

Модель формирования простого вида импринтинга в процессе обучения и эволюции автономных агентов

В.Г. Редько¹, О.В. Редько

1 – доктор физико-математических наук

Построена и исследована компьютерная модель происхождения простой формы импринтинга в процессе обучения и эволюции автономных агентов. Импринтинг – известное в биологии явление запечатления в памяти животных определенных объектов. В настоящей работе моделируется формирование простого вида импринтинга – запечатления образа матери только что родившимися птенцами и последующего использования этого образа. Продемонстрировало формирование этого вида импринтинга и проанализированы механизмы взаимодействия между обучением и эволюционной оптимизацией автономных агентов в процессе формирования импринтинга. Модель строится на базе предыдущей формальной модели генерации простых навыков автономных агентов в процессе их обучения и эволюции.

1. Введение

В работе [1] была построена и исследована формальная модель взаимодействия между обучением и эволюцией автономных агентов. В модели предполагалось, что имеется популяция агентов, каждый агент имеет генотип и фенотип. Генотип S_{Gk} и фенотип S_{Pk} представляют собой одинаковые по форме цепочки бинарных символов (k – номер агента в популяции). И имеется оптимальная цепочка S_m , которая ищется в процессе обучения и эволюционной оптимизации агентов. Процесс обучения осуществляется путем проб и ошибок, эволюция состоит в отборе и мутациях агентов. Генотип агента S_{Gk} не меняется в процессе его жизни, а фенотип S_{Pk} совершенствуется в процессе обучения. Потомки агента наследуют генотип своего родителя, отбор осуществляется в соответствии с конечным фенотипом агента, полученным в результате обучения. Начальный фенотип вновь рожденного агента определяется его генотипом, полученным от родителя. В [1] было показано, что обучение может радикально ускорять процесс эволюционной оптимизации. Кроме этого, были проанализированы механизмы взаимодействия между обучением и эволюционной оптимизацией: механизм генетической ассимиляции, при которой индивидуально приобретаемые путем обучения навыки «переизобретаются» эволюцией и записываются непосредственно в генотип агентов, механизм эффекта экранирования, приводящего к тому, что сильное обучение тормозит эволюционную оптимизацию генотипов и механизм влияния нагрузки на обучение, которая приводит к значительному ускорению эволюции и генетической ассимиляции.

Отметим, что хотя генетическая ассимиляция и эффекты экранирования и нагрузки на обучение были ранее известны и в определенной степени изучены путем компьютерного моделирования [2,3], однако механизмы этих эффектов не были четко раскрыты. Эти механизмы в [1] удалось количественно проанализировать, благодаря использованию одной из наиболее четких моделей эволюции – модели квазивидов [4,5] и количественных оценок эффективности эволюционных

алгоритмов, полученных в [6-8].

Настоящая статья посвящена развитию работы [1], в ней построена и исследуется связанная с биологическим прототипом компьютерная модель формирования простого вида импринтинга в процессе обучения и эволюции популяции автономных агентов, аналогов цыплят, которым целесообразно следовать за своей матерью, курицей. Анализируется, каким образом может происходить формирование навыка следования за матерью в процессе обучения и эволюции. Также анализируется, как в этой модели работают механизмы взаимодействия между обучением и эволюционной оптимизацией автономных агентов.

В разделе 2 излагаются основные результаты работы [1]. В разделе 3 построена и исследована компьютерная модель формирования импринтинга в процессах обучения и эволюции.

2. Формальная модель взаимодействия между обучением и эволюцией

2.1. Описание формальной модели

Рассматривается эволюционирующая популяция автономных агентов или модельных особей. Генотип S_{Gk} и фенотип S_{Pk} агента определяются единообразно, а именно, задаются цепочками символов, равных 0 либо 1. Длина этих цепочек N одинакова. Например, как и в [2], можно считать, что генотип кодируется модельной цепочкой ДНК, «буквы» которой равны 0 либо 1, а фенотип определяется нейронной сетью агента, при этом веса синапсов нейронной сети также равны 0 либо 1. Начальные веса синапсов, полученные при рождении, равны генотипу, а в процессе жизни агента эти веса меняются путем обучения.

