

## Логика в действии. Введение<sup>1</sup>

*Й. ван Бентем*, Амстердам & Стэнфорд, Декабрь 2003.

**Аннотация.** *Динамический поворот* в логике направлен на то, чтобы сделать понятия коммуникации и информирования объектами логического исследования. В данной работе хотелось бы коротко изложить новую версию этой программы и показать, каким образом могут сочетаться различные идеи из логики, философии, теории игр и computer science. В частности, мы обсудим эпистемические, динамические и временные логики, проблему ревизии убеждений (belief revision) и игровые логики.

---

### 1 Динамический поворот

Классическая логика нацелена на исследование пропозиций, которые мы можем знать или в которых быть уверенными, а также на исследование неизменных отношений следования между ними. Но следование – это в первую очередь некоторая *деятельность*, где пропозиции задействованы на входе и на выходе. В последние годы стало крепнуть мнение, что разнообразные формы деятельности, такие как рассуждение, оценка, ревизия убеждений (belief revision) или коммуникация, сами по себе являются объектами исследования логики, и что их динамическая структура может изучаться чисто логическими средствами<sup>2</sup>. Например: представляется странным исследовать вопрос «знания» некоего утверждения исключительно с точки зрения статики, поскольку «знания» обычно являются результатом различных познавательных процессов, которыми мы постоянно заняты – хотя бы потому, что мы регулярно задаем вопросы и получаем ответы. Так вот, есть основание считать, что спрашивание и получение ответов являются такими же важными логическими формами деятельности, что и извлечение следствий! Эту линию мысли можно продолжить: естественным динамическим дополнением статической эпистемической логике является теория индивидуальных и социальных *механизмов обучения*. Аналогичные траектории, ведущие от статики к динамике, можно было бы проследить, если посмотреть на идею следования (inference) следующим образом: сначала как на «без-агентное» математическое отношение между статичными пропозициями, затем как на одно-агентную форму деятельности, нацеленную на выведение умозаключений, а потом и как на мульти-агентный процесс аргументации. Эту расширенную перспективу логических исследований иногда называют «Динамическим поворотом», начавшимся в восьмидесятые годы с изучения процедур интерпретации в естественном языке, а также с проблемы ревизии убеждений в рамках исследований по «искусственному интеллекту». Но как именно следовало бы включать понятие действия (first-class citizens) в область компетенции логики? Довольно правдоподобным выглядит предположение, что подходящий формальный инструментарий здесь могут

---

<sup>1</sup> Перевод осуществлен К.А. Павловым по статье *J. van Benthem*. A Mini-Guide to Logic in Action, Philosophical Researches, Supp: 21-30, Beijing, 2004. Переводчик и редколлегия благодарят Ontos Verlag, Frankfurt за возможность опубликования статьи.

<sup>2</sup> Различие между «динамическим» и «статическим» имеет смысл и тогда, когда мы расширяем понятие классической логики, например, за счет введения в игру понятия определения или же понятия выразительной силы языка. Выразительная сила языка имеет прямое отношение к *деятельности*, связанной с процессом оценки утверждений, процессом различения между ситуациями, и т.п.

предоставить темпоральная логика, философия действия, а также и динамическая логика, нацеленная на системный анализ работы компьютерных программ. Нельзя не учитывать и влияния, оказываемого теориями процессов (process theory) в computer science, теорией игр и прочими дисциплинами. В этой статье я сосредоточусь лишь на той траектории «динамификации» логики, которая отражает мои собственные интересы. Значительно более широкий обзор этой темы можно найти в [van Benthem, 1996].

## 2 От эпистемической логики к коммуникации

**Вопросы и ответы** Симптоматичная иллюстрация, характеризующая Динамический поворот, возникает, к примеру, в рамках эпистемической логики. Давайте попробуем выйти за рамки привычных вопросов, в пределах которых большинство из нас было воспитано, связанных с проблемами типа «является ли знание истинным верифицированным убеждением?», или «какие модальные аксиомы следовало бы выбрать для работы с эпистемическим оператором  $K$ ?». Вместо этого лучше посмотрим на самые распространенные коммуникативные ситуации. Предположим, что я задал вам простой вопрос, предполагающий ответы либо да, либо нет:

«Находится ли Амстердам на той же широте что и Пекин?»

и предположим также, что вы мне ответили правдиво (с вашей точки зрения). Между прочим, правильным ответом бы было сказать «нет»<sup>3</sup>. Обратим теперь внимание на некоторые сложности, скрытые за обманчивой простотой этого примера. При нормальных обстоятельствах мой вопрос является уместным, только если имеют место *определенные условия* (т.н. «входные данные», preconditions). Во-первых, спрашивая, я тем самым даю понять, что правильный ответ мне неизвестен. Но не только это. То, что я спрашиваю именно *вас*, означает, что с моей точки зрения является бессмысленным предполагать, что вы можете знать ответ на заданный вопрос<sup>4</sup>. Теперь об ответах. Отвечая, вы делаете меня осведомленным о некоем факте  $P$ . Помимо этого вы теперь знаете, что я знаю  $P$ , а я знаю что вы знаете, что я знаю, и т.д. Мы достигаем, таким образом, того, что в логической и психологической литературе называется *общим знанием* (common knowledge). Это – так называемые «*выходные данные*» (postconditions), полученные в результате правдивого ответа<sup>5</sup>.

Заметим, между прочим, что в большинстве случаев входные и выходные данные несут в себе знание о знании других людей. Это может показаться несколько избыточным – побочным – социальным эффектом коммуникации. Но в действительности эти итеративные уровни знания о знании нередко оказываются решающими в плане успешности совершаемых действий. Предположим, что я знаю не только то, что вы украли у меня часы и теперь носите их, но также и то, что вы не знаете, что я это знаю. Скорее всего, в этом случае мне достаточно будет рассчитывать на то, чтобы просто отобрать их у вас. Однако в том случае, если вы совершенно открыто носите украденные часы, фактически давая мне знать, что вы знаете, что я знаю, что они у вас (обратите внимания, что здесь целых 3 итерации!), то это есть повод для меня задуматься о каких-то более замысловатых стратегиях возвращения

<sup>3</sup> По крайней мере, если верить маленькому глобусу, стоящему у меня на столе.

<sup>4</sup> Это относится к типу «кооперативных» вопросов. Ни одно из перечисленных условий не работает, скажем, в случае дидактических вопросов, задаваемых учителем в классе, или же в случае *игр*, там где намеренное сбивание с толку может иметь место.

<sup>5</sup> Термины «входные данные» и «выходные данные» распространены в программном анализе.

собственных часов. Иными словами, успешная коммуникация и адекватные практические действия напрямую зависят от аккуратности нашего обращения с различными оттенками смысла понятия «знание».

