mathbf{R}_a$ рефлексивно. Прочие его свойства не уточняются.

Понятие истинности обобщается для модели и класса моделей. Будем обозначать как $\mathrm{V}(\varphi)$ множество пар $<i, w>$ таких, что i, wё φ.

$\otimes_a$ – это эпистемический оператор, описывающий «мир актуального знания субъекта a» на паре $<w, i>$. Этот мир соответствует содержанию динамических эпистемических модальностей. Действует следующее определение:

w, iё $\mathrm{P}_a\pi$ $\Leftrightarrow_{\mathrm{def}}$ i, w_a*ё $\langle\pi\rangle$T.

Перевод выражений естественного языка на язык $L_{ЛДОА}$ осуществляем в соответствии с правилом: каждому неопределенному имени сопоставляется новая переменная, для которой вводится оператор динамического означивания с характеристикой этой переменной. Для всякого анафорически определенного имени также вводим новую переменную.

Для стандартного примера анафоры

Если у фермера ест ь осел, т о он его бьет

получаем следующий перевод (π):

$$((\eta x\colon \mathit{Ф}x) ; (\eta y\colon \mathit{О}y) ; \mathit{П}xy) \Rightarrow \mathit{Б}x_1x_2.$$

Теперь, чтобы анафорически связать переменные x_1 и x_2 с x и y, осуществляем анафорическое доопределение, т. е. приписываем к π слева $\uparrow(x_1{\leftarrow}x)\uparrow(x_2{\leftarrow}y)$ и получаем программу π':

$$((\eta x\colon \mathit{Ф}x) ; (\eta y\colon \mathit{О}y) ; \mathit{П}xy) \Rightarrow$$
$$((\eta x_1\colon x{=}x_1) ; (\eta x_2\colon y{=}x_2) ; \mathit{Б}x_1x_2)).$$

Значение π' в (i, w) определяется так (индекс w опускаем):

$$(i)(((\eta x\colon \mathit{Ф}x) ; (\eta y\colon \mathit{О}y) ; \mathit{П}xy) \Rightarrow$$
$$((\eta x_1\colon x{=}x_1) ; (\eta x_2\colon y{=}x_2) ; \mathit{Б}x_1x_2))=\{i\},$$

если для любого $j\in(i)\,|\,(\eta x\colon \mathit{Ф}x) ; (\eta y\colon \mathit{О}y) ; \mathit{П}xy|$ верно, что $(j)\,|\,(\eta x_1\colon x{=}x_1) ; (\eta x_2\colon y{=}x_2) ; \mathit{Б}x_1x_2\,|\neq\varnothing$. Последнее неравенство имеет следующий вид:

$$\cup\{j(x_1/\mathbf{d}')(x_2/\mathbf{d}''),$$

(*)

где $<j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')(x_1), j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')(x_2)>\in\mathbf{I}(\mathit{Б})\}$,

где $j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')\in\cup\{j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')$:

$<j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')(y),$

$j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')(x_2)>\in\mathbf{I}(=)\}$,

где $j(x_1/\mathbf{d}')\in\cup\{j(x_1/\mathbf{d}') : <j(x_1/\mathbf{d}')(x), j(x_1/\mathbf{d}')(x_1)\in\mathbf{I}(=)\}\neq\varnothing$.

Если бы наш пример вместо $\mathit{Б}x_1x_2$ содержал в себе $\mathit{Б}xy$, то мы получили бы неравенство:

$(j)\,|\,\mathit{Б}xy|\neq\varnothing$, откуда $(j)\,|\,\mathit{Б}xy|=\{j\}$ и $<j(x), j(y)>\in\mathbf{I}(\mathit{Б})$. (**)

Условия (*) и (**) эквивалентны. В самом деле, поскольку x_1 и x_2 по условию суть новые переменные, множество выполняющих условие означиваний вида $j(x_1/\mathbf{d}')(x_2/\mathbf{d}'')$ совпадает со множеством означиваний вида j. Таким образом, доопределение, которое мы осуществили, равнозначно установлению анафоры между определенными и неопределенными именами. Помещая π' в рамки эпистемической установки, получим $(\mathrm{P}_a\pi')\mathrm{T}$, которое истинно тогда и только тогда, когда существует j такое, что $i\mathbf{R}_a^{\pi}j$, т.е. когда π' имеет непротиворечивое завершение.

Оператор $\uparrow$ вводит анафорическую связь переменных как элемент программы, по которой вычисляется значение выражения. В отличие от оператора динамического означивания ЛДО, оператор $\uparrow$ элиминативен. Это обнаруживает существенную черту анафоры: она действует как тест.

Каждое конкретное анафорическое доопределение программы π связано с тем, что реципиент ожидает непустого выполнения программы π'. Введение автором анафорически определенных имен равнозначно замене высказывания

У фермера есть осел. Фермер бьет осла.