Популяция состоит из n агентов, генотипы которых равны S_{Gk} , $k = 1, \dots, n$. S_{Gk} представляет собой цепочку символов, S_{Gki} , $i = 1, \dots, N$. Длина цепочки N и численность популяции n велики: $N, n \gg 1$, величины N и n не меняются в ходе эволюции. Символы S_{Gki} рав-

ны 0 либо 1. Предполагается, что N настолько велико, что только малая часть из всех возможных 2^N генотипов может быть представлена в отдельной популяции: $2^N \gg n$. Эволюционный процесс состоит из последовательности поколений. Новое поколение получается из старого путем отбора и мутаций агентов. В нулевом поколении генотипы агентов случайны.

При размножении потомки наследуют генотипы своих родителей S_{Gk} ; генотип агента не меняется в течение его жизни и передается (с малыми мутациями) его потомкам. Мутации состоят в случайной замене символов S_{Gki} . В момент рождения агент получает от своего родителя генотип S_{Gk} , а фенотип S_{Pk} в этот момент времени t равен генотипу: $S_{Pk}(t=1) = S_{Gk}$ (время t дискретно). После рождения агенты живут в течение T_g тактов времени, T_g – время жизни одного поколения. Цепочка S_{Pk} меняется в течение жизни агента путем обучения.

Предполагается, что имеется оптимальная цепочка S_m длины N (компоненты которой также равны 0 или 1), которая ищется в процессе эволюции и обучения агентов. При моделировании цепочка S_m задавалась случайным образом.

Обучение происходит методом проб и ошибок. При обучении меняются символы фенотипов S_{Pk} : каждый такт времени t каждый символ цепочки S_{Pk} с определенной вероятностью p_i меняется на случайный (равный 0 либо 1), и если новый символ совпадает с соответствующим символом оптимальной цепочки S_m , то этот символ сохраняется; в противном случае происходит возврат к старому символу цепочки S_{Pk} . Таким образом, при обучении происходит приближение фенотипов S_{Pk} к оптимальной цепочке S_m .

Отметим, что при интерпретации символов цепочек S_{Pk} как весов синапсов нейронной сети процесс обучения имеет простой смысл: обучение происходит путем поиска оптимальных весов синапсов.

В конце поколения происходит отбор агентов в соответствии с их приспособленностями, определяемыми фенотипами в такт времени $t = T_g$. Обозначим этот конечный фенотип S_{Fk} , т.е. полагаем $S_{Fk} = S_{Pk}(t = T_g)$. Приспособленность произвольного агента S_k определяется расстоянием по Хеммингу $\rho(S_{Fk}, S_m)$ между цепочками S_{Fk} и S_m :

$$f(S_k) = \exp[-\beta\rho(S_{Fk}, S_m)] + \varepsilon, \quad (1)$$

где β – положительный параметр, характеризующий интенсивность отбора, $0 < \varepsilon \ll 1$. Параметр ε характеризует то, что нет сильного различия в приспособленностях агентов, имеющих цепочки S_{Fk} , удаленные на большое расстояние ρ от оптимальной. Роль слагаемого ε можно рассматривать как влияние случайных факторов окружающей среды на приспособленность агентов.

Отбор агентов в новое поколение производится известным рулеточным методом, при котором агенты отбираются в новое поколение вероятностным способом. Выбор агентов в следующее поколение производится n раз, в результате численность популяции во всех поколениях равна n . При каждом выборе вероятность селекции определенного агента пропорциональна его приспособленности. Подробнее о рулеточном методе отбора см., например, [6].

Таким образом, обучаемые агенты отбираются по

конечным фенотипам $S_{Fk} = S_{Pk}(t = T_g)$, т.е. по конечному результату обучения, а по наследству передаются потомкам (с малыми мутациями) генотипы S_{Gk} .

Поскольку потомкам передаются именно генотипы S_{Gk} , полученные от родителей, а не конечные фенотипы S_{Fk} , полученные в результате обучения, то эволюционный процесс имеет дарвиновский характер.

Дополнительно учтено то, что процесс обучения имеет определенную нагрузку на агента и приспособленность агента может уменьшаться под влиянием этой нагрузки. С этой целью рассматриваются модифицированные следующим образом приспособленности агентов:

$$f_m(S_k) = \exp(-\alpha d) \{ \exp[-\beta\rho(S_{Fk}, S_m)] + \varepsilon \}, \quad (2)$$

где α – положительный параметр, учитывающий нагрузку, связанную с обучением, $d = \rho(S_{Gk}, S_{Fk})$ – расстояние по Хеммингу между начальной и конечной цепочками S_{Pk} рассматриваемого агента, т.е. величина, характеризующая интенсивность всего процесса обучения агента ($S_{Gk} = S_{Pk}(t=1)$). Введенный в (2) фактор $\exp(-\alpha d)$ приводит к уменьшению приспособленности агента, это уменьшение тем больше, чем больше было изменение фенотипа S_{Pk} в процессе обучения, т.е. чем больше была нагрузка на обучение.

