# 2. «Непредвзятый» универсальный алгоритмический интеллект

## 2.2. Простейшая модель с ограничением ресурсов

*LSearch и* $\xi^{tl}$

Очевидной причиной невозможности применения универсального предсказания с помощью алгоритмической вероятности, а также моделей ИМИ на его основе, является его невычислимость. Невычислимость связана с двумя обстоятельствами: 1) усреднение проводится по бесконечному множеству моделей; 2) среди этих моделей есть безостановочные алгоритмы (причем в силу неразрешимости проблемы останова во всех случаях невозможно определить, зацикливается ли тот или иной алгоритм на имеющихся данных или нет).

Р. Соломонов указал [Solomonoff, 1986], что большим шагом вперед по использованию алгоритмической вероятности при решении практических проблем является процедура поиска, предложенная Л.А. Левиным (учеником А.Н. Колмогорова) [Левин, 1973] и сейчас называемая LSearch, которая позволяет получать решение любой P или NP проблемы за время, пропорциональное оптимальному времени, умноженному на константный сомножитель.

Общая идея этой процедуры заключается в том, чтобы выделять на проверку каждой алгоритмической модели $q$ время, пропорциональное ее априорной вероятности $\xi(q)=2^{-l(q)}$. При фиксированных ресурсах можно заранее рассчитать, сколько тактов процессорного времени выделить на выполнение той или иной модели (алгоритма). Если же решение ищется до достижения какого-то удовлетворительного качества, то модели могут выполняться параллельно с приоритетами $\xi(q)$ (или в случае последовательного исполнения на каждом такте с вероятностью $\xi(q)$ выбирается модель $q$ для продолжения вычислений).

Очевидно, LSearch требует затрат времени, экспоненциальных по сложности задачи, и может использоваться только для задач очень низкой размерности (это неоднократно отмечал и Соломонов, указывая на то, что LSearch – слепой поиск, лучше которого, однако, для точного решения задач без дополнительной априорной информации ничего предложено быть не может). Но, по крайней мере, на основе LSearch может быть построен вычислимый вариант алгоритмической вероятности (или алгоритмической сложности), который в указанном смысле является оптимальным. В частности, можно отметить, что наличие безостановочных алгоритмов не является проблемой для LSearch, поскольку они не будут отличимы от просто долго работающих алгоритмов, выполнение которых по истечении имеющихся ресурсов тоже не успеет завершиться.

Тонкий вопрос, однако, заключается в том, учитывать ли в предсказании результаты, сформированные алгоритмами, которые не успели остановиться. Случай предсказания особый: от предсказываемых естественных процессов мы вполне может ожидать безостановочности, которая, однако, связана не с постоянной модификацией «истории», а с формированием «будущего». При этом, скажем, алгоритм, который выводит бесконечную последовательность 010101… будет проще алгоритма, который выводит эту же последовательность, но некоторого фиксированного размера. В этом смысле требование от $q$ вывода истории $x_{<k}$ и останавливаться (что делается в AIξ) может быть не вполне естественным. Данная тонкость на текущем уровне детальности моделей универсального интеллекта принципиальной роли не играет.

Поскольку на основе LSearch может быть построена вычислимая модификация алгоритмической вероятности, для которой доказывается оптимальность по времени вычислений с точностью до мультипликативной константы, вполне естественно, что следующее по детальности расширение моделей ИМИ может включать LSearch.

Очевидно, процедура LSearch при ограниченном времени успеет выполнить только более короткие и более быстрые алгоритмы. Несколько обобщая, можно ввести следующее распределение $\xi^{tl}$ на множестве программ, ограниченных по длине и времени выполнения:

$$\xi^{TL}(q)=\begin{cases}0, \neg(q: l(q)\le L \,\&\, t(q)\le T)\\ 2^{-l(q)}, l(q)\le L \,\&\, t(q)\le T\end{cases}. \qquad (2.2.1)$$

Как и раньше, сами алгоритмические вероятности введем как

$$P_U^{tl}(x)=\sum_{q:U(q)=x}\xi^{tl}(q). \qquad (2.2.2)$$

Следует иметь в виду, что распределение, записанное в таком виде, не является нормированным. Нормировку выполнить несложно, но без этого также можно и обойтись. Также отметим, что интересующие нас условные вероятности легко ввести по аналогии с $P_{ALP}(x'|x)$, что даст некий аналог формулы (42) в [Hutter, 2007].

Очевидно,

$$\xi^{tl}(q)\xrightarrow[t,l\to\infty]{}\xi(q),$$

то есть при использовании этого распределения интеллектуальный агент будет стремиться к (неограниченному) универсальному при стремлении ресурсов к бесконечности.

Это распределение может напрямую использоваться для модификации моделей типа AIξ при поиске моделей среды. Совместное увеличение значений $t$ и $l$ может проводиться так же, как в LSearch, чтобы в момент $T$ каждая программа $q$ выполнилась на $2^{-l(q)}T$ шагов; в момент $T$ успеют выполниться те программы, для которых $2^{-l(q)}T>t(q)$, то есть программы окажутся отсортированными по величине $2^{l(q)}t(q)$. При этом первая модель $q$, которая воспроизводит историю взаимодействия агента со средой и выполняет предсказание, будет обладать наименьшим значением $2^{l(q)}t(q)$, логарифм от которой можно назвать *сложностью по Левину* (такое определение сложности имеет глубокий смысл, который требует отдельного обсуждения).

После нахождения модели с минимальной сложностью перебор может не прекращаться до исчерпания вычислительных ресурсов, поскольку универсальное предсказание в моделях ИМИ подразумевает использование многих моделей, совместимых с историей. Однако и при использовании единственной лучшей модели время вычисления может оказаться $O(2^L T)$, где $L$ и $T$ – соответствующие характеристики минимальной модели. Как видно, такая схема вычислений (аналогичная LSearch) имеет замедляющий множитель $2^L$, связанный с тем, что до обнаружения этой модели придется выполнить порядка $2^L$ программ и потратить на каждую из них до $T$ тактов времени. Кроме того, для универсального агента необходимо выполнять предсказание для разных цепочек своих действий (здесь возникает соблазн разделить задачи предсказания и выбора действий, но, как мы покажем, наивное их разделение нарушает универсальность агента).

В любом случае мультипликативная замедляющая константа $2^L$ является гигантской. И, кроме того, в случае реального мира величина $L$ (длина «истинной» модели среды) может быть неограниченной сверху. Как результат, сложность поиска минимальной детерминированной модели среды будет экспоненциально возрастать с увеличением длины истории. А в случае ограниченных ресурсов это сделает недостижимой сходимость поведения агента, основанного на распределении $\xi^{tl}$, к поведению агента, использующего некое «правильное» априорное распределение $\mu$ (поскольку сходимость в случае $\xi$ как раз и возникала при увеличении объема наблюдательных данных, когда этот объем начинает превосходить сложность источника данных).

