

МОДЕЛИ ОБУЧЕНИЯ В ПРОЦЕССЕ РАБОТЫ

Новиков Д.А.

(Институт проблем управления
им. В.А. Трапезникова РАН, Москва)
novikov@ipu.ru

Рассмотрены модели итеративного научения, осуществляемого в процессе деятельности. Сформулирована и решена задача об оптимальном обучении коллектива агентов – выборе выполняемых ими в те или иные промежутки времени объемов работ.

Ключевые слова: обучение, коллектив, оптимизационная модель.

1. Введение

Члены коллектива (*агенты*) в процессе совместной деятельности сознательно или неосознаваемо приобретают опыт как индивидуальной, так и совместной деятельности, то есть имеет место их *научение* (под научением понимается «процесс и результат приобретения индивидуального опыта» [3]). Научение является частным случаем *обучения* – процесса овладения знаниями, умениями, навыками [1]. Рассмотрим последовательно ряд моделей, отражающих эффекты научения членов коллектива в процессе их работы. Начнем с общей постановки задачи и количественного описания процесса научения, затем рассмотрим модель индивидуального обучения и, наконец, модель обучения коллектива агентов.

2. Общая постановка задачи и модель процесса научения

Качественно в общем виде задача об оптимальном обучении может быть сформулирована следующим образом. Каждый агент, входящий в коллектив, характеризуется некоторым первоначальным уровнем навыка (например, производительностью

труда). В процессе осуществления деятельности производительность труда агента растет по мере приобретения опыта, совершенствования навыков и т.д. (имеет место обучение в процессе работы – learning by doing), причем скорость этого роста (так называемая *скорость научения* – см. формальное определение ниже) у каждого агента индивидуальна. Спрашивается, как оптимальным образом распределить во времени работу между агентами. Ведь один агент, уровень начальной квалификации которого низок, будучи с самого начала сильно загружен, быстро повысит свою квалификацию и сможет потом работать эффективно. С другой стороны, быть может, рациональным является загрузить сразу тех агентов, которые имеют более высокую начальную квалификацию? Ответы на эти вопросы не очевидны, тем более, что необходимо доопределить, что понимается под «оптимальным» распределением работы между агентами. В качестве критерия эффективности могут выступать суммарные затраты агентов, время выполнения коллективом заданного объема работ, результат, достигнутый за фиксированное время, и т.д.

Перейдем к формализации рассмотренной ситуации – начав с наиболее простой модели, будем потом постепенно ее усложнять. При этом ограничимся случаем *итеративного научения* [6], соответствующего достаточно рутинным видам деятельности и представляющего собой многократное повторение обучаемой системой действий, проб, попыток и т.д. для достижения фиксированной цели при постоянных внешних условиях. Итеративное научение (ИН) лежит в основе формирования навыков у человека, условных рефлексов у животных, обучения многих технических (материализованных) и кибернетических (абстрактно-логических) систем и является предметом исследования педагогической и инженерной психологии, психофизиологии, педагогики, теории управления и других наук (см. обзор в [6]).

Постоянство как внешних условий, так и цели, позволяет проводить количественное описание ИН в виде *кривых научения* – зависимостей критерия уровня научения от времени или от числа повторений (итераций).

Многочисленные экспериментальные данные (см. [5, 6]) свидетельствуют, что важнейшей общей закономерностью

итеративного научения является замедленно-асимптотический характер кривых научения: они монотонны, скорость изменения критерия уровня научения со временем уменьшается, а сама кривая асимптотически стремится к некоторому пределу. В большинстве случаев кривые итеративного научения аппроксимируются экспоненциальными кривыми.

Различают два аспекта научения. Первый аспект – результативный: при научении система должна достичь требуемого результата – качества выполнения действий – с приемлемыми затратами времени, энергии и т.д. Второй аспект – процессуальный: адаптация, приспособление научаемой системы к некоторому виду действий в процессе деятельности (например, упражнения) и т.д. Соответственно, выделяют результативные характеристики итеративного научения и характеристики адаптации [5], которые относятся, как правило, к физиологическим компонентам деятельности – утомляемость и т.п. В настоящей работе речь идет именно о результативных характеристиках научения (характеристики адаптации зачастую имеют совсем другую динамику).