**Эпистемическая логика.** Входные и выходные даны предыдущего эпизода могут быть записаны в терминах стандартной эпистемической логики, формализм которой вполне позволяет кодировать знания о фактах и о той информации, которой владеют другие люди. К примеру, с помощью эпистемического модального оператора  $K$ , читаемого как «знает что», и дуальной модальности  $\Diamond$ , означающей «считает возможным что», предыдущая ситуация может быть описана так:

$$\neg K_I P \ \& \ \neg K_I \neg P \\ \langle I \rangle (K_{you} P \ \vee \ K_{you} \neg P)$$

Более того, после получения ответа, становятся истинными еще и следующие утверждения:

$$K_{you} P, \ K_I P, \ K_{you} K_I P, \ \text{и т.д.}$$

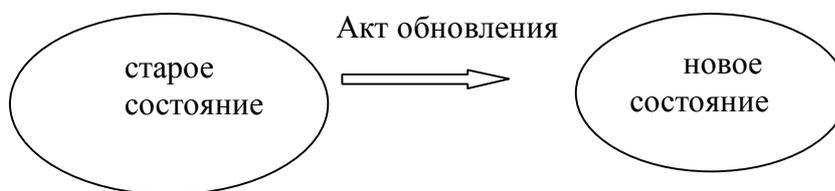
вплоть до утверждения об общем для нас обоих знании (common knowledge)  $C_{I, you} P$ .

Говоря точнее, подобные эпистемические формулы отсылают к типичным семантическим моделям

$$M = (W, \{\sim_j\}_j, V)$$

для эпистемической логики, состоящей из множества возможных миров  $W$  (описывающих возможные состояния действительного мира), отношений достижимости  $\sim_j$  для каждого агента  $j$ , и оценочной функции  $V$ , присваивающей каждой пропозиции истинностное значение в каждом из возможных миров. В таком случае формула  $K_j P$  истинна в мире  $s$  если  $P$  истинно во всех мирах  $t$ , таких что  $s \sim_j t$ . Значительно более сильная формула  $C_G P$  (выражающая общее знание) является истинной тогда, когда  $P$  истинно во всех мирах, достижимых из  $s$  за любое конечное число шагов  $s \sim_j t \sim_j k \dots$  с произвольным набором агентов. Иногда бывает удобным считать отношения достижимости  $\sim_j$  просто отношениями эквивалентности, что превращает нашу логику в поли-модальную систему  $S5$  с оператором общего знания. Однако аналогичные идеи будут работать и в случае существенно более слабых логик, моделирующих убеждения агентов, а не их знания.

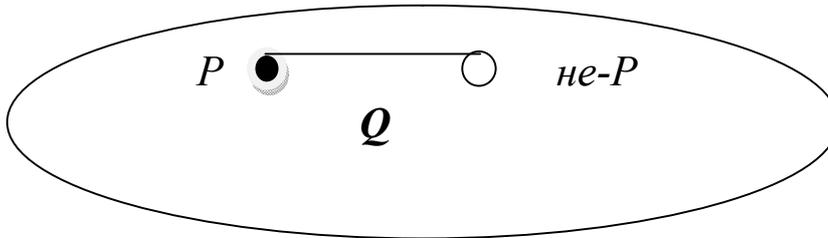
**Динамика: смена информационных состояний.** Но это только начало истории. Полное выявление того, что происходит в вопрос-ответных ситуациях не должно ограничиваться регистрацией истинности утверждений, имевших место до и после обмена информацией. Необходимо также смоделировать процесс смены информационных состояний непосредственно в терминах переходов между ними, представив их как компоненты некоторого объемлющего информационного пространства.



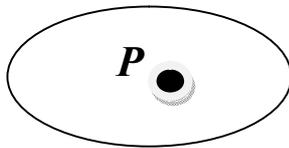
С этой целью нам необходимо будет «динамизировать» традиционную эпистемическую логику. Во-первых, последовательность информационных состояний, соответствующих порядку развертки любого разговора, может быть представлена эпистемическими моделями  $(M, s)$ , с миром  $s$  означающим реальное положение дел. Эти модели представляют собой как бы «мгновенные снимки» тех фрагментов

информации, которые на данный момент доступны рассматриваемым агентам. Обычно фиксируется какая-нибудь одна модель  $M$  и затем производится оценка истинностных значений формул  $\phi$  в некотором возможном мире. Однако в нашем случае необходимо будет рассматривать целые последовательности таких моделей, поскольку речевые акты (в которых что-то утверждается или отрицается) приводят к изменению модели в соответствии с определенными правилами обновления (update rule).

Рассмотрим простейший пример вопрос-ответной ситуации, в которой Спрашивающий ( $C$ ) не знает о том, имеет ли место  $P$ , а Отвечающий ( $O$ ) заведомо знает правильный ответ:

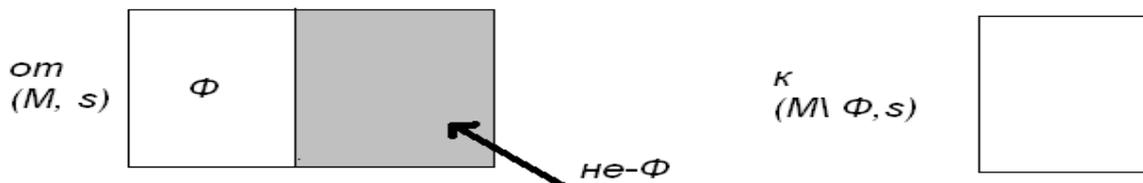


Чёрная точка соответствует действительному положению дел. (В данной модели подразумевается также и то, что – как того требуют правила эпистемической логики –  $C$  с определенностью знает, что  $O$  знает ответ на задаваемый  $C$  вопрос; тем не менее ясно, что в общем случае вовсе не обязательно, чтобы это ограничение имело место). На следующем шаге, в результате ответа  $O$  на заданный ему вопрос, вследствие элиминации варианта  $не-P$  происходит смена информационной модели, и в итоге имеет место следующая диаграмма:



На данном этапе оказывается, что  $P$  стало общим для  $C$  и  $O$  знанием.

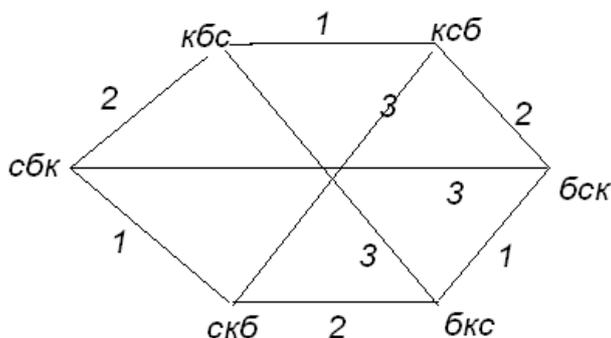
В общем случае динамика здесь может быть описана следующим образом. Публичное оглашение  $\phi!$  истинного утверждения  $\phi$  элиминирует все те миры, достижимые из наличного мира, которые не удовлетворяют  $\phi$ :



Там, где эпистемические модели оказываются значительно более сложными, подобная элиминация может сыграть существенную роль.

**Игры.** Карточные игры являются весьма показательным примером нетривиального обмена информацией даже в самых простейших случаях. Пусть имеется три игрока 1, 2 и 3, и три типа карт – красные ( $\kappa$ ), белые ( $\delta$ ) и синие ( $\sigma$ ), которые оказались розданными

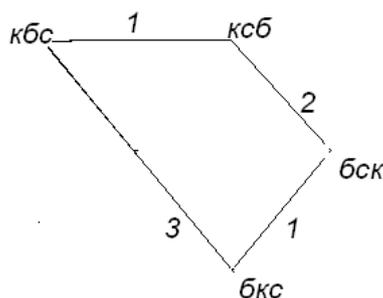
игрокам в следующем порядке: *кбс*. Каждый игрок видит только свою карту. Эпистемическая модель иллюстрируется следующей диаграммой:



Эту диаграмму следует читать следующим образом. Несмотря на то, что на руки сдана комбинация *кбс*, ни один игрок этого не знает. Поэтому анализ ситуации должен включать в себя все 6 возможностей (6 миров). Предположим теперь что

игрок 1 правдиво говорит: «У меня нет синей карты».