на

У фермера есть осел. Он его бьет.

Эти имена выполняют функцию логических операторов (индексов или маркеров), поскольку инициируют процесс поиска источника референции в контексте. Автор высказывания, вводя определенное имя, подразумевает для него анафорическую связь, причем так, что предполагаемый автором процесс поиска этой связи реципиентом однозначно завершается. Этот процесс поиска есть перебор и сортировка возможных доопределений программы, от которых реципиент, исходя из своей эпистемической установки, ожидает положительного результата. В семантическом анализе анафоры, таким образом, присутствуют план автора и план реципиента. Динамические свойства интерпретации проявляются в обоих случаях, но по-разному. Анализ статических значений программ в ЛДО ван Эйка и де Вриза, принадлежащий к плану автора, приводит к формулировке выражаемой им в высказывании пропозиции в виде замкнутой формулы языка логики предикатов. Исходя из такой формулировки, автор задает анафорические отношения имен, где консеквентам анафор могут сопоставляться свободные переменные, в действительности не являющиеся таковыми, в силу динамических свойств связок и присутствия операторов динамического означивания. Определение «извне» статического значения является в этом случае достаточно тривиальной задачей, которую можно решать и не прибегая к механизму редукции ЛДО, а рассматривая ЛДО как динамическую логику и пользуясь только семантическими определениями.

Для реципиента проблема интерпретации состоит в том, чтобы, соотнося логическую форму высказывания в ЛДО с пропозициями, входящими в его эпистемическую установку, восстановить подразумеваемую анафору или хотя бы свести к минимуму возможные варианты такого восстановления. Нас интересуют здесь, конечно же, чисто логические критерии приемлемости анафоры, а не грамматические, прагматические или какие-либо иные.

Попробуем теперь, пользуясь методами ЛДО и логики модификации ван Эйка и де Вриза [4], определить понятие слабейшего эпистемического условия для программы π и формулы B, которое будет рассматриваться как соответствующее статическому значению программы π в эпистемической логике с оператором $\otimes_a$:

слабейшее эпистемическое условие программы π и формулы B есть формула $\varphi \in L_{ДЭ}$ такая, что $V(\varphi)=V(\langle\pi\rangle B)$.

Определим функцию **су**, задающую φ по π и B (коллизии переменных, свободно входящих в B, избегаются):

1. **су**$(\bot, B) = \bot$
2. **су**$(P^n(t_1...t_n), B) = P^n(t_1...t_n)\wedge B$
3. **су**$((t_m=t_n), B) = (t_m=t_n)\wedge B$
4. **су**$((C\wedge D), B) = (C\wedge D)\wedge B$
5. **су**$((\pi_1 ; \pi_2), B) =$ **су**$(\pi_1,$ **су**$(\pi_2, B))$
6. **су**$((\pi_1 \Rightarrow \pi_2), B) = B\wedge($**су**$(\neg\pi_1, \top)\vee$**су**$(\pi_2, \top))$
7. **су**$(\neg\pi, B) = B\wedge\neg$**су**$(\pi,\top)$
8. **су**$(\eta x : \pi, B) = \exists x($**су**$(\pi, B))$
9. **су**$(\exists x\pi, B) = \exists x($**су**$(\pi, \top))\wedge B$
10. **су**$(P_a\pi, B) = \otimes_a$**су**$(\pi, \top)\wedge B$
11. **су**$(\pi\uparrow(y\leftarrow x), B) =$ **су**(π', B), где π' получено из π соответствующей вставкой $(\eta y\!: x=y)$.

Функция **су** дает неэпистемическое статическое значение программы π как формулу логики предикатов, если π не содержит в себе оператора P_a, а в качестве B взята формула $\top$.

Рассмотрим неэпистемическое **су** для доопределения логической формы π примера (2)

$(((\eta x\!: Фx) ; (\eta y\!: Oy) ; Пxy) \Rightarrow Бx_1x_2)\uparrow(x_1\leftarrow x)\uparrow(x_2\leftarrow y)$,

где π' это

$((\eta x\!: Фx) ; (\eta y\!: Oy) ; Пxy) \Rightarrow$
$((\eta x_1\!: x=x_1) ; (\eta x_2\!: x_2=y) ; Бx_1x_2)$.

Тогда **су**$(\pi',\top) =$

$\exists x(Фx\wedge\exists y(Oy\wedge Пxy)\rightarrow\exists x_1(x=x_1\wedge\exists x_2(y=x_2\wedge Бx_1x_2))$,

что эквивалентно формуле $\exists x\exists y((Фx\wedge Oy\wedge Пxy)\rightarrow Бxy)$. При введении внешнего оператора P_a для π' мы получили бы в качестве **су**$(\pi',\top)$ формулу

$\otimes_a(\exists x(Фx\wedge\exists y(Oy\wedge Пxy)\rightarrow\exists x_1(x=x_1\wedge\exists x_2(y=x_2\wedge Бx_1x_2))))$,

являющуюся эпистемическим статическим значением π'.