Как отмечалось выше, итеративное научение, как правило, характеризуется замедленно-асимптотическими кривыми научения, аппроксимируемыми экспоненциальными кривыми (см. рис. 1):

$$(1) \quad r(t) = r^\infty + (r^0 - r^\infty) e^{-\gamma t}, \quad t \geq 0,$$

или дискретной последовательностью¹

$$(2) \quad r^k = r^\infty + (r^0 - r^\infty) e^{-\gamma k}, \quad k = 1, 2, \dots,$$

где t – время научения, k – число итераций (проб, попыток) с момента начала научения, $r(t)$ (r_k) – тип агента (уровень навыка, квалификация) в момент времени t (на k -ой итерации), $r^0 > 0$ – начальная квалификация (соответствующее моменту начала научения – первому периоду времени – значение типа), r^∞ – «конечное» значение, $r^\infty \geq r^0$, γ – некоторая неотрицательная константа, определяющая скорость изменения типа и называемая скоростью научения [6].

¹ Условимся верхним индексом обозначать номер периода времени, а нижним индексом – номер агента. В случае, когда рассматривается единственный агент, нижний индекс будем опускать.

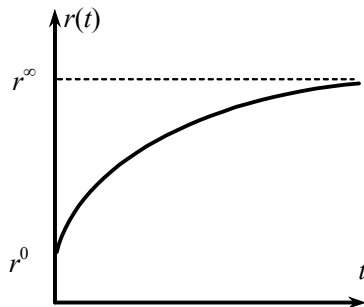


Рис. 1. Экспоненциальная кривая обучения

3. Обучение одного агента

Рассмотрим сначала *модель обучения* (обучения – так как мы рассматриваем только процесс обучения, то различий между этими терминами делать не будем) единственного агента. Обозначим $y^k \geq 0$ – выполняемый им в k -ом периоде времени *объем работ*. Если интерпретировать тип агента (уровень навыка) $r^k \in [0; 1]$ как долю успешных действий агента, то, выполняя в периоде k объем работ y^k , агент достигнет *результата* $z^k = r^k y^k$.

Тогда результат агента – суммарный объем работ, успешно выполненных агентом за k периодов времени, равен

$$(2) \quad Z^k = \sum_{l=1}^k r^l y^l .$$

С другой стороны, агентом выполнен большой объем (успешных и неуспешных) работ:

$$(3) \quad Y^k = \sum_{l=1}^k y^l .$$

Этот объем работ условно можно считать тем «*опытом*», который приобрел агент (см. [2, 7], а также обзор в [6]), то есть, его «эффективным внутренним временем» (прошедшим с момента начала обучения и потраченным на обучение), и подставить в показатель экспоненты (1). Получим:

$$(4) \quad r^k = 1 - (1 - r^0) \exp(-\gamma Y^{k-1}), \quad k = 2, 3, \dots .$$

Обозначим $y^{1:\tau} = (y^1, y^2, \dots, y^\tau)$, $\tau = 1, 2, \dots$ и условимся считать, что $y^0 = 0$.

Объединяя (3) и (4), получим следующие выражения для соответственно объемов успешно выполненных работ и типов агента:

$$(5) \quad Z^k = \sum_{l=1}^k y^l \left\{ 1 - (1 - r^0) \exp\left(-\gamma \sum_{m=1}^{l-1} y^m\right) \right\},$$

$$(6) \quad r^k = 1 - (1 - r^0) \exp\left(-\gamma \sum_{l=1}^{k-1} y^l\right), \quad k = 2, 3, \dots$$

Отметим, что при фиксированном суммарном объеме работ тип агента определяется выражением (4) однозначно и не зависит от того, как объемы работ распределены по периодам времени. Поэтому задача максимизации типа агента при фиксированном суммарном объеме работ в рамках рассматриваемой модели не имеет смысла.

В модели фигурируют три «макропараметра»: суммарный объем работ Y , число периодов T и результат Z . Искомой переменной является «траектория обучения» $y^{1:T}$.

Задачи оптимального обучения могут заключаться в экстремизации одной из переменных при фиксированных остальных переменных¹. Таким образом, получаем, что целесообразно рассматривать следующие постановки.