М. Хаттером вводится модель $AI\xi^{tl}$ [Hutter, 2005]. Интересно, однако, то, что в ней не просто используется $\xi^{tl}$ вместо $\xi$, но и предлагается метод поиска, оптимальный по времени выполнения с точностью до аддитивной, а не мультипликативной константы. Основная идея этого метода заключается в том, чтобы перебирать не все, а только те алгоритмы, для которых доказано, что предсказываемые ими значения целевой функции не являются завышенными. При этом перебираются не модели среды, а программы самого агента, которые, правда, выводят не только действия, но и значения подкрепления (присутствующие в истории и ожидаемые). Доказательства генерируются в некоторой формальной системе отдельным алгоритмом, работающим параллельно с процедурами исполнения моделей.

Некоторое сомнение вызывает возможность строгого конструктивного доказательства требуемого свойства для программ рационального поведения без перебора прочих программ. В этом смысле выглядящая привлекательной замена мультипликативной замедляющей константы на аддитивную на практике может оказаться неэффективной. В любом случае, это дополнительное время, требуемое на поиск доказательств, является слишком большим, чтобы реальный ИИ мог действовать в соответствии с этим подходом. Поскольку в будущем мы не будем использовать саму модель AI$\xi^{tl}$, здесь не приводятся детальные разъяснения на ее счет, за которыми лучше обратиться к первоисточнику.

*Эффективный и прагматический универсальный интеллект*

В самом начале мы упоминали о данном в [Wang, 2007] понятии эффективного интеллекта, как интеллекта, выполняющего свое предназначение в условиях ограниченных вычислительных ресурсов и информации. AI$\xi^{tl}$ работает в условиях ограниченных ресурсов и информации (среда полагается априорно неопределенной), причем оптимально по времени с точностью до некоторого замедляющего фактора. Однако подобная оптимальность является слишком слабой. Но можно ли ее улучшить?

Явным ограничением AI$\xi^{tl}$ является то, что выбор каждого последующего действия при увеличении истории не упрощается, а усложняется, поскольку увеличивается длина истории, а вместе с ней и сложность наикратчайших алгоритмов, способных воспроизвести эту историю (конечно, это замечание не относится к AI$\xi$, в которой перебираются все алгоритмы, что, однако, абсолютно нереалистично). Выбор каждого нового действия рассматривается независимо от предыдущего посредством одного и того же метода. Как результат, константный замедляющий фактор не уменьшается от такта к такту, то есть, по сути, он умножается на то, сколько раз агент осуществляет выбор действия. Вполне естественно рассматривать не просто оптимальность отдельно каждого выбора действия, но и всего процесса функционирования агента. Эффективность интеллектуального агента может, в принципе, повышаться от такта к такту. Указанный аддитивный замедляющий фактор при этом не исчезнет, а просто перестанет умножаться на число тактов. И этот результат в принципе улучшить нельзя. Его можно охарактеризовать в смысле оптимальности по Парето, как и распределение $\xi$: не существует алгоритма интеллекта, который был бы не хуже данного во всех средах и лучше хотя бы в одной.

Можно, конечно, ослабить требование к сходимости поведения агента к оптимальному. Ведь обычно требуется не оптимальное поведение, а просто достаточно хорошее. Однако в пределе (при бесконечных ресурсах) нет причин отказываться от оптимального поведения. Проблема, скорее, в использовании оптимальности по Парето.

Действительная эффективность для всех сред сразу достигнута быть не может. Однако нас не интересуют вообще все возможные среды. Вернее, нам бы хотелось сохранить оптимальность по Парето и универсальность моделей ИМИ, но добавить к этому более высокую эффективность для нашего конкретного мира. В этой связи полезным является понятие прагматического (эффективного) ИИ, введенного в [Goertzel, 2010]. В более привычной терминологии такой ИИ следует назвать «предвзятым» (biased). Подчеркнем, что «предвзятость» вовсе не обязательно означает потерю универсальности. Парето-множество включает многие варианты оптимального ИИ, из которых какие-то оказываются гораздо более «прагматическими» (то есть эффективными в нашем конкретном мире), чем просто некоторый произвольно выбранный вариант.

В частности, мы отмечали, что распределения $\xi(q)$ могут быть разными в зависимости от выбранной универсальной машины. Выбор этой машины не имеет значения с точки зрения оптимальности по Парето, но имеет большое значение для увеличения прагматичности ИИ.

Вернемся к нашим методологическим принципам, введенным в самом начале и гласящим, что ИИ должен быть универсальным, и эта универсальность должна сохраняться при вводе ресурсных ограничений, во многом определяющих архитектуру реального эффективного ИИ, которая также должна основываться на априорной информации о мире для ускорения обучения. Теперь мы можем пояснить, что они означают создание универсального эффективного прагматического ИИ. Касательно $\xi(q)$ сохранение универсальности означает $\xi(q)\neq 0$ для любого алгоритма $q$; введение априорной информации для увеличения скорости обучения (увеличение прагматичности) означает выбор универсальной машины, определяющей наиболее адекватное (нашему миру) априорное распределение $\xi(q)$; учет ресурсных ограничений (увеличение эффективности) означает использование распределения $\xi^{tl}(q)$ и его уточнение по мере получения новой информации.

*Выбор опорной машины*

Возможность выбора опорной машины для внесения априорной информации уже отмечалась разными авторами [Solomonoff, 2010], [Pankov, 2008], и даже указывалось, что ее выбор влияет на «относительную интеллектуальность агентов» [Hibbard, 2009]. Но что это означает на практике?

Есть опасность понять выбор опорной машины слишком буквально, то есть выбор, скажем, между машинами Тьюринга, лямбда-исчислениями, формальными грамматиками, алгорифмами Маркова и т.д. Все эти формализмы, действительно, задают разные $\xi(q)$, но вряд ли стоит ожидать, что один из них существенно «прагматичнее» других. Аналогичное понимание выбора опорной машины, но в несколько более практическом варианте, приводит к попытке выбора (или создания) наиболее подходящего языка программирования. Действительно, любой язык также задает некоторое $\xi(q)$. И выбор подходящего языка может облегчить для конкретных случаев задачу построения алгоритмических моделей. Существуют различные работы, исходящие из сходных соображений, например [Hall, 2008], [Looks, 2009].