Что игроки теперь знают о розданных им картах? Устраняя из диаграммы комбинации, начинающиеся с «с», мы получаем следующую диаграмму:



Отсюда сразу становится понятным, что игрок 2 может знать действительный расклад, игрок 3 знает, что 2 это знает, а игрок 1 знает только то, что либо 2 либо 3 могут знать правильный ответ. Заметим, между прочим, что знание игрока 2 нельзя считать «общим»! Ведь первый игрок считает возможной такую ситуацию, что синяя карта может находиться у игрока 2, и в этом случае одного лишь первого высказывания явно не достаточно, чтобы помочь ему распознать действительный расклад. Диаграмма позволяет предсказать и дальнейшие варианты событий. Например, если бы теперь игрок 3 честно сказал «Я всё еще не знаю ответа», то это позволило бы оставить на диаграмме только самые левые миры, и игрок 2 сразу бы сумел вычислить действительный расклад.

**Более общий случай обновления информации.** Подобные модели помогают прояснить принцип устройства таких головоломок как «Чумазные дети» и проч., как показано в Fagin, Halpern, Moses & Vardi 1995. Простое изложение этих идей и тех общих вопросов, которые отсюда вытекают, можно найти в van Benthem 2001. Такого рода задачи послужили отправной точкой для целого ряда исследований, нацеленных на изучение более изощренных форм коммуникации, включающих в себя утаивание, забывание и намеренное введение в заблуждение. Здесь может сочетаться как публичная, так и приватная информация (как в случае с защищенными протоколами в

Интернете), где агенты коммуникации могут систематически вводиться в заблуждение. Одной из лучших систем такого рода является система обновления продуктов (*product update*) для состояний с действиями: см. Baltag-Moss-Solecki 1998, van Ditmarsch 2000<sup>6</sup>.

### 3 Логика эпистемических процессов

*Динамическая логика* Пришло время поговорить об исчислениях, отражающих природу процессов обновления, и способы рассуждения о них. Адекватной двухуровневой статично-динамической системой, отвечающей целям Динамического поворота, могла бы послужить система, заимствующая определенные идеи из computer science, главным образом, идею сосуществования пропозиций и выражений действия в *динамических логиках*.

Эти логики позволяют описывать истинные положения дел, возникающие по ходу совершения определенных действий:

$[a]\phi$  это читается так:  $\phi$  имеет место после каждого удачного осуществления действия  $a$ .

По аналогии, мы можем фиксировать эпистемические эффекты коммуникации, такие как

$[A!]K_j\phi$  это читается так: после правдивого публичного оглашения  $A$ , агент  $j$  знает, что  $\phi$ .

Этот комбинированный язык сочетает в себе модальности динамической логики с эпистемическими модальностями. Порядок их употребления регистрирует процесс взаимодействия входных данных с выходными. Вот пример, который может показаться тривиальным на первый взгляд.

$[A!]C_GA$  публичное оглашение  $A$  приводит к появлению общего знания  $C_GA$ .

Ниже мы еще попытаемся разобраться в том, в какой мере эта формула может претендовать на роль общезначимого логического закона коммуникации. Другим примером может служить формальный принцип коммуникации, соотносящий исходное знание агентов коммуникации с тем знанием, которое может иметь место после публичного оглашения  $A$ :

$$[A!]K_j\phi \leftrightarrow (A \rightarrow K_j(A \rightarrow [A!]\phi))$$

Эта формула утверждает, что знание  $\phi$  (на выходе) соответствует (определенным образом релятивизированной) версии исходного знания о  $\phi$ . Это лишь один из законов рассуждения о коммуникации в рамках полной динамико-эпистемической системы, отражающей логику процессов публичного оглашения, о которой известно, что она является аксиоматизируемой и разрешимой. По-видимому, эта система является простейшим логическим исчислением процедур коммуникации<sup>7</sup>. Для более сложных информационных процессов существуют и более сложные системы их описания. В любом случае, динамико-эпистемическая логика обещает быть более или менее

<sup>6</sup> В общем случае, по ходу развития разговора или игры, эпистемические модели, генерируемые продуктами обновления, могут разрастаться в размерах так, что это может даже приводить к ситуациям, когда будет отсутствовать непосредственная отсылка к общим знаниям о действительном мире!

<sup>7</sup> Это исчисление состоит из базовой эпистемической логики и простых аксиом редукции, рекурсивно разлагающих выходные данные. Но тут имеются некоторые тонкости, поскольку аксиома редукции общеизвестного (уже оглашенного) знания требует обогащения статичного базового языка до более богатой эпистемической логики с оператором  $CG(A, \phi)$  (оператор общепринятого знания, релятивизированного по отношению к мирам, удовлетворяющих  $A$ ). Таким образом, динамическая структурная надстройка может служить способом модификации базовой статичной структуры.

систематическим источником логической таксономии и понимания общих свойств коммуникации.

**Анализ речевых актов** Всё сказанное выше, по-видимому, вызывает к значительному пересмотру старых философских проблем. Рассмотрим, к примеру, *полные эпистемические спецификации* в теории речевых актов. Что нам становится известным благодаря публичному оглашению некоего  $\phi$ ? Перечисленные выше принципы предлагают нам считать, что в результате мы всегда получаем некое *общеизвестное знание* касательно  $\phi$ . Но это, очевидно, не верно в общем случае! Предположим, что на некий вопрос Отвечающий ( $O$ ) ответил:

$\phi$  «Хоть вы и не знаете, но на самом деле  $P$ »,

Это утверждение было бы истинным лишь в момент его произнесения, поскольку информационное воздействие, которое несет с собой самый акт утверждения высказывания  $\phi$ , немедленно привел бы к смене истинностного значения  $\phi$ , потому что вопрошающий  $C$  теперь уже знает, что  $P$ ! Философы, разумеется, легко распознают здесь т.н. парадокс Мура, только воспроизведенный уже в рамках динамико-эпистемической логики<sup>8</sup>.

Иными словами, те логики, которые разрабатываются в рамках Динамического поворота, позволяют подойти к старым проблемам с новыми техническими средствами. Даже в такой простой формуле как  $[A!K_i \phi$  оказываются закодированными идеи, взятые из философской эпистемологии, из лингвистики с ее понятием речевых актов, и логики программирования из computer science.

**Структура программ** Аналогия между коммуникативными актами и компьютерными программами может быть проведена и далее. Компьютерные программы можно представлять себе как последовательность базовых действий, осуществляемых компьютерным «железом» благодаря программным *конструкциям*, таким как

сочетание	$S; T$
условный выбор	$IF P THEN S ELSE N,$
управляемая итерация	$WHILE P DO S.$

В особенности последняя структура является типичной для тех вычислений, когда заранее невозможно сказать, сколько раз компьютеру необходимо будет повторить определенную инструкцию. Но такого рода ситуации изобилуют и в коммуникации. Публичное оглашение тоже является базовой инструкцией, в соответствии с которой происходят изменения информационных состояний, скажем, на уровне наших общественных соглашений или же на уровне нашего мозга. Помимо этого, разумеется, можно говорить и об аналоге компьютерному софту – о «коммуникативных программах».