Отметим, что поскольку начальные генотипы агентов в исходной популяции S_{Gk} случайны, то среднее расстояние по Хеммингу между этими цепочками и оптимальной S_m равно $N/2$. Такое расстояние нужно преодолеть путем обучения и эволюционного поиска.

2.2. Основные результаты исследования формальной модели

2.2.1. Схема и параметры моделирования

Рассматривается два режима работы модели: 1) эволюция с обучением, как изложено выше, 2) «чистая эволюция», в этом режиме обучения не происходит и считается, что фенотипы не меняются и в каждый момент времени $S_{Pk} = S_{Gk}$. Анализ модели был проведен путем компьютерного моделирования.

Параметры расчетов выбирались таким образом, чтобы эволюционный поиск проходил достаточно эффективно, при этом использовался опыт работы [7], в которой сделаны оценки скорости и эффективности эволюционных процессов. Полагалось, что длина цепочек достаточно велика: $N = 100$. Параметр отбора считался равным $\beta = 1$, это соответствует достаточно высокой интенсивности отбора, при которой скорость эволюционного поиска определяется в основном интенсивностью мутаций. Интенсивность мутаций должна быть не слишком высока, чтобы не было мутационных потерь уже найденных в процессе эволюции удачных агентов, но и не слишком мала, для того, чтобы мутационный поиск происходил достаточно быстро. С учетом этого полагалось, что вероятность замены любого символа в цепочках S_{Gk} при мутациях равна $p_m = N^{-1} = 0.01$. При такой интенсивности мутаций p_m за одно поколение происходит замена примерно одного символа в генотипе агента, т.е. в течение одного поколения за счет мутаций расстояние по Хеммингу ρ между цепочками агентов S_{Gk} и оптимальной цепочкой S_m изменяется в среднем на 1. Отбор приводит к умень-

шению этого расстояния. Поскольку расстояние по Хеммингу от цепочек \mathbf{S}_{Gk} агентов нулевого поколения до оптимальной \mathbf{S}_m порядка N , то весь процесс эволюции займет порядка $G_T \sim N$ поколений. Такая оценка скорости эволюции справедлива, если численность популяции достаточно велика и флуктуационными эффектами и нейтральной фиксацией агентов (независящей от приспособленности агентов) можно пренебречь. Для выполнения этого условия достаточно потребовать, чтобы характерное время нейтрального отбора (по порядку величины равное n [6,9]) было не меньше G_T , поэтому полагалось $n = G_T = N$.

Итак, параметры расчета были выбраны следующим образом: $N = 100$, $\beta = 1$, $p_m = N^{-1} = 0.01$, $n = N = 100$.

Дополнительно считалось, что вероятность случайной замены символов при обучении достаточно велика: $p_l = 1$, число тактов времени в течение поколения равно $T_g = 2$ (выбор таких параметров p_l и T_g приводит к тому, что обучение происходит достаточно быстро), параметр ε достаточно мал: $\varepsilon = 10^{-6}$. Полученные зависимости были усреднены по 1000 или 10000 расчетам с разными генераторами случайных чисел.

2.2.2. Сопоставление режимов с обучением и без обучения

На рис. 1 представлена зависимость от номера поколения G среднего по популяции расстояния по Хеммингу $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ между генотипом агента \mathbf{S}_{Gk} и оптимальной цепочкой \mathbf{S}_m . Кривая 1 характеризует режим эволюции с обучением, кривая 2 – режим чистой эволюции (без обучения). Приспособленности агентов определялись выражением (1). Видно, что без обучения при даже малом ε чисто эволюционный процесс (кривая 2) не обеспечивает оптимизации генотипов агентов \mathbf{S}_{Gk} , а при наличии обучения (кривая 1) явно происходит приближение к оптимальному агенту \mathbf{S}_m .