Все это, конечно, имеет отношение к увеличению прагматичности универсального ИИ. Даже интерпретация некоторых строк данных как чисел с введенными для них операциями уже задает немалое индуктивное смещение, облегчающее построение моделей нашего мира. И даже просто использование упомянутых выше формализмов уже обеспечивает некоторое индуктивное смещение в направление нашего мира, поскольку они имеют относительно простую физическую реализацию (то есть сам мир является такой универсальной «машиной», которая «эмулирует» другие универсальные машины, имеющие разную сложность конструкции). Однако даже самый беглый анализ задачи, скажем, анализа изображений покажет, что такого индуктивного смещения далеко не достаточно. Действительно, минимальная программа, порождающая некоторое изображение и написанная на каком-то более или менее обычном языке, будет иметь огромный размер, и построить ее прямым перебором будет невозможно.

В мозгу имеется огромное число специализированных процедур обработки сенсорной информации, без априорного задания которых интеллект не может быть прагматическим. Аналогичный вывод можно сделать и про другие, как низкоуровневые, так и высшие когнитивные функции. Может даже показаться, что только этими специализированными (возникшими для нашего мира) функциями только и стоит заниматься, а модели универсального ИИ оказываются как будто нерелевантными. Но мы в очередной раз повторим, что для создания сильного ИИ нужно всеми силами пытаться сохранить свойство универсальности и не поддаваться искушению идти по более простому пути эффективного прагматического, но не универсального интеллекта. Действительно, для человеческого интеллекта (несмотря на всю его «прагматичность») характерна универсальность даже в самом малом. Без такой универсальности человек никогда бы не смог увидеть, скажем, битого пикселя на мониторе или креста, образованного звездами на небе. Ни одна система компьютерного зрения не способна формировать столь произвольные зрительные концепты, если они не заложены априорно в ее пространство описаний. Вполне прозрачная причина этого в неуниверсальности.

Таким образом, естественный интеллект обладает огромной «предвзятостью» (в частности, в форме индуктивного смещения), обеспечивающей его «прагматичность» и определяющей его структуру. Это смещение можно описывать в терминах задания опорной машины при определении $\xi(q)$, однако такая трактовка не должна пониматься упрощенно. Чтобы подчеркнуть это, будем эту «предвзятость» называть *когнитивным смещением* (cognitive bias), поскольку это смещение у человека реализовано в когнитивных функциях (таких как память, внимание, планирование и т.д.), отсутствующих у моделей ИМИ.

Отметим, что понятие когнитивного смещения (когнитивных искажений) уже «занято» в психологии. Там оно носит в большей степени негативный характер, указывая на систематические «ошибки», совершаемые человеком в процессе мышления. Особенно примечательны ошибки в рассуждениях о вероятностях, что как будто свидетельствует о том, что человеческое мышление не использует правил статистического вывода. Когнитивные искажения заслуживают отдельного рассмотрения. Здесь же отметим, что вводимое нами понятие когнитивного смещения хотя и более общее, чем понятие когнитивных искажений в психологии, но между ними противоречий нет. Напротив, можно надеяться, что они позволять прояснить содержание друг друга.

*Адаптивность и самооптимизация*

Выше было отмечено, что рассмотренные модели ИМИ по своей структуре являются неинкрементными: каждая из них представляет собой вызов в бесконечном цикле одного и того же алгоритма; от такта к такту меняется только входная строка для этого алгоритма (представляющая собой историю взаимодействия агента со средой). Интуитивно понятно, что ИМИ в такой форме будет многократно проделывать работу по предсказанию $x$ и выбору $y$. Если в интеллект можно закладывать априорную информацию путем выбора распределения $\xi(q)$ в целях повышения эффективности, то вполне естественно модифицировать это распределение и по мере накопления информации.

Подобное инкрементное обучение (правда, в более узком контексте задачи индукции) рассматривал и Соломонов [Solomonoff, 2003]. Его подход содержит важные идеи (которые мы еще обсудим), но в целом он эвристичен (его идеи проистекают во многом из эвристического программирования) и недостаточен.

Сходные идеи высказывал и Шмидхубер, предлагая адаптивный LSearch, идея которого заключается в том, что как только найдена программа $q$ при нахождении решения текущей задачи, вероятность $\xi(q)$ существенно увеличивается; однако сам способ модификации $\xi(q)$ не гарантирует оптимальности [Schmidhuber, 2007a].

Шмидхубер предложил [Schmidhuber, 2007b] иной подход (названным им «машина Гёделя»), подразумевающий строго обоснованную самооптимизацию (включающую не только распределение $\xi(q)$, но и другие компоненты интеллекта агента, и даже целевую функцию). Однако хотя этот подход и выглядит более теоретически обоснованным, в нем акт самооптимизации выполняется только тогда, когда может быть строго доказано, что он ведет к увеличению будущего подкрепления. Хотя это гарантирует, что подобная самооптимизация, по крайней мере, не ухудшит поведение агента, сомнительно, что она окажется достаточно эффективной, поскольку логически доказать, что те или иные шаги самомодификации приведут агента к гарантированному увеличению будущего подкрепления, вряд ли будет возможно во многих случаях в силу индуктивного характера самой проблемы (аналогичное методологическое сомнение нами высказывалось и относительно AI$\xi^{tl}$). Хотя здесь могут использоваться и доказательства о том, что та или иная самооптимизация обязательно повысить не само будущее подкрепление, а вероятность его увеличения, ответ на вопрос о выборе надлежащей аксиоматики для машины Гёделя не дается, что делает эту машину «рамочной моделью», требующей существенного развития. В то же время, указание на необходимость самооптимизации, включающей не только уточнение $\xi(q)$, но также и программ поиска как моделей, так и оптимальных цепочек действий, представляется совершенно справедливым (стоит отметить, что в условиях бесконечных ресурсов даже модификация $\xi(q)$ с учетом имеющейся истории не нужна, поскольку предсказание выполняется по той же полной истории). К сожалению, подходящей теории самооптимизации в литературе отсутствует и требует разработки.

*Избыточность ИМИ*

Мы уже проводили предварительное обсуждение того, являются ли модели ИМИ в достаточной степени универсальными. Однако с учетом сложности достижения подобной универсальности также уместно спросить, не является ли она избыточной? Возможно, погоня за такой универсальностью бессмысленна? Хотя алгоритмическая неполнота влечет за собой слабость методов, например машинного обучения, и следует согласиться, что достижение алгоритмической полноты так или иначе необходимо, можно высказать ряд соображений, в соответствии с которыми должны существовать более общие принципы, предшествующие алгоритмической универсальности. В первую очередь, речь идет об эволюционных соображениях.

Действительно, на первый взгляд кажется, что, по крайней мере, животные не обладают универсальным интеллектом. Получается, универсальный интеллект возникает как поздняя эволюционная надстройка. Это как будто не согласуется не только с нашим желанием проводить разработку сильного ИИ путем повышения эффективности универсального интеллекта (вместо того, чтобы постепенно повышать универсальность эффективного прагматического интеллекта), но и подразумевает первичность принципов эволюционного развития.