В процессе коммуникации имеет место обмен комплексными инструкциями типа «Сначала спроси, как у нее дела, а уж потом обращай со своей просьбой», или «Если учитель спросит  $A$ , то нужно сказать  $B$ , а иначе  $C$ ».

В этих инструкциях могут быть задействованы и итерации. Иными словами, процесс разговора можно представлять себе как разновидность императивного программирования, где роль «машин» играют социальные установки, на которые мы оказываем определенное воздействие.

<sup>8</sup> Технический вопрос о том, какого рода формы эпистемических высказываний, по мере их публичного оглашения, продуцируют общеизвестное знание, до сих пор является открытым. О связи этой проблемы с парадоксом Фитча (Fitch paradox) см. van Benthem 2003.

Замечательным примером подобного итерирования может служить хорошо известная головоломка:

*Чумазные дети: После игры на улице, двое из троих детей вернулись с грязью на лбу. Поскольку все они видят друг друга, но ни один из них не видит сам себя, то Отцу приходит в голову мысль задать им такую загадку: «По крайней мере один из вас с грязью на лбу. Кто-либо из вас сумеет догадаться, грязный ты или нет?». В правила игры входит то, что все должны говорить честно. Итак, как будет развиваться ход игры?*

На первом этапе, разумеется, никто из них не знает правильного ответа. Но именно это обстоятельство служит сигналом к размышлению на втором этапе, поскольку каждый из двух детей, вернувшихся домой с грязью на лбу, теперь может рассуждать так:

*«если бы я был чистым, то грязнуля, которую я вижу, видела бы только чистых детей перед ней, и тогда она сразу бы догадалась, что именно она вернулась с грязью на лбу. Но она молчит. Значит я тоже грязнуля!».* Это рассуждение в одинаковой степени проходит для обоих чумазных детей, поэтому на втором этапе они оба могли бы догадаться о том, что именно они вернулись чумазными. А третий ребенок сумел бы вычислить то, что он чистый только на третьем этапе, после того, как первые два огласили бы свои рассуждения.

Эта головоломка с легкостью распространяется на любое количество грязнуль и чистюль. В ее решение всегда будет входить итерация вида «продолжай показывать свое незнание до тех пор пока нет возможности дедуцировать знание», которая может быть повторена произвольное число раз, в зависимости от комбинации детей.

Для того, чтобы произвести полный анализ этой головоломки, необходимо прибегнуть к средствам динамико-эпистемической логики, позволяющей конвертировать сложные действия  $\pi$  в утверждения типа  $[\pi]\phi$ . Аксиомы такого конструирования хорошо известны из программирования, как, например, закон программной редукции (program reduction law)  $[S;T]\phi \leftrightarrow [S][T]\phi$ .

**Общая логика коммуникации** Логика коммуникации – очень обширная тема. В работе van Benthem 2002 исследуются следующие типы новых задач:

Каким образом следует описывать наилучший из возможных исходов, достигаемых группой агентов коммуникации, нацеленных на оптимальное информирование друг друга?

Очень богатым источником является книга Baltag-Moss-Solecki 2003. Там, к примеру, содержится результат, что динамико-эпистемическая логика публичного информирования (logic of public update), допускающей итерации Клини \* на поступающие утверждения, является неразрешимой. Это означает, в частности, что логика, лежащая в основании головоломок типа Чумазных детей<sup>9</sup>, достаточно богата для того, чтобы на ее языке можно было кодировать значимые математические проблемы! И это лишь один из многих «порогов сложности», очерчивающих спектр человеческой коммуникативной деятельности.

<sup>9</sup> а также и в основании тех хмельных полуночных разговоров, когда мы начинаем многократно повторять сами себя...

## 4 Ревизия убеждений и ожиданий

**От обновления к пересмотру** Обновление – это лишь одна из форм нашей когнитивной деятельности. Другим ключевым моментом Динамического поворота является теория ревизии убеждений (belief revision), разработанная в (Gärdenfors & Rott 1995), и базирующаяся на взаимодействии трех процессов:

- (a) *обновления*, т.е. учета новых утверждений
- (b) *стирания информации*, т.е. игнорировании определенных утверждений
- (c) *ревизии убеждений*, т.е. смене исходных убеждений на новые.

Теория ревизии убеждений предлагает рассматривать не только определенные формы представлений информационных состояний, но и сам процесс ревизии основных постулатов, приводящий к альтернативным основоположениям, отражающим как консервативные, так и более радикальные стратегии трансформации убеждений. И здесь речь идет не просто об изменении информации, несомой пропозициями. Меняется также и структура убедительных, с точки зрения данных агентов, связей между возможными мирами, и система их предпочтений, и вообще любой параметр наличной логической семантики, допускающий какую бы то ни было разумную вариацию во времени<sup>10</sup>.

Остается еще открытым вопрос о том, как именно следует синтезировать эти идеи с эпистемическими логиками обновления, обсуждавшимися выше. Один из способов осуществления этого синтеза может оказаться связанным с «динамификацией» условной логики (conditional logic), благодаря тому, чтобы трактовать импликацию  $A \rightarrow B$  так:

*B* истинно во всех наиболее предпочтительных (или наиболее вероятных) мирах, удовлетворяющих *A*.

Итак, динамика совершаемых актов приводит к изменениям в структуре правдоподобных убеждений, а не только лишь к устранению миров или появлению неопределенных связей. На эту тему можно посмотреть Veltman 1996, Aucher 2003. Нечто подобное происходит и тогда, когда предметом нашего внимания становятся простейшие разговоры, как это будет показано ниже на примере игровых ситуаций.

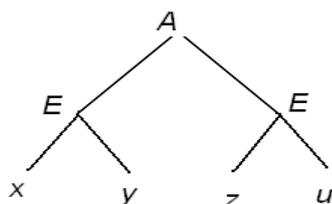
**Теория обучения** Очевидно, что люди придерживаются самых разных стратегий, когда дело доходит до пересмотра теорий или самых обычных мнений. Теория ревизии убеждений не является полноценной динамикой, поскольку сами процессы не представлены в ней в качестве полноправных (first class) объектов исчисления. Примером последнего может служить современная теория *механизмов обучения*, сочетающая в себе идеи из философии науки, математической топологии и computer science. В работе Hendricks 2002 показана неустраняемая значимость этого хода для весьма широкого спектра эпистемологических исследований. Информирование, ревизия и обучение совместно образуют согласованное проблемное поле, возрастающее по своей сложности и охватывающее самые разные формы когнитивного поведения, от краткосрочных до весьма продолжительных. Некоторые вопросы этого типа более детально рассматриваются в van Benthem 2003A.

## 5 Цели, стратегии и игры

**Более широкий взгляд на коммуникацию** Публичные оглашения являются строительными блоками процессов разговора или аргументации. Здесь однако существенным является не только то, *что* люди нам сообщают, но еще и вопрос *почему*

<sup>10</sup> Подвергаться изменениям может даже сам язык, в котором представлены все базовые концепции.