Верна следующая теорема о **су**: $V($**су**$(\pi, B))=V(\langle\pi\rangle B)$. Она представляет собой модификацию для $L_{ДЭ}$ теоремы, доказанной ван Эйком и де Вризом для логики модификации. Рассмотрим только случай **су**$(P_a\pi, B)$.

Пусть $V($**су**$(P_a\pi, B) = V(\otimes_a$**су**$(\pi, \top)\wedge B)$,
$= V(\otimes_a$**су**$(\pi, \top))\cap V(B)$.

Докажем утверждение слева направо. Допустим противоположное, а именно, что существует пара $<i, w>$ такая, что

i, *w*ё $\otimes_a$**су**$(\pi, \mathrm{T})\wedge B$, но неверно *i*, *w*ё $-\langle P_a\pi\rangle B$, т. е.
i, *w*ё $\otimes_a$**су**(π, T), *i*, *w*¦ $-B$ и неверно, что *i*, *w*ё $-P_a\pi\wedge B$. Это значит, что неверно *i*, *w*ё $P_a\pi$. Поскольку по предположению индукции V(**су**(π, T))=V($\langle\pi\rangle\mathrm{T}$)), имеем *i*, *w*ё $\otimes_a\langle\pi\rangle\mathrm{T}$, откуда по определению $\otimes_a$ получаем *i*, w_a*ё $\langle\pi\rangle\mathrm{T}$, откуда по определению w_a* следует *i*, *w*ё $P_a\pi$. Противоречие.

В другую строну. Пусть неверно, что *i*, *w*ё $\otimes_a$**су**$(\pi, \mathrm{T})\wedge B$, но *i*, *w*ё $\langle P_a\pi\rangle B$. Тогда существует *j* такое, что $i\mathbf{R}_a^{\delta}j$, где $\delta=P_a\pi$ и *i*, *w*ё *B*. Тогда *i*, *w*ё $P_a\pi\wedge B$ и, следовательно, неверно, что *i*, *w*ё $\otimes_a$**су**(π, T), откуда по предположению индукции неверно, что *i*, *w*ё $\otimes_a\langle\pi\rangle\mathrm{T}$, откуда по определению $\otimes_a$ неверно, что *i*, w_a*ё $\langle\pi\rangle\mathrm{T}$ и по определению w_a* неверно, что *i*, *w*ё $P_a\pi$. Противоречие.

В доказательстве того, что V(**су**($P_a\pi$, *B*)=V($\langle P_a\pi\rangle B$) мы использовали сведение содержания динамической эпистемической установки к миру w_a*. Отношения модальности $\otimes_a$ с эпистемическими и динамическими модальностями $\Diamond_a$ и $\langle\pi\rangle$ нуждается в дальнейшем уточнении. Введение оператора $\otimes_a$ вызвано тем, что прямое совмещение эпистемического и динамического универсума невозможно. Семантика $\otimes_a$ описана только в общих чертах, но она генерирует новую структуру эпистемических миров, выражающих содержание динамических эпистемических установок.

Трактуя анафору как особый оператор доопределения программы вычисления значения выражения в ЛДО, мы пришли к логике, в которой комбинируются эпистемические и динамические модальности. Она необходима не только для того, чтобы формулировать выражаемые программами пропозиции, но и для осуществления анализа логических отношений этих пропозиций. Рассматривая возможные доопределения и стоящие за ними пропозиции, реципиент осуществляет сложный процесс, в ходе которого анализируются следствия принятия той или иной пропозиции как соответствующей содержанию высказывания. Описание критериев такого процесса, по которым, например, можно было бы выделить наиболее консервативную или наиболее либеральную версию доопределения, возможно в рамках исчисления, адекватного какому-либо уточнению описанной семантики.

ЛИТЕРАТУРА

1. *Groenendijk J.*, *Stokhof M.* Dynamic Predicate Logic // Linguistics and Philosophy. 1991. Vol. 14. P. 39-100.

2. *Heim I.* File Change Semantics and the Familiarity Theory of Definiteness // Meaning, Use and Interpretation of Language / Ed. by R. Bäuerle, C. Schwarze, A.von Stechow. Berlin, 1983. P. 164-189.
3. *Eijck J.van*, *Vries F.-J.de*. Dynamic Interpretation and Hoare Deduction // J. of Logic, Language and Information. 1992. Vol. 1. P. 1-44.
4. *Eijck J.van*, *Vries F.-J.de*. Reasoning about Update Logic // J. of Philosophical Logic. 1995. Vol. 24. P. 19-45.