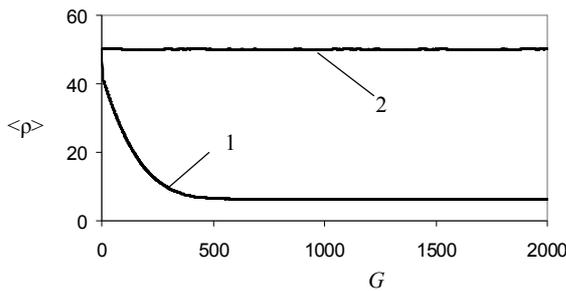


Рис. 1. Зависимость среднего по популяции расстояния генотипов до оптимума $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ от номера поколения G ; 1 – режим с обучением, 2 – режим чистой эволюции (усреднено по 1000 расчетам).

Анализ расчетов показывает, что постепенное уменьшение величин $\rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ при наличии обучения происходит следующим образом. Сначала при обучении распределение агентов $n(\rho)$ смещается в сторону меньших ρ , в результате чего величины $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$ становятся достаточно малыми, такими, что $\exp[-\beta \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)] \sim \varepsilon$. Далее в соответствии с приспособленностями (1) из этих агентов в следующее поколение

отбираются агенты с малыми значениями $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$. Генотипы \mathbf{S}_{Gk} отобранных агентов должны быть достаточно близки к конечным фенотипам \mathbf{S}_{Fk} (полученным в результате обучения и отбора) этих агентов. Поэтому в результате отбора происходит селекция агентов, генотипы которых \mathbf{S}_{Gk} также приближаются к \mathbf{S}_m . То есть происходит генетическая ассимиляция: уменьшение величин $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$, полученное в результате обучения, переходит в уменьшение величин $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$. Отсутствие уменьшения ρ в процессе чистой эволюции обусловлено следующим. Все агенты исходной популяции удалены от оптимальной, для них $\rho \sim N/2 = 50$, следовательно, $\exp(-\rho) \sim 10^{-22}$ и $\exp(-\rho) \ll \varepsilon$. Это означает, что все агенты популяции имеют приближенно одинаковое значение приспособленности, равное ε , поэтому в случае чистой эволюции вообще не происходит селекции агентов и приближения к \mathbf{S}_m . Приближение к \mathbf{S}_m происходит только при наличии обучения.

Механизм генетической ассимиляции представлен на рис. 2, на котором показано распределение агентов $n(\rho)$ по ρ в разные моменты первого поколения. Кривая 1 на рис. 2 показывает распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для исходных генотипов агентов. Кривая 2 показывает распределение по $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$ для агентов после обучения, но еще до отбора. Кривая 3 показывает распределение по $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$ для агентов, отобранных в соответствии с приспособленностями (1). Кривая 4 показывает распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для генотипов отобранных агентов в конце поколения. Генотипы отобранных агентов \mathbf{S}_{Gk} достаточно близки к цепочкам обученных и отобранных агентов \mathbf{S}_{Fk} , поэтому распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для генотипов (кривая 4) перемещается в сторону распределения для конечных фенотипов \mathbf{S}_{Fk} (кривая 3). В итоге после отбора в конце поколения формируется распределение генотипов по ρ (кривая 4), более близкое к распределению, представленному кривой 3, чем начальное распределение генотипов (кривая 1). В следующих поколениях происходит аналогичное дальнейшее смещение распределения $n(\rho)$ в сторону меньших значений ρ .

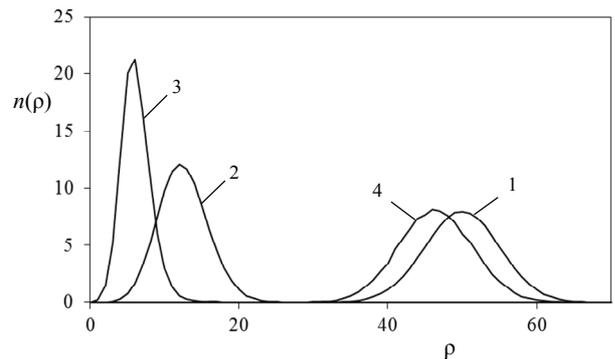


Рис. 2. Распределение агентов $n(\rho)$ по величинам ρ в первом поколении эволюции: 1 – распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для исходных генотипов до обучения, 2 – распределение по $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$ для агентов после обучения, но еще до отбора, 3 – распределение по $\rho = \rho(\mathbf{S}_{Fk}, \mathbf{S}_m)$ для отобранных агентов, 4 – распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для генотипов отобранных агентов в конце поколения (усреднено по 10000 расчетам).