Более того, можно сказать, что интеллект каждого человека в отдельности не столь универсален. Конечно, люди в принципе могут научиться чему угодно. В истории науки люди изобретали самые разные модели (в том числе и универсальной машины). Дети могут обучиться любому языку или освоить приемы мышления (скажем, в форме математических методов), которые еще век назад были людям неизвестны. Одна лишь способность человека к программированию практически безоговорочно доказывает алгоритмическую универсальность его мышления. Но в то же время эта универсальность весьма относительна: люди чаще демонстрируют ограниченность своего мышления, будучи неспособными освоить что-то новое или выйти за рамки привычных стереотипов. Даже в науке нередко требуется смена поколений ученых, чтобы некая новая парадигма стала в полной мере воспринята.

Конечно, это можно списать на то, что даже от универсального интеллекта в условиях ограниченных ресурсов и сравнительно небольшого времени жизни не следует ожидать слишком многого. Но при этом на себя обращает внимание тот банальный факт, что человеческое мышление – во многом социальный феномен. Мы уже отмечали необходимость учета взаимодействий между интеллектуальными агентами, но несколько в ином аспекте (скорее «соревновательном»). Здесь же речь идет о том, что универсальный интеллект, возможно, имеет смысл приписывать не отдельному индивиду, а социуму. Однако и социум «стремится» к выживанию (или «оптимизирует» какую-то неявную «целевую функцию»), а его мультиагентность – лишь одна из «архитектурных особенностей» (хотя, возможно, и очень важная), поэтому приведенные соображения не входят в противоречие с рассмотрением общих моделей универсального алгоритмического интеллекта.

Возвращаясь к эволюционным соображениям, можно также заметить, что эволюционный «поиск» идет в «алгоритмически полном пространстве» (эти термины мы берем в кавычки, чтобы подчеркнуть модельность такого представления). Действительно, в генетическом коде могут быть записаны произвольные «алгоритмы» (в частности, программы поведения животных и человека), то есть прочтение генетического кода осуществляется «универсальной машиной». В этом смысле эволюция также «алгоритмически универсальна». Более того, она оказывается «самооптимизирующимся поиском», ведь такие широко используемые в ИИ метаэвристики поиска, как скрещивание, в эволюции присутствовали не априорно, а были «изобретены». И подобного рода «изобретений» очень много. При этом на «изобретение» первых метаэвристик уходили миллиарды лет, тогда как в последствии темпы эволюции существенно возросли (особенно если считать научно-технический прогресс продолжением единого глобального эволюционного процесса). В этом смысле глобальная эволюция (начало которой, возможно, следует отнести и до момента возникновения репликаторов, также являющихся нетривиальными «машинами») исходно и представляла собой неэффективный непредвзятый, но самооптимизирующийся универсальный «интеллект», который постепенно создал эффективный прагматический интеллект – человеческий интеллект. Возможно, аналогия между эволюцией и интеллектом не слишком полная, однако свойства «алгоритмической универсальности» и «самооптимизирующегося поиска» в обоих случаях прослеживаются весьма четко. Из этого, в частности, можно сделать вывод, что непредвзятому интеллекту придется, как минимум, повторить всю эволюцию, что делает этот интеллект непрагматичным и практически нереализуемым.

Рассмотрение социума или эволюции может позволить прояснить некоторые моменты относительно целевой функции индивидуальных агентов, но при этом поставит дополнительные вопросы относительно критериев социального и эволюционного развития. В остальном же такое рассмотрение не избавит нас от необходимости решать проблему самооптимизирующегося поиска в алгоритмически полных пространствах. Можно эту проблему изучать и на примере моделирования эволюции, но в конечном итоге нам потребуется создавать прагматический ИИ. Таким образом, рассмотрение эволюционных и социальных аспектов интеллекта не дает каких-то более общих принципов его исследования и разработки, а, напротив, еще больше подчеркивает значимость рассматриваемой проблематики.

*Опасность упрощения ИМИ на примере элементарной среды*

И все же, смотря на сверхзатратные модели ИМИ, возникает естественное желание произвести их упрощение, например, уменьшить количество используемых в предсказании моделей или отделить предсказание от выбора действий. Хотя такое упрощение и необходимо, оно может быть весьма опасным даже в случае простейших сред.

Рассмотрим очень простую среду, которая на каждое действие агента $y$ выдает подкрепление $r$ с неизменной, но не известной заранее вероятностью $P(r|y)$. При этом и действия, и величины подкреплений будут принадлежать конечным множествам. Данная среда стохастична, и алгоритмическая сложность истории взаимодействия с такой средой будет пропорциональна длине самой истории. Однако в связи с известностью и элементарностью регулярной части модели среды, полное ее восстановление достаточно просто.

Рассматриваемые модели сред представляют собой двумерные массивы $P(r|y)$. Все модели будут совместимы с историей, за исключением тех, у которых $P(r|y)=0$ для таких пар $(y, r)$, которые в данной истории встречаются; а с учетом того, что вероятности являются вещественными, таких моделей будет множество меры ноль. Массив $P(r|y)$ является лишь регулярной частью модели, которая должна быть дополнена данными, необходимыми для генерации истории по распределению $P(r|y)$. Каждый элемент истории $(y, r)$ кодируется словом $c$ длиной $-\log_2 P(r|y)$ бит (здесь мы не будем обращать внимания на то, что длина кодового слова может составлять дробное число бит). Закодированная таким образом история в совокупности с описанием распределения $P(r|y)$ и является полной моделью среды, согласованной с ее историей: $q=P(r|y),\{c_{<k}\}$. Длина такой модели будет равна сумме длины описания $P(r|y)$ и длины закодированной истории. Эта модель алгоритмична и детерминирована в следующем смысле: ее можно представить как самораспаковывающийся архив – программу, внутри которой прописаны массивы $c_{<k}$ и $P(r|y)$, по которым происходит «распаковка» истории. Длина самой программы распаковки также должна была бы учитываться; но поскольку алгоритм распаковки один и тот же для всего множества рассматриваемых моделей, мы ее опускаем. Также мы положим длины описаний распределений $P(r|y)$ равными, хотя это дополнительное упрощение. Тогда длина полной модели будет равна:

$$l(q) = \sum_i c_i = -\sum_i \log_2 P(r_i \mid y_i) \text{ и } 2^{-l(q)} = \prod_i P(r_i \mid y_i).$$

Весьма примечательно, что сложность модели определяется по истории, хотя в AIξ просто перебираются модели, имеющие «безусловную» длину. Откуда берется такой эффект? Его причина достаточна прозрачна. Действуя в точности по схеме AIξ, мы должны были бы перебирать для каждого распределения $P(r|y)$ все возможные цепочки кодов истории $c_{<k}$, среди которых лишь одна оказалась бы согласованной с имеющейся реальной историей. Здесь же мы в явном виде без перебора восстанавливаем оставшуюся (случайную) часть модели по ее началу (регулярной части). То есть мы сразу выделяем из всего множества моделей подмножество таких моделей, которые согласованы с историей. Из-за этого, правда, исчезает возможность перебора моделей в порядке возрастания их длины. Такая ситуация достаточно характерна при использовании стохастических моделей, что мы более детально обсудим позднее.