1. Фиксируем суммарный объем работ Y , который может выполнить агент, и результат Z , который необходимо достичь. Требуется найти траекторию, минимизирующую время достижения результата:

$$(7) \quad \begin{cases} T \rightarrow \min \\ Y^T \leq Y, \\ Z^T \geq Z. \end{cases}$$

¹ В более общем случае может существовать некоторый функционал, который подлежит экстремизации (например, затраты на обучение, качество обучения и т.д.), могут приниматься во внимание дополнительные ограничения, варьироваться могут одновременно несколько переменных и т.д. – все эти задачи представляют перспективный предмет будущих исследований.

Задачу (7) можно условно назвать задачей о быстродействии.

2. Фиксируем суммарный объем работ Y , которые может выполнить агент, и время обучения T . Требуется найти траекторию, максимизирующую результат Z :

$$(8) \quad \begin{cases} Z^\tau \rightarrow \max \\ Y^\tau \leq Y, \\ \tau \leq T. \end{cases}$$

Задачу (8) можно условно назвать *задачей об оптимальном обучении агента*. Наверное, именно эта задача наиболее близка к проблемам педагогики, когда требуется за фиксированное время и при фиксированном объеме учебного материала так распределить его во времени (дидактические аспекты, то есть содержание, в силу рутинности предмета научения практически не имеют значения), чтобы максимизировать «объем усвоенного материала» (максимизировать «качество обучения»).

Так как выражение (5) монотонно по сумме объемов работ агента и длительности периода обучения, то задачу (8) можно записать в виде:

$$(9) \quad \sum_{l=1}^T y^l \exp(-\gamma \sum_{m=1}^{l-1} y^m) \rightarrow \min_{\{y^{1,T} | \sum_{\tau=1}^T y^\tau = Y\}} .$$

В выражение (9) уже не входит начальная квалификация агента r^0 , то есть справедливо следующее утверждение.

Утверждение 1. Решение задачи об оптимальном обучении не зависит от начальной квалификации агента.

Данный вывод представляет интерес для методики обучения, так как с точки зрения результатов отдельных независимых агентов существенны только индивидуальные различия скоростей их научения.

3. Фиксируем время обучения T и результат Z , который требуется достичь. Требуется найти траекторию обучения, минимизирующую суммарный объем работ:

$$(10) \quad \begin{cases} Y^\tau \rightarrow \min \\ \tau \leq T \\ Z^T \geq Z \end{cases} .$$

Каждая из задач (7)-(10) может быть сведена к задаче (или набору задач) динамического программирования.

Пример 1. Решим задачу (9) для случая двух периодов. Если $T = 2$, то имеем два варианта. В первом $\tau = 1$, то есть $y^1 = Y$, тогда $Z^1 = r^0 Y$. Во втором $\tau = 2$, и, учитывая, что $y^2 = Y - y^1$, получим:

$$Z^2(y^1) = Y - (1 - r^0) [y^1 + (Y - y^1) \exp(-\gamma y^1)].$$

Легко видеть, что максимум этого выражения по $y^1 \in [0; Y]$ не зависит от r^0 . То есть оптимальное распределение объемов работ по периодам не зависит от начальной квалификации агента.

Кроме того, чем больше γ , тем меньше оптимальное значение объема работ, выполняемого в первом периоде. То есть чем выше скорость научения агента, тем больший объем работ он должен выполнять в последнем периоде (и, соответственно, тем меньший объем работ необходимо выделять на начальный период для повышения его начальной квалификации). •¹

Пример 2. Решим задачу (9) для случая $T = 11$, $r^0 = 0,1$, $\gamma = 0,5$, $Y = 10$.