они это делают. Ведь существенным является и то, чего хотят достигнуть собеседники, ибо на этой основе будут определяться и мои цели, в зависимости от которых я буду знать, что говорить или спрашивать. Рассмотрим, к примеру, такой игровой сценарий. Предположим *A* должен скрыто от *E* выбрать либо левую либо правую ветку диаграммы, и затем передать ход игроку *E*. В свою очередь *E* должен загадать одну из букв на левой и правой ветках – т.е одну из *x* или *y*, и одну из *z* или *u*. Понятно, что таким образом оказывается «объективно» заданной однозначно определяемая буква. Помимо этого игроки могут по очереди задавать друг другу вопросы. Выигрывает тот, кто первым вычислит букву, детерминированную исходными данными. Имеется в виду следующая диаграмма:



Предположим, что *A* спрашивает *E*:

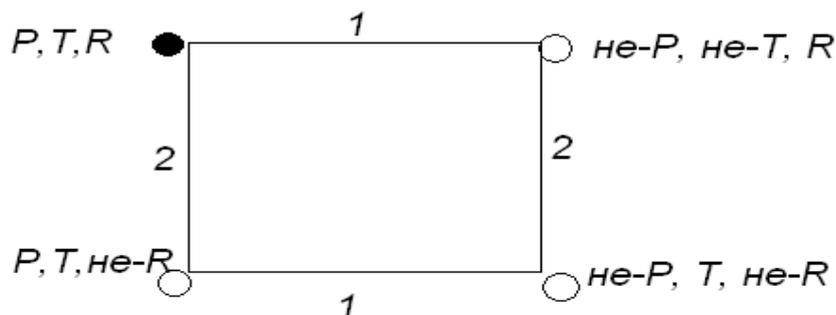
«Какую букву ты выбрал слева?»

Этот вопрос правомерен, так как *A* не знает ответа на него, и знает также, что *E* знает ответ на него. Однако за этим вопросом скрывается еще и дополнительная информация, которую не учитывают ранние версии эпистемической логики, ограничивающиеся подобными входными данными. А именно: *почему* *A* задал этот вопрос? Может показаться, что этот вопрос осмыслен только тогда, если *A* выбрал левую сторону диаграммы. Но тогда эта информация мгновенно бы сообщила *E* о действительном исходе игры, поскольку именно *E* знает итоговую комбинацию букв, и тогда для выигрыша у *E* не было бы даже необходимости задавать встречный вопрос. Напрашивается вопрос: обоснованным ли будет с точки зрения *E* такое рассуждение? Это зависит от множества факторов. Во-первых, от того, как именно оценивает *E* своего партнера: рациональным, тупым, и т.п. Во-вторых, существенным является и система распределения выигрыша. Если, к примеру, оглашение неверного ответа автоматически означает выигрыш противника, то *A* мог намеренно задать сбивающий с толку вопрос, чтобы вынудить *E* к оглашению неверного варианта...

Целью этого простого примера было обратить внимание на идею стратегического взаимодействия, возникающего на каждом очередном этапе разговора, и на необходимость учета возможности возникновения непредвиденных обстоятельств в будущем. Направление, в котором следовало бы развивать исследования по логикам обновления (*update logics*), подсказывается результатами *теории игр*. Фактически игры – это определенные модели поведения группы агентов, стремящихся достичь соответствующих целей посредством взаимодействия. Благодаря играм в дело может быть вовлечено еще два понятия, не использовавшихся ранее: понятие *предпочтений* агентов относительно возможных исходов и понятие долгосрочных *стратегий*, обеспечивающих успешные ответные действия. В частности, именно понятие стратегии позволяет перейти с микроуровня описания на уровень долгосрочного описания поведения.

**Разговорные игры** Рассмотрим беседу двух людей, информированных не одинаково. Предположим, что я не знаю, находимся ли мы уже в Голландии (*P*) или

еще нет ( $\neg P$ ), и продолжается ли еще 2004 год ( $T$ ) или уже нет ( $\neg T$ ). Предположим, что вы точно знаете, что я не знаю наше местонахождение, но думаете, что я всё-таки могу знать точное время. Помимо этого я точно знаю, что мы с вами не просто так оказались в одном месте и одно время ( $R$ ), в то время как вы этого не знаете. Эпистемическая модель данной ситуации представима следующей диаграммой, с черной точкой, указующей на действительное положение дел:



Задача стоит определить действительное положение дел. Выигрывает тот, кто первым сумеет его распознать. Правила игры такие: мы можем по очереди задавать друг другу вопросы, но только вопросы должны быть настоящими – т.е., в частности такими, на которые спрашивающий действительно не знает ответа. На любом этапе тот, кто догадался о правильном ответе, может огласить его и таким образом выиграть. (Здесь может быть предусмотрена и жеребьевка на тот случай, если оба игрока огласят свой ответ одновременно).

Предположим, что я спросил вас о времени. Отсюда вы немедленно заключите, что я не знаю, имеет ли место  $T$ , вследствие чего для вас устранился два нижних варианта диаграммы. Но тогда вам автоматически становятся известными все факты (включая неизвестное вам до этого отношение  $R$ , отрицание которого элиминировалось предыдущим ходом) и вы получите шанс мгновенно выиграть. Отсюда следует, что я должен спросить о месте ( $P$ ), а не о времени. Благодаря этому вопросу вы не получите никакой дополнительной информации, которой бы у вас не было ранее, поскольку этот вариант совместим со всеми четырьмя вершинами диаграммы. Однако ваш положительный ответ тут же устранил два правых варианта, после чего я сумею заключить о реальном положении дел, в то время как вы этого сделать не сможете, поскольку для вас всё еще останется открытым вопрос об отношении  $R$ .

Выбор между удачными и неудачными вопросами (точнее, выбор между тем, что нужно и что ненужно говорить вообще) является краеугольным камнем разговорно-игровой динамики, совершаемой игроками посредством выбора таких вопросов, которые способствовали бы приросту информации у одного игрока, оставляя при этом противника в неведении как можно дольше.

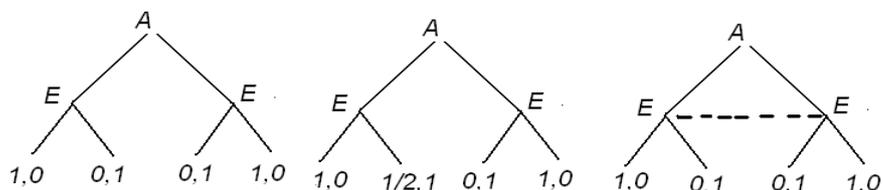
Осуществимо ли это или нет зависит не только от эпистемической модели, но и от графика вопросов и ответов. (Ясно, что в данной игре могли бы выиграть и вы, если первый ход числился за вами). Существенную роль играют также и ограничения по времени<sup>11</sup>.

**Теория игр и логика** Теория игр занимается исследованием разнообразных стратегий, нацеленных на оптимизацию долгосрочного поведения игроков, выигрыш которых определяется либо в терминах равновесия Нэша, либо в терминах иных

<sup>11</sup> Вопрос о стратегиях успешного вопрошания в разговорных играх очень обширен; исследование подобных игр требует привлечения вероятностных методов (см. например van Rooy 2003).