В рассматриваемом элементарном случае легко в явном виде построить и модель минимальной длины, для чего нужно $P(r|y)$ оценить как частоту подкрепления $r$ при условии выполнения действия $y$.

Структура AIξ в форме (2.1.9) (или (2.1.5) с учетом того, что здесь мы используем суженное пространство моделей) указывает, что мы должны перебирать все возможные модели среды, то есть в данном случае все возможные распределения $P(r|y)$, и для каждой из них перебирать все возможные цепочки действий. Понятно, что для рассматриваемого класса элементарных сред эту схему можно сильно упростить. Однако здесь мы хотим показать, что даже для такого банального случая есть опасность выполнить неправильное упрощение. Вполне может показаться интуитивно очевидным, что при имеющемся распределении $P(r|y)$, полученном по частотной оценке, легко выбрать оптимальное действие, как действие с максимальным математическим ожиданием подкрепления. Однако это неверно. Ведь, с другой стороны, также может показаться очевидным, что постоянное совершение действий с высокой оценкой ожидаемого непосредственного подкрепления будет препятствовать совершению действий, которые еще не были опробованы или которые при первом своем применении привели к плохому результату, что не исключает возможности того, что такие действия обладают очень высоким истинным значением среднего подкрепления. В теории обучения с подкреплением это хорошо известная дилемма исследования/использования (exploration/exploitation). Но в AIξ такая дилемма не возникала. Либо она там решается сама собой (то есть действия, приводящие к высокому подкреплению или к получению новой информации, выбираются оптимальным образом), либо в AIξ что-то упущено.

Попробуем разобрать этот вопрос чуть детальнее. Если у нас есть некоторая история $y_{<k}$, $r_{<k}$, то вероятность соответствующей ей модели с

распределением $P(r|y)$ будет $\prod_{i=1}^{k-1} P(r_i \mid y_i)$ с точностью до нормирующей константы. Однако даже такая полная модель не описывает предсказания. Нас же интересуют $r_{k:m}$, а они могут быть любыми особенно с учетом произвольности выбора $y_{k:m}$. Значит, они также должны быть включены в модель (здесь видно, что использование детерминированных моделей для сред с возрастающей сложностью является несколько искусственным). Из-за этого для каждой цепочки будущих действий удобнее рассматривать все возможные отклики среды и по ним уже строить согласованные с ними модели. Для этого случая используем формулу (2.1.5):

$$y_k^* = \arg\max_{y_k} \max_{y_{>k}} \sum_{q'} \sum_{t \ge k} \mu_k(q') q'_r(y_{\le t}).$$

Здесь разные $q'$ определяются разными $P(r|y)$ – конечным множеством вещественных переменных, и разными $r_{k:m}$ – конечным множеством дискретных переменных. Суммирование по $q'$ распадается на интегрирование по всем переменным $P(r|y)$ и суммирование по всем возможным $r_{k:m}$:

$$y_k^* = \arg\max_{y_k} \max_{y_{>k}} \int_{\vec{P}} d\vec{P} \sum_{r_{k:m}} \sum_{t \ge k} \mu_k(q') q'_r(y_{\le t}).$$

При этом $q'_r(y_{\le t})=r_t$. А чему равна вероятность модели $\mu_k(q')$? Поскольку модель включает в себя коды для всей цепочки $r_{1:m}$ (так как если бы она их не включала, это была бы уже другая модель), кажется естественным считать, что

$$\mu_k(q') = \mu_m(q') = \prod_{i=1}^{m} P(r_i \mid y_i) \text{ и}$$

$$y_k^* = \arg\max_{y_k} \max_{y_{>k}} \int_{\vec{P}} d\vec{P} \sum_{r_{k:m}} \left( \left( \prod_{i=1}^{m} P(r_i \mid y_i) \right) \left( \sum_{t \ge k} r_t \right) \right).$$

Для простоты будем считать, что у нас есть всего два действия с индексами 1 и 2, а подкрепление может принимать значения только 0 и 1. Тогда регулярная компонента модели задается значениями вероятностей, которые мы обозначим как $P_{01}$, $P_{11}$, $P_{02}$, $P_{22}$, где первый индекс отвечает за значение $r$, а второй – за номер действия. При этом $P_{01} = 1 - P_{11}$ и $P_{02} = 1 - P_{22}$, то есть интеграл будет двойным. Далее можно заметить, что вероятность модели не зависит от порядка действий и подкреплений, поэтому можно различать модели не по $r_{1:m}$, а только по числу раз, когда некоторое действие сопровождалось определенным подкреплением. Пусть $n_{01}$, $n_{11}$, $n_{02}$, $n_{12}$ – количество соответствующих пар (подкрепление | действие) в истории, а $N_{01}$, $N_{11}$, $N_{02}$, $N_{12}$ – в будущем; при этом значения $N_1=N_{01}+N_{11}$ и $N_2=N_{02}+N_{12}$ определяются при переборе действий, а при переборе моделей имеются только две свободные переменные, например $N_{01}$ и $N_{02}$. Тогда

$$\mu(q') = C_{N_1}^{N_{01}} C_{N_2}^{N_{02}} P_{01}^{\,n_{01}+N_{01}} (1-P_{01})^{n_{11}+N_{11}} P_{02}^{\,n_{02}+N_{02}} (1-P_{02})^{n_{12}+N_{12}} \text{ и}$$

$$y_k^* = \arg\max_{y_k} \max_{N_1,N_2} \int_{P_{01}} \int_{P_{02}} \sum_{N_{01},N_{02}} \mu(q')(N_{11}+N_{12}) dP_{01} dP_{02}.$$