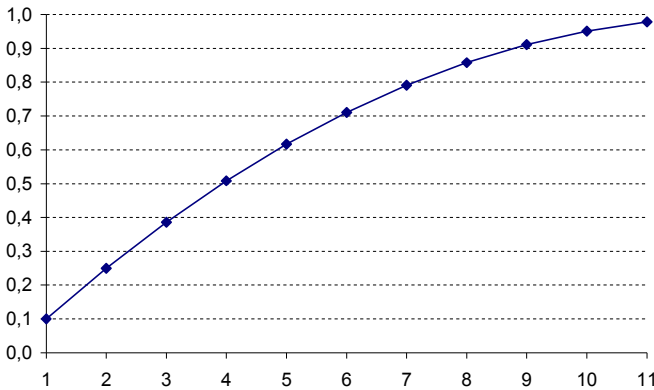


Рис. 2. Динамика типов агента в примере 2

Динамика типов агента представлена на рис. 2. Динамика оптимальных объемов работ представлена на рис. 3.

¹ Символ «•» здесь и далее обозначает окончание примера.

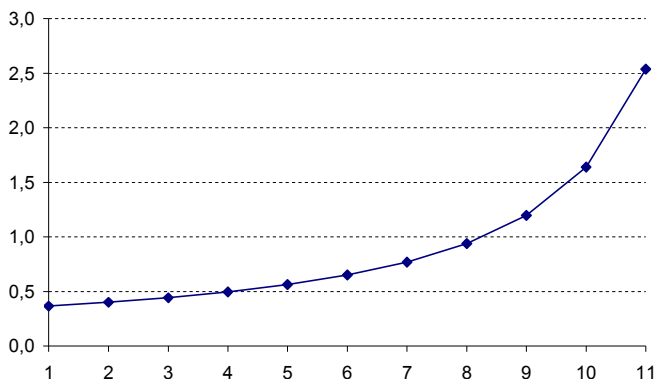


Рис. 3. Динамика оптимальных объемов работ в примере 2

Оптимальной стратегией обучения является увеличение объема работ агента со временем, причем, чем выше скорость обучения, тем более «выпуклой» является оптимальная траектория обучения. •

4. Обучение нескольких агентов

До сих пор мы рассматривали одного агента. Обобщим полученные результаты на случай нескольких одновременно работающих агентов, причем сначала рассмотрим ситуацию, когда агенты полностью независимы (результаты и тип каждого не зависят от результатов и типов других), а потом проанализируем задачу об обучении зависимых агентов.

Рассмотрим коллектив – множество $N = \{1, 2, \dots, n\}$, состоящее из n агентов. По аналогии с выражениями (5) и (6), получим следующие выражения для соответственно объемов успешно выполненных работ и типов агентов:

$$(11) Z_i^k = \sum_{l=1}^k y_i^l \left\{ 1 - (1 - r_i^0) \exp\left(-\gamma_i \sum_{m=1}^{l-1} y_i^m\right) \right\},$$

$$(12) r_i^k = 1 - (1 - r_i^0) \exp\left(-\gamma_i \sum_{l=1}^{k-1} y_i^l\right), k = 2, 3, \dots, i \in N.$$

Если результат коллектива является суммой результатов входящих в этот коллектив агентов, то есть

$$(13) Z^k = \sum_{i=1}^n Z_i^k, k = 1, 2, \dots,$$

то задача об оптимальном обучении коллектива (ср. с (8)) примет вид:

$$(14) Z^T \rightarrow \max_{\{y_i^{l,T} | \sum_{\tau=1}^T \sum_{i=1}^N y_i^\tau = Y\}},$$

то есть:

$$(15) \sum_{i=1}^n \sum_{l=1}^T y_i^l \{1 - (1 - r_i^0) \exp(-\gamma_i \sum_{m=1}^{l-1} y_i^m)\} \rightarrow \max_{\{y_i^{l,T} | \sum_{\tau=1}^T \sum_{i=1}^N y_i^\tau = Y\}}.$$

Задача (15) может быть решена методом динамического программирования.

Легко видеть, что оптимальное решение задачи (15) в общем случае зависит и от индивидуальных скоростей научения агентов $\{\gamma_i\}$, и от их начальных квалификаций $\{r_i^0\}$.

Утверждение 2. Если скорости научения агентов одинаковы, то оптимальным распределением работ является выполнение всего объема работ агентом с максимальной начальной квалификацией. Если начальные квалификации агентов одинаковы, то оптимальным распределением работ является выполнение всего объема работ агентом с максимальной скоростью научения.