подходящих критериев. В рамках такого подхода можно говорить практически о любых конкретных играх (экономика, война, состязание), а также и о более абстрактных феноменах, где игровая терминология может использоваться для исследования таких социальных действий как употребление языка или процесс логического рассуждения. Значительная часть математики, которую привлекают для теоретико-игрового анализа, нацелена как раз на поиск равновесия и анализ его свойств, проблематичность которых проистекает из того, что игроки обладают неодинаковым доступом к информации. В этом направлении уже разработано множество техник, начиная с анализа типа «от вершин-к-корню» (leaf-to-root analysis) до существенно более сложных результатов (Osborne & Rubinstein 1994). Несмотря на очевидные различия в сфере деятельности и целях, теория игр и логика всегда имели множество пересечений. В работах van Benthem 1999-2003 представлена панорама тех игр, которые оказываются причастными логике благодаря процессам моделирования семантических оценок, аргументации и другим формам деятельности. Идея *логической игры* может оказаться полезной при использовании игры как модели интерактивных вычислений. В результате мы получим слияние логического исчисления для программ и логического исчисления, определяющего ход игр, и сумеем изучить их вычислительные возможности (Parikh 1985, Abramsky 1998).

**Игровые логики** Другим аспектом взаимосвязи логики и теории игр являются логические исследования того, каким образом игроки должны рассуждать, принимать решения и осуществлять соответствующие действия. По сравнению с более ранними версиями логики обновления, случай с играми требует принятия во внимание процесса абстрагирования. Пусть у нас имеется дерево, отображающее последовательность всех возможных игровых комбинаций, вершины которого обозначают ходы игроков. Мы хотим проанализировать процесс выбора последовательности действий, который могли бы осуществлять игроки, способные к рефлексии над своими стратегиями. Рассмотрим для примера три следующих игровых дерева, с парами чисел внизу, соответствующими исходам игры для  $A$  и  $E$  соответственно:



Каждая из игр является моделью модальной логики, соответствующей основным типам игровых действий – «левых» и «правых» в данном случае. Игровые структуры и стратегии могут быть сформулированы здесь в стандартных терминах. К примеру, исходя из 4 возможных стратегий в первой игре (a), наилучшей стратегией для  $E$  была бы следующая:

«если  $A$  идет направо, то иди налево – если  $A$  идет налево, то иди направо».

#

Эта стратегия является простой программой, для исследования которой вполне достаточно средств стандартной динамической логики (van Benthem 2001B). Интерпретируя значение «1» как «выигрыш», мы видим, что это является *выигрышной стратегией* для игрока  $E$ : следуя ей,  $E$  выигрывает у  $A$  независимо от того, что делает  $A$ . Большинство логических игр не идет далее этого понятия. Однако во второй игре (b), с более сложным набором предпочтений относительно исходов игры, предсказания

могут быть улучшены. В случае «правого» хода  $A$ , игрок  $A$  также будет придерживаться стратегии  $\#$  (при условии, что  $A$  действует рационально). Однако игрок  $A$ , с учетом сказанного, предпочтет сделать «левый» ход, поскольку это дает ему возможность сыграть ничью ( $1/2$ ), в противоположность проигрышу ( $0$ ) при «правом» ходе. Такой расклад предсказывает возможность получения равновесия Нэша, тем не менее, «левый» ход дает возможность игроку  $A$  реализовать выигрышную стратегию. В логических терминах, подобный аргумент вовлекает в себя выражения значений вершин (expressions for values of nodes), и возможно даже полноценную логику предпочтений (preference logic).

И наконец, игра (с) демонстрирует нам новое свойство: *неполную информацию*. На своем ходе игрок  $E$  не знает, какой именно ход был сделан игроком  $A$  (обстоятельство, отображаемое пунктирной линией, соединяющей средние вершины). Ситуация неполной информации имеет место во многих играх, как правило, вследствие ограниченных возможностей наблюдения – как в карточных играх. Подобные игры являются моделями комбинированного динамико-эпистемического языка с базовыми действиями  $a, b, \dots$ , обозначающими ходы игроков, и эпистемическими операторами  $K_j$ , указующими на знания игроков. В этом языке могут быть представлены специфические образцы информации об игре, такие как факт незнания игроком  $E$  того хода, который мог бы привести его к выигрышу:

$$\neg K_E \langle \text{right} \rangle WIN_E \wedge \neg K_E \langle \text{left} \rangle WIN_E$$

На нем также могут быть выражены общие законы, описывающие специальные типы агентов. В работе van Benthem 2001A приводятся типичные случаи взаимодействия логических и теоретико-игровых понятий. К примеру, у игроков  $j$  с *идеальной памятью* в отношении истории игры модель незнания будет удовлетворять следующей перестановочной аксиоме:

$$K_j[a]\phi \rightarrow [a]K_j\phi$$

Это похоже на приводившуюся выше аксиому, связанную с актами публичного оглашения, соотносящую  $[A!]K_j\phi$  с  $K_j[A!]\phi$ .<sup>12</sup> В обоих случаях подразумевается, что все задействованные в игре агенты, обладают идеальной памятью. В противоположность этому случаю, поведение игроков с конечной *ограниченной памятью*, способных помнить историю игры лишь с точностью до некоторого фиксированного «просвета», будет подчиняться другим логическим законам.<sup>13</sup>

**Информационное обновление в играх** В играх, подобных обсуждаемым выше, ранние версии логики обновления всё еще имеют смысл. Интуитивно говоря, происходит следующее. Игроки продвигаются по игровому дереву по мере осуществления ходов. С каждым шагом увеличиваются их знания о происходящих событиях, в то время как горизонт возможных будущих сценариев постоянно сужается. Для начала рассмотрим первый процесс.

<sup>12</sup> Аксиома логики обновления устанавливает отношение *эквивалентности* между операторами  $[K]$  и  $K[ ]$ . А новая импликация отображает еще и то, что устранение незнания у игроков никогда не происходит «спонтанно».

<sup>13</sup> На тему взаимодействия логики и теории игр см. также van Benthem 1999-2003, 2003B; а также на тему силы игроков, структурных понятий эквивалентности игр, операций конструирования новых игр и их алгебр, и анализа теоретико-игрового равновесия в терминах «логики неподвижных точек» (fixed-point logics). Множество прочих интересных направлений исследования можно найти в Stalnaker 1999, Pauly 2001.

Игры с неполной информацией кодируют структурную неопределенность процесса игры, систематически обновляющейся с каждым наблюдаемым ходом. Эту задачу можно пытаться решать средствами эпистемического обновления, механизм которой позволяет объяснять появление пунктирных линий (линий неопределенности) на вышеприведенных диаграммах. Движение начинается от корня дерева, возможно, с учетом наличия некоей исходной эпистемической модели  $M$ . В общем случае у игроков имеется лишь частичный доступ к наблюдению действий, совершаемых другими игроками. Это обстоятельство может быть закодировано с помощью эпистемической модели действий  $A$ , состоящей из конкретных событий, связанных отношениями неопределенности.

К примеру, я вижу, что вы сносит карту, но я при этом знаю лишь то, что вы сносит либо червовую королеву, либо пикового короля. Оба действия предусмотрены в  $A$ , но для меня здесь имеет место отношение неопределенности между соответствующими событиями. Последовательные слои игрового дерева есть результат последовательного вычисления update products в смысле разделе 2:

$M, M \times A, (M \times A) \times A$ , и т.д.