Множители $C_{N_1}^{N_{01}} C_{N_2}^{N_{02}}$ появляются в связи с тем, что суммирование по $r_{k:m}$ нельзя просто заменить суммированием по $N_{01}$ и $N_{02}$, поскольку одной модели, задаваемой через $N_{01}$ и $N_{02}$, будет соответствовать разное число моделей, задаваемых через $r_{k:m}$. Это число и выражается через количество сочетаний. Эти сомножители можно было опустить, что дало бы тоже корректную формулу, но соответствующую немного другому распределению $\mu(q')$, задающему достаточно интересное индуктивное смещение, которое, однако, могло бы вызвать лишние вопросы. Теперь можно поменять порядок интегрирования и суммирования и в явном виде вычислить интегралы вида:

$$P(n_0,n_1) = \int_0^1 P^{n_0}(1-P)^{n_1} dP = \int_0^1 \sum_{j=0}^{n_1} C_{n_1}^j (-1)^j P^{j+n_0} dP = \sum_{j=0}^{n_1} C_{n_1}^j (-1)^j \Big/ (1+j+n_0).$$

Теперь мы можем окончательно получить:

$$y_k^* = \arg\max_{y_k} \max_{N_1,N_2} \sum_{N_{01}=0}^{N_1} \sum_{N_{02}=0}^{N_2} C_{N_1}^{N_{01}} C_{N_2}^{N_{02}} P(n_{01}+N_{01}, n_{11}+N_{11}) \times$$
$$\times P(n_{02}+N_{02}, n_{12}+N_{12})(N_{11}+N_{12}).$$

Максимизируемое выражение вычисляется в явном виде. Остается только перебрать $N_1$ и $N_2$ при том, что $N_1+N_2=m-k$, то есть перебор линеен по глубине предсказания. Еще раз отметим, что для получения математического ожидания подкреплений вероятности надо нормировать.

Посмотрим, какое ожидание подкрепления получится при разных историях и при разных $N_1$, $N_2$. При $n_{01}=n_{11}=n_{02}=n_{12}=0$ получаем средний ожидаемый выигрыш 0.5 вне зависимости от того, сколько каких действий предполагается выполнить в будущем. Это кажется естественным, но некоторые сомнения могут закрасться уже здесь: хотя у нас нет априорного предпочтения какой-то среде, кажется, что возможность выбора лучших действий должна как-то смещать ожидаемое подкрепление в сторону большего суммарного подкрепления. Далее, при $n_{01}=n_{11}\neq 0$ и $n_{02}=n_{12}=0$ (одно из действий дало равное число подкреплений 0 и 1, другое не опробовано) имеем ту же картину, то есть действия оказываются равнозначными. С одной стороны, кажется, что здесь также нет предпочтения в выборе какого-то действия, поскольку текущие частоты подкреплений для них одинаковы. С другой стороны, кажется, что опробование нового ($n_{02}=n_{12}=0$) действия должно иметь больший потенциал: если действие окажется «плохим», то его можно будет дальше не повторять, и проигрыш будет незначительным; если же оно окажется «хорошим», то его можно будет повторять много раз, получая от этого выигрыш.

Каково будет ожидаемое подкрепление при $n_{01}=n_{02}$, $n_{11}=n_{12}$, но $n_{01}\neq n_{11}$ (то есть, если судить по истории, действия равнозначны, но частота подкреплений

0 и 1 разная)? Расчеты показывают, что математическое ожидание подкрепления не зависит от соотношения будущих действий $N_1$ и $N_2$, хотя оно при этом отличается от соотношения $n_{11}/(n_{01}+n_{11})$. В частности, при $n_{01}$=2, $n_{11}$=1 математическое ожидание подкрепления в расчете на одно действие оказывается 0.4, а не 1/3. Однако при увеличении длины истории различие это уменьшается. Так, при $n_{01}$=8, $n_{11}$=4 матожидание уже 0.357, а при $n_{01}$=16, $n_{11}$=8 оно составляет 0.346. На самом деле, это достаточно естественно: ведь, скажем, история с $n_{01}$=1, $n_{11}$=0 могла быть порождена с некоторой вероятностью почти любым распределением, а вовсе не только распределением с $P_{01}$=1, поэтому чистая частотная оценка является смещенной.

В остальном различий нет: если мы возьмем какие-то значения $n_{01}$, $n_{11}$, $n_{02}$, $n_{12}$, то при $N_2$=0, $N_1$=$m$–$k$ ожидаемое подкрепление будет определяться той же оценкой $P_{01}$ по $n_{01}$, $n_{11}$, а при $N_1$=0, $N_2$=$m$–$k$ – оценкой $P_{02}$ по $n_{02}$, $n_{12}$. К примеру, при $n_{01}$=2, $n_{11}$=1 и $n_{02}$=8, $n_{12}$=4 получим ожидание 0.4 при $N_2$=0 и 0.357 при $N_1$=0, которые плавно переходят друг в друга при других соотношениях $N_1$ и $N_2$. При этом стоит отметить, что при $n_{01}$=1, $n_{11}$=2 соответствующая оценка получается также не 2/3, а 0.6, то есть в этом случае она оказывается «заниженной». Иными словами, отличие ожидаемого подкрепления от простой частотной оценки всецело обусловлено смещением оценок вероятностей, а вовсе не общим увеличением подкрепления за счет учета новой информации в процессе последовательного выбора действий.

Можно подумать, что так и должно быть (то есть данные оценки корректны, и никакая стратегия выбора действий не может дать лучшего результата в силу стохастичности среды), либо же, что в самой AIξ что-то не учтено. Однако простое рассмотрение выбора первого действия показывает следующее. Пусть, к примеру, $n_{01}$=5, $n_{11}$=5, $n_{02}$=0, $n_{12}$=0 и $m$–$k$=10. У агента есть возможность выбрать одно из двух действий, и для каждого из них он с равной вероятностью получит один из двух исходов. Но будут ли при этом действия равноценными? Посмотрим, какое ожидание подкрепления будет в каждом случае:

– если выбрано первое действие:

$P = 0.5 : n_{01} = 6, n_{11} = 5, n_{02} = 0, n_{12} = 0$ => M[$r$]=0.500 при $N_2$=10;

$P = 0.5 : n_{01} = 5, n_{11} = 6, n_{02} = 0, n_{12} = 0$ => M[$r$]=0.538 при $N_1$=10;

– если выбрано второе действие:

$P = 0.5 : n_{01} = 5, n_{11} = 5, n_{02} = 1, n_{12} = 0$ => M[$r$]=0.500 при $N_1$=10;

$P = 0.5 : n_{01} = 5, n_{11} = 5, n_{02} = 0, n_{12} = 1$ => M[$r$]=0.666 при $N_2$=10.

Здесь мы просто предполагаем выбор одного из действий и одного из откликов среды и применяем к каждому из случаев рассмотренную выше модель. При этом в зависимости от выбора действия и отклика получается разная гипотетическая история, для которой модель дает разные оценки дальнейших действий.