Пример 3. Рассмотрим задачу (15) для случая двух агентов при $T = 11$, $r_1^0 = 0,1$, $r_2^0 = 0,3$, $\gamma_1 = \gamma_2 = 0,75$, $Y = 10$. При этом в соответствии с утверждением 2 оптимальным является выполнение всего объема работ вторым агентом, то есть тем, чья начальная квалификация выше (напомним, что в настоящем примере скорости научения агентов одинаковы). На рис. 4 приведена динамика оптимальных объемов работ (квадратики соответствуют второму агенту, ромбики – первому).

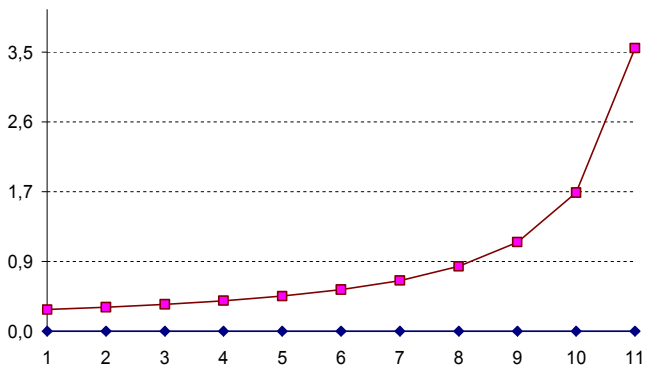


Рис. 4. Динамика оптимальных объемов работ в примере 3

Получили, что первый агент не выполняет никаких работ и не обучается. •

Решение задачи об оптимальном обучении в случае, когда все агенты имеют одинаковые скорости научения, получилось «вырожденным» – работает и обучается один агент, а остальные не работают и не обучаются. С одной стороны, такой коллектив вряд ли можно назвать полноценным, с другой стороны, необходимо признать, что в жизни такие ситуации встречаются нередко.

Рассмотрим, что произойдет, если агенты различаются и начальными квалификациями, и скоростями научения.

Пример 4. Если в условиях примера 3 (при неизменных всех остальных параметрах) выбрать скорость научения первого агента (чья начальная квалификация ниже, чем у второго агента) равной 3,0, то есть сделать ее существенно больше скорости научения второго агента, то оптимальным решением будет выполнение всего объема работ уже не вторым, а первым агентом. •

Формально, структура решения задачи (15) – то, что весь объем работ выполняет «лучший» (с точки зрения комбинации начальной квалификации и скорости научения) агент – обусловлена наличием большого числа переменных при единственном ограничении. Содержательно, в задаче могут присутствовать и

другие ограничения, помимо ограничения на суммарный объем работ, выполняемый членами коллектива. Наиболее естественным представляется ограничение на максимальный объем работ, который каждый агент может выполнить за одну итерацию (за один период времени).

Пример 5. Если в условиях примера 3 добавить ограничение на максимальный объем работ (равный, например, 0,5), который каждый агент может выполнить за один период времени, то в оптимальном решении будут загружены уже оба агента – динамика их типов представлена на рис. 5, а динамика оптимальных объемов работ – на рис. 6.