Учитывая специфику этого механизма обновления, можно показать, что возникновение пунктирных линий в полном игровом дереве основывается на ряде допущений (одним из которых является идеальная память), которые можно выписать в явном виде. Более подробно это рассматривается в van Benthem 2001A.

**Координация ожиданий** Однако же обновление информации, сопутствующее ходу наблюдаемых событий, отнюдь не исчерпывает того многообразия явлений, которые задействуются в процессах рассуждения в играх. Даже тогда, когда игроки в совершенстве представляют себе игровую структуру, включая всю цепочку уже осуществленных ходов, к сути игры относятся еще и ожидания игроков относительно своих и чужих образов действия. Такого рода предвосхищение вообще является одним из ключевых моментов любой формы человеческой деятельности. Теоретико-игровое понятие стратегического равновесия как раз связано с идеей устойчивого прогноза. Разумеется, существуют самые разные типы ожиданий. Предположим, что у вас есть основания считать, что моя память вмещает лишь один бит информации, т.е. что у меня в памяти остается лишь последний ход игры, и что мои действия зависят лишь от самого последнего вашего действия. Однако, возможно, что по ходу игры придется пересмотреть некоторые из ваших ожиданий. Предположим, что  $E$  ожидал от  $A$  «левого» хода в игре (b), однако игрок  $A$  вместо этого сделал «правый» ход. В этом случае  $E$  будет вынужден пересмотреть свои ожидания относительно стратегии игрока  $A$ . Иными словами, мы здесь входим в область *ревизии убеждений*, которая вкратце уже упоминалась в разделе 4. Таким образом, можно сказать, что логические исследования игр находятся на пороге нового проблемного поля: необходимости учета и соответствующего синтеза типов анализа – обновления информации и координации ожиданий.<sup>14</sup>

<sup>14</sup> Абстрактные игры и процессы обновления имеют многочисленные точки соприкосновения с ситуациями конкретного разговора. Допустим, что имеется игровое дерево и игроки уже определились с выбором стратегий, но теперь задача заключается в том, чтобы распознать, где именно завершится игра. Выигрывает тот, кто первым догадается об этом. Это, конечно же, является игрой с неполной информацией, поскольку информация здесь раскрывается постепенно по мере вопрошания и получения ответов. В частности, признание в

## 6 Эволюция во времени

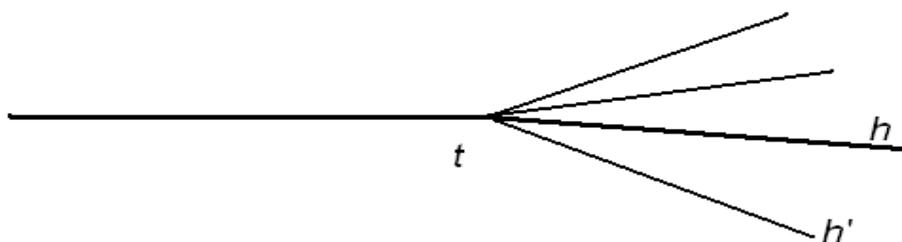
Мы начали с рассмотрения логики простых актов коммуникации и соответствующих им обновлений «информационных состояний», совершаемой агентами коммуникации. Далее мы рассмотрели закономерности более длительных форм игрового поведения, где игроки пытаются добиться своих целей посредством конечных последовательностей действий, осуществляемых в ответ на действия противников. Это уже потребовало введения в игру более сильных типов логики, включающих в себя рассуждения о стратегиях. Но и это еще не всё: ведь коммуникация и игры включены в еще более широкий временной контекст, охватывающий самые разнообразные человеческие практики. Ниже мы коротко рассмотрим некоторые аспекты этой более широкой перспективы.

*Конечное vs. бесконечное* Игры, как и разговоры, и доказательства, ограничены во времени. Однако computer science изучает также и различные бесконечные процессы работы операционной системы, позволяющие осуществлять массу специальных конечных заданий. Подобную дихотомию можно усмотреть и в отношении когнитивных процессов в рамках Динамического поворота. Некоторые типы деятельности подразумеваются конечными во времени, а некоторые же играющими роль оперативной системы для первых. Примерами последних являются логические системы доказательств, или же, скажем, хорошо известные максимы Грайса, организующие возможность разговора. По аналогии с этим, теория игр изучает также и бесконечные игры вместе с соответствующими формами поведения игроков, такие как Дилемма заключенного, постоянно возобновляемая в рамках социального взаимодействия.

*Временная логика* Для изучения подобных феноменов логические системы, описанные выше, необходимо погрузить в подходящие временные системы, которые позволили бы ввести в игру, например, понятие эпистемического мульти-агентного протокола (и т.п.) Понятие протокола используется для того, чтобы кодировать самые общие закономерности, относящиеся к коммуникации, такие как знание одним агентом того, что его собеседник говорит правду только в половине случаев. Типичной картинкой здесь является ветвящееся дерево:

---

неспособности распознать точку завершения игры, содержит в себе полезную информацию, позволяющую исключить определенные ходы (см. van Benthem 2004).



Подобные временные универсы, с дополнительной эпистемической структурой, являются подходящими аренами, в рамках которых можно пытаться соединить отдельные акты обновления, конечные формы игровой деятельности и соответствующие бесконечные процессы, на фоне которых всё это происходит.<sup>15 16</sup>

**Динамическая логика и динамические системы** Бесконечные аспекты той арены, где разворачивается современная теория игр, приводят нас к *эволюционным* идеям, где коллективное поведение анализируется в терминах равновесных свойств, которыми обладают бесконечные динамические системы, функция переходных состояний которых нередко аналогична функциям биологическим (см. Osborne & Rubinstein 1994). Это существенно иной стиль математического мышления о долгосрочных формах поведения, где устойчивые структуры являются результатом статистических свойств рассматриваемых популяций. В свете этой трудности задача согласования логического подхода, защищаемого в данной статье, с эволюционными процессами становится еще более заманчивой.

## 7 От анализа к синтезу

Последняя существенная деталь Динамического поворота, на которую хотелось бы обратить особое внимание, лежит в иной плоскости. Большая часть логических исследований до сих пор была нацелена на анализ и понимание *данных форм поведения*, говорят ли тут об использовании языка или о процессах рассуждения. Однако не меньший интерес представляет и тот неопровержимый факт, что результатом логических исследований является также и *создание новых способов*

<sup>15</sup> Здесь следует упомянуть модели компьютерных вычислений Fagin et al. 1995, бесконечные игры Abramsky 1996, модель протоколирования сообщений Parikh & Ramanujam 2003, универсам обучающих механизмов Kelly 1996, и, скажем, философскую теорию рассуждения и действия Belnap et al. 2001.

<sup>16</sup> Неопределенность, возникающая между конечными последовательностями действий в этих моделях естественным образом обобщает понятия динамической логики, введенные ранее. Например, в древообразных структурах эпистемическое обновление продуктов гарантирует, что две последовательности  $X$  и  $Y$  неразличимы, если они имеют одинаковую длину и все их соответствующие элементы  $X_i$  и  $Y_i$  неразличимы. Однако те системы, в основе памяти которых лежат конечные автоматы, неразличимы уже с точностью до некоторого фиксированного конечного числа предшествующих состояний.