Усредняя, получаем, что выбор первого действия дает ожидание подкрепления (с учетом подкрепления от выбираемого действия), равное 0.515, а второго – 0.575. Не только выбор второго (не совершавшегося ранее) действия приводит к большему значению ожидаемого подкрепления (то есть действия, направленные на приобретение новой информации, «сами собой» увеличивают ожидаемое подкрепление, что не требует каких-то специальных механизмов), но даже выбор первого действия дает ожидание, превышающее 0.5. И это именно то, чего следовало ожидать! Ведь математическое ожидание максимума из двух случайных величин, равномерно распределенных в диапазоне [0, 1], больше, чем 0.5. И именно этот эффект мы сейчас наблюдаем при рассмотрении выбора текущего лучшего действия.

Так почему же мы получали равные 0.5 значения ожидаемого подкрепления при выборе разных соотношений $N_1$ и $N_2$? Поскольку расчет максимального ожидаемого подкрепления не отличается по своей сути от максимизации подкрепления при выборе текущего действия (max вместо argmax), причина должна быть в том, что было упущено нечто важное. Здесь некорректный переход достаточно очевиден, поскольку он должен нарушать учет получаемой информации при максимизации последующего выбора. В частности, порядок действий здесь должен иметь значение: несмотря на то, что у нашей среды полностью отсутствует память, и в наперед выбранной совокупности действий порядок действий действительно не имеет значения, но при последовательном выборе действий их порядок играет решающее значение для обеспечения *индуктивного поведения*, то есть поведения, направленного на получение информации.

Таким образом, нельзя было отождествлять $\mu_k(q')$ и $\mu_m(q')$, хотя это не слишком прозрачно при рассмотрении детерминированных моделей, поскольку программы, содержание закодированную историю и закодированное предсказание будущего формально являются разными. Это приводит к некорректному разделению задач предсказания и выбора оптимальных действий, из-за чего в последствии и возникает возможность замены перебора по $r_{k:m}$ перебором по $N_1$ и $N_2$. На самом деле, и в AI$\xi$, заданной в форме уравнения (20) из [Hutter, 2007], не производится учета гипотетического изменения распределения $\mu$ или $\xi$ при получении новой информации. Более аккуратное воплощение модели AI$\xi$ в указанной форме также показывает, что она не реализует индуктивного поведения. В равной мере нельзя использовать ни $\mu_m(q')$, ни $\mu_k(q')$, а нужно для каждого варианта прогнозируемой истории на момент $t$ использовать собственное уточненное распределение $\mu_t(q')$.

Для получения правильного эффекта индуктивного поведения удобнее использовать рекурсивную форму AI$\xi$ (точнее, AI$\mu$), в которой шаги максимизации при выборе одного действия и суммирования по разным возможным откликам среды в явном виде чередуются. У Хаттера [Hutter, 2007] формула (22) имеет вид, аналогичный следующему уравнению:

$$y_k^* = \arg\max_{y_k} \sum_{x_k} \max_{y_{k+1}} \sum_{x_{k+1}} \ldots \max_{y_m} \sum_{x_m} P(x_{k:m} \mid x_{<k}, y_{1:m}) \sum_{t \geq k} r_t \ ,$$

где вероятности $P$ могут вычисляться через универсальное распределение ξ или какое-то конкретное распределение μ в зависимости от постановки задачи.

Стоит отметить, что разные формы AIξ в [Hutter, 2007], а, именно, (20) и (22), вряд ли могут быть эквиваленты. Последняя форма предпочтительнее, поскольку она учитывает возможные разные отклики среды в зависимости от совершаемых действий, но она также не содержит в явном виде учета уточнения распределения вероятностей. Мы предлагаем следующую форму, которая представляется наиболее корректной:

$$y_k^* = \arg\max_{y_k} \sum_{x_k} \left( P_k \left( r_k + \max_{y_{k+1}} \sum_{x_{k+1}} \left( P_{k+1} \left( r_{k+1} + \ldots + \left( \max_{y_m} \sum_{x_m} P_m r_m \right) \ldots \right) \right) \right) \right), \quad (2.2.3)$$

где $P_t = P(x_t \mid x_{<t}, y_{1:t})$. Стоит отметить, что это должны быть нормированные распределения; игнорирование нормировки (как это делается в AIξ из-за использования только $P_k$) недопустимо в силу того, что нормировка оказывается разной в зависимости от рассматриваемого варианта будущего взаимодействия агента и среды.

С учетом нормировки $P_t$ можно все $r_t$ и $P_t$ собрать внутри последней суммы и переписать формулу (2.2.3) в виде

$$y_k^* = \arg\max_{y_k} \sum_{x_k} \max_{y_{k+1}} \sum_{x_{k+1}} \ldots \max_{y_m} \sum_{x_m} P_k \cdot \ldots \cdot P_m (r_k + \ldots + r_m), \quad (2.2.4)$$

который наиболее близок к AIξ в упомянутой форме (22). Если бы все $P_t$ зависели от всех последующих действий, то из свойств условной вероятности выполнялось бы соотношение $P_k \cdot \ldots \cdot P_m = P(x_{k:m} \mid x_{<k}, y_{\leq m})$, и уравнение (2.2.4) было бы в точности эквивалентно уравнению (22) из [Hutter, 2007]. Но важная тонкость заключается в том, что $P_t = P(x_t \mid x_{<t}, y_{1:t})$ зависит только от совершенных к моменту $t$ действий, а не всех действий до момента $m$ (то есть действия выбираются не априорно, а после получения новой информации), поэтому $P_k \cdot \ldots \cdot P_m \neq P(x_{k:m} \mid x_{<k}, y_{\leq m})$ , и (2.2.3) обладает свойствами, отсутствующими в AIξ.

То, что такой агент (при правильной реализации) будет проявлять «любопытство», должно быть понятно в связи с тем, что этот эффект виден уже при анализе влияния выбора текущего действия при котором учитывается пересчет распределения вероятностей. При этом оказывается принципиальным повторный перебор всех моделей для всех возможных откликов среды на каждом такте времени при предсказании будущего подкрепления. Необдуманный отказ от столь избыточного на первый взгляд перебора приведет к нерациональному поведению (отсутствию «любопытства»). Ошибочным является не только использование для предсказания одной лучшей модели, построенной по текущей истории, но и устранение возможности учета новой информации при последовательном принятии решений и предсказании.

Можно задаться вопросом, не получится ли, наоборот, что агент всегда будет пытаться опробовать незнакомые действия? Такое поведение тоже выглядело бы нерациональным (не случайно у людей есть такие пословицы, как «от добра добра не ищут» или «лучше синица в руках, чем журавль в небе»). Для «бессмертного» агента, возможно, так и будет: если «плохое» действие всегда имеет только локальный эффект, а «хорошее» действие можно эксплуатировать неограниченно, то проверка нового будет всегда оправданной. Однако уже в случае простых Марковских сред вполне могут существовать действия, приводящие к «безвыходным положениям» (совокупности состояний среды, имеющих вероятные переходы только друг между другом, но не другими состояниями) с низким значением целевой функции. Стремление избежать таких ситуаций в условиях априорной неопределенности должно ограничивать «любопытство». Как именно агент будет себя вести в конкретных ситуациях, существенно зависит от целевой функции. Но можно надеяться, что в рамках некоторой заданной целевой функции ИМИ-агент будет оптимально разрешать дилемму исследования/использования.