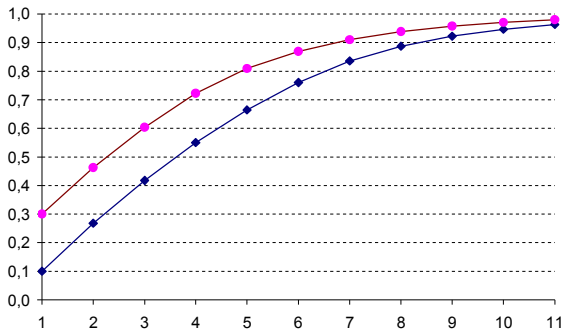


Рис. 5. Динамика типов агентов в примере 5

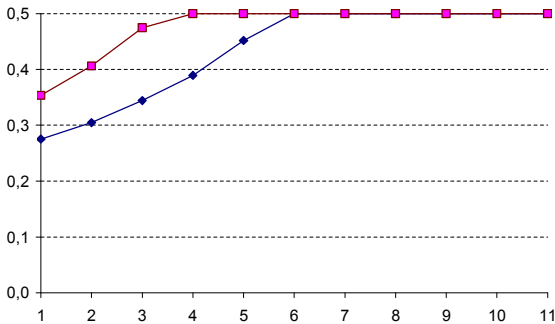


Рис. 6. Динамика оптимальных объемов работ в примере 5

5. Обучение в коллективе

До сих пор при рассмотрении научения агентов в процессе работы мы считали, что каждый агент учится только «на собственном опыте». Тем не менее, в коллективах имеет место обмен опытом, и агенты, наблюдая за деятельностью других (их успехами и трудностями), могут также приобретать опыт. Для того, чтобы отразить этот эффект, будем описывать «опыт», накопленный агентом, не только как сумму его собственных действий, но и добавим к этой сумме взвешенную сумму действий других агентов. В результате получим следующие выражения для соответственно объемов успешно выполненных работ и типов агентов:

$$(16) Z_i^k = \sum_{l=1}^k y_i^l \left\{ 1 - (1 - r_i^0) \exp\left(-\gamma_i \sum_{j=1}^n \alpha_{ij} \sum_{m=1}^{l-1} y_j^m\right) \right\},$$

$$(17) r_i^k = 1 - (1 - r_i^0) \exp\left(-\gamma_i \sum_{j=1}^n \alpha_{ij} \sum_{l=1}^{k-1} y_j^l\right), k = 2, 3, \dots, i \in N,$$

где константы $\{\alpha_{ij} \geq 0\}$ могут интерпретироваться как *эффективности передачи опыта* от j -го агента i -му, $i, j \in N$.

Тогда задача об оптимальном обучении примет вид:

$$(18) \sum_{i=1}^n \sum_{l=1}^T y_i^l \left\{ 1 - (1 - r_i^0) \exp\left(-\gamma_i \sum_{j=1}^n \alpha_{ij} \sum_{m=1}^{l-1} y_j^m\right) \right\} \rightarrow \max_{\{y_i^{1,T} | \sum_{i=1}^n \sum_{l=1}^T y_i^l = Y\}} .$$

Пример 6. Рассмотрим задачу (18) в условиях примера 3 (скорости научения обоих агентов одинаковы, второй агент обладает большей начальной квалификацией) при матрице $\|\alpha_{ij}\| = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$. Качественно, первый агент обучается на своем

опыте и на опыте второго агента (даже более эффективно, чем на своем). Второй же агент обучается только на своем собственном опыте. Динамика типов агентов представлена на рис. 7, а динамика оптимальных объемов работ – на рис. 8.

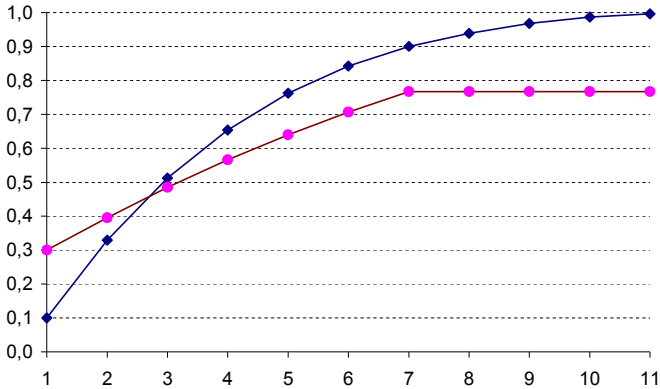


Рис. 7. Динамика типов агентов в примере 6

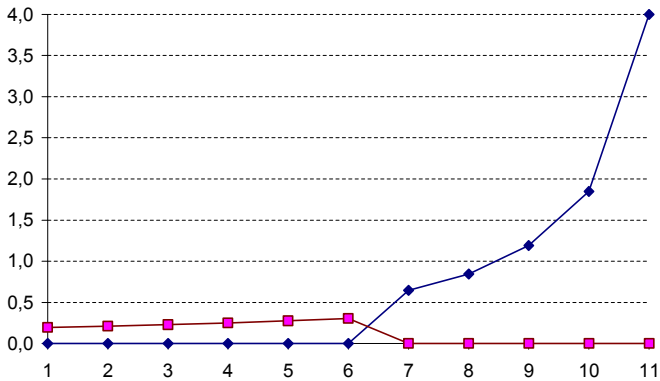


Рис. 8. Динамика оптимальных объемов работ в примере 6

Первые шесть периодов первый агент не выполняет работ сам, а «наблюдает» за действиями второго агента. При этом квалификация первого агента растет гораздо быстрее, чем второго. Начиная с седьмого периода, оптимальным оказывается выполнение всего объема работ первым, а не вторым агентом.