выражения, рассуждения и вычисления. Хорошо известными примерами из computer science являются формальные языки спецификации, или логические программы. Но аналогичный ход от анализа к *конструированию* имеет смысл во всех когнитивных сферах. К примеру, любая целесообразная процедура голосования является конструктивным фрагментом «социального софта» (Parikh 2002), где нами создаются новые образцы поведения, нацеленные на извлечение определенной выгоды. Даже анализ подобных вещей является очень непростым занятием<sup>17</sup>: что уж говорить о том, насколько трудной может оказаться задача конструирования такого рода процессов! То же самое может быть сказано и о потоке новых игр, непрестанно зарождающихся в нашем мире, которые затем включаются в естественный ассортимент человеческой деятельности<sup>18</sup>. Системы, рассмотренные в данной статье, могут использоваться и в более «активистском» ключе, как способ формирования нового образа действий и, стало быть, изменения мира. Вот пример одной из самых первых логик обновления, так называемая Московская головоломка (van Ditmarsch 2002):

«Игрок *A* получает 1 карту, *B* и *C* получают по 3 карты. Что должны *B* и *C* сказать друг другу в присутствии *A*, чтобы они могли догадаться о выпавшем раскладе, а *A* нет?».

Размышляя о подобных головоломках, возникает вопрос и о процессе создания новых игр, или новых практик, прояснение которого представляется возможным методами динамических логик, средства которых позволяют очерчивать горизонт возможностей, а также целенаправленно держать в поле зрения сферу намеренных следствий.

## 8 Заключение

Задачей статьи было создание максимально широкой перспективы видения логики, какой она рисуется в контексте коммуникации, вычислительной и познавательной деятельности. Это позволяет синтезировать традиционные формы анализа рассуждения и определения с такими идеями как ревизия убеждений, планирование действий, игровые практики, а также рассмотреть эти феномены на фоне закономерностей, возникающих у длительных социальных процессов. Мы привели некоторые примеры того, как это можно было бы осуществить – хотя, разумеется, большей частью мы здесь имеем дело с красивыми пожеланиями, а не с устоявшейся практикой. Тем не менее, даже имеющийся в этом вопросе незначительный опыт подсказывает нам, что мечты имеют шанс осуществиться...

---

<sup>17</sup> Хорошим примером тут является непостижимая процедура выбора Дожа, официально принятая в Венеции в 1268 г., со всеми ее защитными механизмами, призванными обеспечивать учет интереса всех сторон, кланов и партий, и включающими в себя многоступенчатое голосование и жеребьевку партиями. В своей *Истории Венеции* (Norwich, History of Venice, NY, 1989) Норвич говорит, что это «можно расценивать как поистине одно из самых сложных институциональных изобретений цивилизованного человечества».

<sup>18</sup> Сравните также с исследованиями по «стратегическому проектированию» (mechanism design) в рамках современной теории игр.

---

## 9 Литература

S. Abramsky, 1998, 'From Computation to Interaction, towards a science of information', CS/IEE Turing Lecture.

G. Aucher, 2003, *A Joint System of Update Logic and Belief Revision*, Master of Logic Thesis, ILLC University of Amsterdam.

A. Baltag, L. Moss & S. Solecki, 1998, 'The Logic of Public Announcements, Common Knowledge and Private Suspicions', *Proceedings TARK 1998*, 43-56, Morgan Kaufmann Publishers, Los Altos. Updated version, 2003, department of cognitive science, Indiana University, Bloomington, and department of computing, Oxford University.

N. Belnap, M. Perloff & M. Xu, 2001, *Facing the Future*, Oxford University Press, Oxford.

J. van Benthem, 1996, *Exploring Logical Dynamics*, CSLI Publications, Stanford.

J. van Benthem, 1999 ÷ 2003, *Logic in Games*, electronic lecture notes, ILLC, Amsterdam & philosophy Stanford (occasional paper versions).

J. van Benthem, 2001A, 'Dynamic Epistemic Logic', *Bulletin of Economic Research* 53:4, 219-248 (Proceedings LOFT-4, Torino).

J. van Benthem, 2001B, 'Extensive Games as Process Models', *Journal of Logic, Language and Information* 11, 289-313.

J. van Benthem, 2001C, 'Logics for Information Update', *Proceedings TARK VIII (Siena 2001)*, Morgan Kaufmann Publishers, Los Altos, 51-88.

J. van Benthem, 2002, 'One is a Lonely Number: on the logic of communication', Tech Report PP-2002-27, ILLC Amsterdam. To appear in P. Koepke et al. eds., *Colloquium Logicum, Muenster 2001*.

J. van Benthem, 2003A, 'Logic and the Dynamics of Information', in L. Floridi, ed., *Minds and Machines* 13:4, 503-519.

J. van Benthem, 2003B, 'Rational Dynamics and Epistemic Logic in Games', in S. Vannucci, ed., *Logic, Game Theory and Social Choice III*, University of Siena, department of political economy, 19-23.

J. van Benthem, 2003C, 'What One may Come to Know', Tech Report PP-2003-22, ILLC Amsterdam. To appear in *Analysis*, 2004.

J. van Benthem, 2004, 'Representing Arbitrary Games through Conversation', manuscript, ILLC Amsterdam.

H. van Ditmarsch, 2000, *Knowledge Games*, dissertation DS-2000-06, ILLC Amsterdam and University of Groningen.

---

H. van Ditmarsch, 2002, 'Keeping Secrets with Public Communication', department of computer science, University of Otago.

R. Fagin, J. Halpern, Y. Moses & M. Vardi, 1995, *Reasoning about Knowledge*, The MIT Press, Cambridge (Mass.).

P. Gärdenfors & H. Rott, 1995, 'Belief Revision', in D. M. Gabbay, C. J. Hogger & J. A. Robinson, eds., *Handbook of Logic in Artificial Intelligence and Logic Programming 4*, Oxford University Press, Oxford 1995.

V. Hendricks, 2002, 'Active Agents', PHILOG Newsletter, Roskilde. In J. van Benthem & R. van Rooy, eds., special issue on Information Theories, *Journal of Logic, Language and Information* 12:4, 469-495.

M. Osborne & A. Rubinstein, 1994, *A Course in Game Theory*, The MIT Press, Cambridge (Mass.).

R. Parikh, 1985, The Logic of Games and its Applications, *Annals of Discrete Mathematics* 24, 111-140.

R. Parikh, 2002, 'Social Software', *Synthese* 132, 187-211.

R. Parikh & R. Ramanujam, 2003, 'A Knowledge Based Semantics of Messages', CUNY New York & Chennai, India. In J. van Benthem & R. van Rooy, eds., special issue on Information Theories, *Journal of Logic, Language and Information* 12:4, 453-467.

M. Pauly, 2001, *Logic for Social Software*, dissertation DS-2001-10, Institute for Logic, Language and Computation, University of Amsterdam.

R. van Rooy, 2003, 'Quality and Quantity of Information Exchange. In J. van Benthem & R. van Rooy, eds., special issue on Information Theories, *Journal of Logic, Language and Information* 12:4, 423-451.

R. Stalnaker, 1999, Extensive and Strategic Form: Games and Models for Games, *Research in Economics* 53:2, 93-291.

F. Veltman, 1996, 'Defaults in Update Semantics', *Journal of Philosophical Logic* 25, 221-261.