Как видно, упрощение ИМИ может приводить к существенно отрицательным последствиям. Это, однако, не означает, что такое упрощение в принципе невозможно. В частности, в нашем примере индуктивное поведение могло быть достигнуто и при использовании одной модели, но представление которой было бы расширено понятием неопределенности. Так, по истории $n_{01}$, $n_{11}$, $n_{02}$, $n_{12}$ можно было бы восстанавливать одну частотную модель, но дополненную информацией о неопределенности, которая тем выше, чем меньше соответствующие значения $n_{01}+n_{11}$, $n_{02}+n_{12}$. Совершение каждого действия приводит к модификации этой модели как в плане $P(r|y)$, так и в плане неопределенностей. Для данного случая несложно разработать частный способ учета неопределенности модели, который заменил бы интегрирование по всем моделям. Однако в общем случае для алгоритмически полного пространства моделей данная задача нетривиальна, и мы к ней вернемся позднее.

*Выводы*

Модели ИМИ достаточно легко сделать вычислимыми и в определенной степени учитывающими ресурсные ограничения. Однако в лучшем случае они оказываются оптимальными по времени выполнения с точностью до мультипликативной замедляющей константы, пропорциональной $2^{-L}$, где $L$ – длина «истинной» модели среды. Помимо того, что значение $L$ может быть большим даже для отдельных частных задач, для стохастической среды оно вообще не ограничено, в связи с чем требуется модификация моделей ИМИ. Мультипликативный замедляющий фактор может быть превращен в аддитивный, который, однако, помимо того, что является еще более значительным, также действует в каждый момент времени.

Уменьшение замедляющего фактора означает самооптимизацию, имеющиеся методы которой, к сожалению, вызывают сомнения. Помимо этого непредвзятый метод универсальной самооптимизации по-прежнему будет обладать огромной аддитивной константой, характеризующей то, что ему для того, чтобы стать эффективным интеллектом, потребуется, по сути, повторить эволюцию. При этом недостаточно аккуратная запись моделей ИМИ (даже в вариантах общеизвестной модели AIξ) вполне может привести к нарушению важных форм разумного поведения.

Вместо непосредственного наивного упрощения ИМИ создание эффективного прагматического интеллекта подразумевает реализацию такого индуктивного смещения и эвристик поиска (в предсказании, выборе действий и самооптимизации), которые бы делали этот интеллект «предвзятым» – заранее адаптированным под конкретный мир (но, еще раз подчеркнем, без потери универсальности, так как в самом мире может встретиться произвольная закономерность). Очевидно, такая «предвзятость» не может быть выведена чисто аналитически и требует учета общих свойств мира, которые, в частности, должны быть отражены в особенностях естественных когнитивных архитектур.

## Литература


(Goertzel, 2010) Goertzel B. *Toward a Formal Characterization of Real-World General Intelligence* // In: E.Baum, M.Hutter, E.Kitzelmann (Eds), Advances in Intelligent Systems Research. 2010. Vol. 10 (Proc. 3rd Conf. on Artificial General Intelligence, AGI 2010, Lugano, Switzerland, March 5-8, 2010.). P. 19–24.

(Hall, 2008) Hall S. *VARIAC: an Autogenous Cognitive Architecture* // In: Frontiers in Artificial Intelligence and Applications (Proc. 1st AGI Conference). 2008. V. 171. P. 176–187.

(Hibbard, 2009) Hibbard B. *Bias and No Free Lunch in Formal Measures of Intelligence* // Journal of Artificial General Intelligence. 2009. V. 1. P. 54–61.

(Hutter, 2005) Hutter M. *Universal Artificial Intelligence. Sequential Decisions Based on Algorithmic Probability* / Springer. 2005. 278 p.

(Hutter, 2007) Hutter M. *Universal Algorithmic Intelligence: A Mathematical Top→Down Approach* // in Artificial General Intelligence. Cognitive Technologies, B. Goertzel and C. Pennachin (Eds.). Springer. 2007. P. 227–290.

(Looks, 2009) Looks M., Goertzel B. *Program Representation for General Intelligence* // B. Goertzel, P. Hitzler, M. Hutter (Eds), Advances in Intelligent Systems Research. V. 8 (Proc. 2nd Conf. on Artificial General Intelligence, AGI 2009, Arlington, Virginia, USA, March 6-9, 2009). P. 114–119.

(Pankov, 2008) Pankov S. *A Computational Approximation to the AIXI Model* // Frontiers in Artificial Intelligence and Applications (Proc. 1st AGI Conference). 2008. V. 171. P. 256–267.

(Schmidhuber, 2007a) Schmidhuber J. *The New AI: General & Sound & Relevant for Physics* // in Artificial General Intelligence. Cognitive Technologies, B. Goertzel and C. Pennachin (Eds.). Springer. 2007. P. 175–198.

(Schmidhuber, 2007b) Schmidhuber J. *Gödel Machines: Fully Self-Referential Optimal Universal Self-improvers* // In: Artificial General Intelligence. Cognitive Technologies, B. Goertzel and C. Pennachin (Eds.). Springer. 2007. P. 199–226.

(Solomonoff, 1986) Solomonoff R. *The Application of Algorithmic Probability to Problems in Artificial Intelligence* // In: L.N. Kanal and J.F. Lemmer (Eds.). Uncertainty in Artificial Intelligence: Elsevier Science Publishers. 1986. P. 473–491.

(Solomonoff, 2003) Solomonoff R. *Progress in Incremental Machine Learning* // Technical Report IDSIA-16-03, IDSIA. 2003.

(Solomonoff, 2010) Solomonoff R. *Algorithmic Probability, Heuristic Programming and AGI* // In: E.Baum, M.Hutter, E.Kitzelmann (Eds), Advances in Intelligent Systems Research. 2010. V. 10 (Proc. 3rd Conf. on Artificial General Intelligence, Lugano, Switzerland, March 5-8, 2010). P. 151–157.

(Wang, 2007) Wang P. *The Logic of Intelligence. In: Artificial General Intelligence* // In: Cognitive Technologies, B. Goertzel and C. Pennachin (Eds.). Springer, 2007. P. 31–62.

(Левин, 1973) Левин Л.А. *Универсальные задачи перебора* // Проблемы передачи информации. 1973. Т. 9. № 3. С. 115–116.