Данный пример наглядно иллюстрирует, как недостаток начальной квалификации может быть успешно компенсирован

эффективным обучением на чужом опыте. Возможна и другая (близкая) интерпретация. Можно считать второго агента учителем, тьютором, наставником, который, имея более высокую начальную квалификацию, обучает первого агента. В какой-то момент ученик «обгоняет» учителя и может работать самостоятельно. •

6. Логистические кривые научения

До сих пор мы предполагали, что уровень навыка каждого агента описывается экспоненциальной зависимостью (1), то есть рассматривали достаточно рутинную деятельность. Рассмотрим несколько более сложный случай.

Так как итеративное научение является одним из частных случаев научения, то, помимо экспоненциальных кривых, соответствующих итеративному научению, встречаются кривые научения других типов, в том числе – *логистические кривые научения*, которые аппроксимируются зависимостью (см. рис. 9):

$$(19) r(t) = r^0 r^\infty / (r^0 + (r^\infty - r^0) e^{-\gamma t}).$$

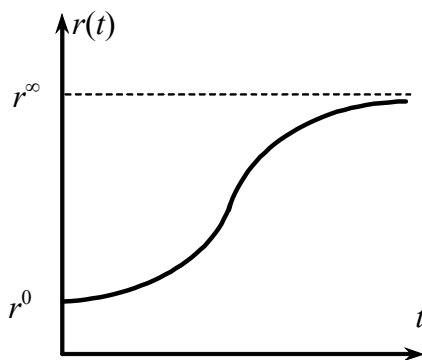


Рис. 9. Логистическая кривая научения

При этом скорость изменения $r(t)$ первоначально мала (некоторое время может требоваться на понимание задачи, идентификацию и осознание ситуации и т.п., то есть на первоначальную *адаптацию*), затем в окрестности точки перегиба скорость

увеличивается (система интенсивно обучается), а потом начинает уменьшаться. Другими словами, имеют место два параллельных процесса – адаптация и обучение. Сначала «все силы» идут на адаптацию, а затем – на обучение.

Пример 7. Решим задачу (9) для случая $T = 10$, $r^0 = 0,1$, $\gamma = 0,75$, $Y = 10$ при условии, что динамика типа агента описывается логистической кривой (19).

Динамика типов агента представлена на рис. 10. Динамика оптимальных объемов работ представлена на рис. 11.

Оптимальная стратегия обучения уже не столь тривиальна, как в примере 2 – сначала объем работ, выполняемых агентом, уменьшается, а затем начинает расти. Интересно отметить, что в точке перегиба (когда скорость обучения – производная кривой научения – максимальна) оптимальный объем работ минимален. Качественно, получается, что в оптимальном решении либо агент учится (адаптируется), либо работает на достижение результата, потому что эффективность каждого из этих процессов не может быть максимальна одновременно. •

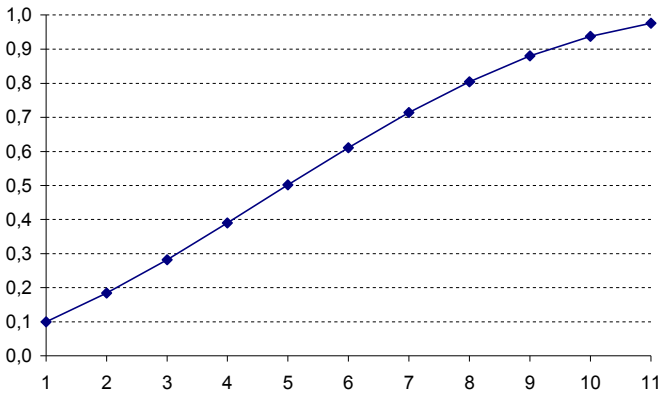


Рис. 10. Динамика типов агента в примере 7

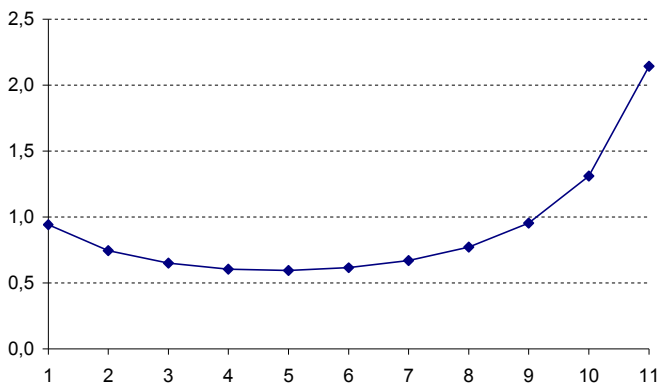


Рис. 11. Динамика оптимальных объемов работ в примере 7

Отметим, что результат утверждения 1 для коллектива, обучение членов которого описывается логистической кривой, не имеет места, то есть оптимальная траектория обучения уже зависит от начальной квалификации агента.

Задачи оптимального распределения работ между членами коллектива, научение которых описывается логистическим законом (19), формулируются аналогично соответствующим рассмотренным выше для экспоненциальных кривых научения задачам.

В заключение настоящего раздела отметим, что существуют и более сложные (чем (1) и (19)) кривые научения – так называемые последовательные логистические кривые [6], соответствующие освоению различных смежных или все более сложных видов деятельности; обобщенные логистические кривые [4] и др. Их подробное рассмотрение выходит за рамки ограниченного объема настоящей работы, хотя, если известны законы научения членов коллектива (пусть даже эти законы довольно сложны), то задача оптимального распределения объемов работ может ставиться так, как это делалось выше. А вот поиск общего решения (желательно – аналитического) этой задачи является предметом будущих исследований.

7. Заключение

Таким образом, в настоящей работе рассмотрены модели обучения в процессе работы. В рамках предположения о том, что объем уже выполненных агентом работ условно отражает накопленный им «опыт», сформулирована и решена задача об оптимальном обучении – выборе объемов работ, выполняемых агентами в те или иные промежутки времени. Проведенный анализ свидетельствует, что моделирование позволяет сделать следующие выводы:

- при фиксированном суммарном объеме работ одного агента результативные характеристики научения не зависят от того, как объемы работ распределены по периодам времени;

- решение задачи об оптимальном итеративном научении одного агента не зависит от его начальной квалификации;

- чем выше скорость научения агента, тем больший объем работ он должен выполнять в последних периодах (и, соответственно, тем меньший объем работ необходимо выделять на начальные периоды для повышения его начальной квалификации);

- оптимальной стратегией обучения является увеличение объема работ агента со временем, причем, чем выше скорость обучения, тем более «выпуклой» является оптимальная траектория обучения;

- если отсутствуют ограничения на индивидуальные объемы работ, то в коллективе весь объем работ выполняет «лучший» (с точки зрения комбинации начальной квалификации и скорости научения) агент;

- недостаток начальной квалификации агента может быть успешно компенсирован эффективным обучением как на его собственном, так и чужом опыте.

Литература

1. *Большой энциклопедический словарь*. М.: Большая российская энциклопедия, 2002. – 827 с.
2. БУШ Р., МОСТЕЛЛЕР Ф. *Стохастические модели обучаемости*. М.: Гос. изд-во физ.-мат. лит., 1962.
3. *Краткий психологический словарь*. М.: ИПЛ, 1985. – 201 с.
4. НИЖЕГОРОДЦЕВ Р. М. *Информационная экономика*. В 3-х кн. – М.: МГУ, 2002.
5. НОВИКОВ А. М. *Процесс и методы формирования трудовых умений: пропедagogика*. М.: Высшая школа, 1986.
6. НОВИКОВ Д. А. *Закономерности итеративного научения*. М.: ИПУ РАН, 1998.
7. HULL C. L. *Principles of behavior and introduction to behavior theory*. New York: Appleton Century Company, 1943.

*Статья представлена к публикации
членом редакционной коллегии Р.М. Нижегородцевым*