Подчеркнем, что в результате обучения происходит смещение к достаточно малым ρ , при которых роль

параметра ε значительно уменьшается и существенно различаются приспособленности (1), так что после обучения происходит эффективный отбор агентов с меньшими значениями $\rho(\mathbf{S}_{F_k}, \mathbf{S}_m)$. Для того, чтобы изложенный механизм работал, обучение должно быть достаточно сильным. О другой роли сильного обучения сказано в следующем подразделе.

Приведенные результаты показывают, что обучение может приводить к эффективной генетической ассимиляции, к радикальному ускорению эволюционного процесса.

2.2.3. Эффект экранирования

Однако сильное обучение может не только способствовать эволюции, но и препятствовать нахождению оптимального генотипа. Как показывает кривая 1 на рис. 1, при больших G уменьшение $\langle \rho \rangle = \langle \rho(\mathbf{S}_{G_k}, \mathbf{S}_m) \rangle$ ограничено: конечное значение $\langle \rho \rangle$ остается достаточно большим, приближенно равным 6.2. Это связано с тем, что для сильного обучения ($p_l = 1, T_g = 2$) при $G \sim 1000$ нахождение фенотипов \mathbf{S}_{F_k} , близких к \mathbf{S}_m , происходит практически независимо от генотипа \mathbf{S}_{G_k} . Поэтому на конечных стадиях эволюции не происходит приближения генотипов к \mathbf{S}_m . Наблюдается эффект экранирования, обсуждавшийся в работе [3]. Как показано в [1], эффект экранирования можно существенно ослабить, если уменьшить интенсивность обучения, например, положив $p_l = 0.5$.

Итак, эффект экранирования приводит к тому, что сильное обучение препятствует эволюционной оптимизации генотипов, так как оно увеличивает шансы нахождения хорошего фенотипа независимо от генотипа агента. В рассмотренном случае экранирование наблюдается в конце эволюционного процесса.

2.2.4. Влияние нагрузки на обучение на моделируемые процессы

В [1] был также проведен анализ влияния нагрузки на обучение на моделируемые процессы эволюции с обучением. Для этого случая приспособленность определялась выражением (2). Расчет проводился для указанных выше параметров ($N = n = 100, \beta = 1, p_m = 0.01, p_l = 1, T_g = 2, \varepsilon = 10^{-6}$), величина α полагалась равной 1. Результаты расчета представлены на рис. 3,4. На рис. 3 показана зависимость среднего по популяции расстояния $\langle \rho \rangle$ между генотипами \mathbf{S}_{G_k} и \mathbf{S}_m от номера поколения G . На рис. 4 представлена динамика распределения особей $n(\rho)$ по величинам ρ в первом поколении эволюции.

Сравнение рис. 1,2 и рис. 3,4 показывает, что учет нагрузки на обучение приводит к дополнительному значительному ускорению эволюции. Это ускорение связано с тем, что с учетом нагрузки при отборе происходит более сильная селекция агентов, для которых мало расстояние между начальной и конечной цепочками $\rho(\mathbf{S}_{G_k}, \mathbf{S}_{F_k})$, чем для случая приспособленности (1), т.е. дополнительно минимизируется изменение фенотипа \mathbf{S}_{F_k} в процессе обучения. В конце эволюционного процесса (при $G = 200$) агенты, для которых $\rho = 0$, в популяции присутствуют в начале поколения, после обучения и после отбора, т.е. при наличии нагрузки на

обучение эффект экранирования не наблюдается.

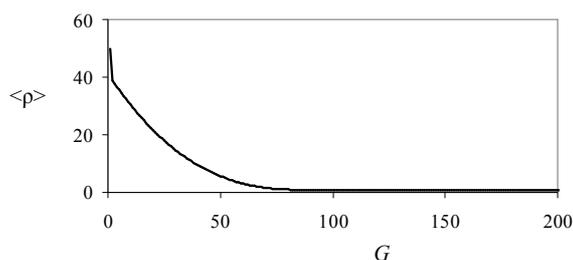


Рис. 3. Зависимость среднего по популяции расстояния до оптимума $\langle \rho \rangle = \langle \rho(\mathbf{S}_{G_k}, \mathbf{S}_m) \rangle$ для генотипов агентов от номера поколения G ; приспособленность агентов определяется выражением (2), уменьшение $\langle \rho \rangle$ происходит намного быстрее, чем на рис. 1 (усреднено по 1000 расчетам).

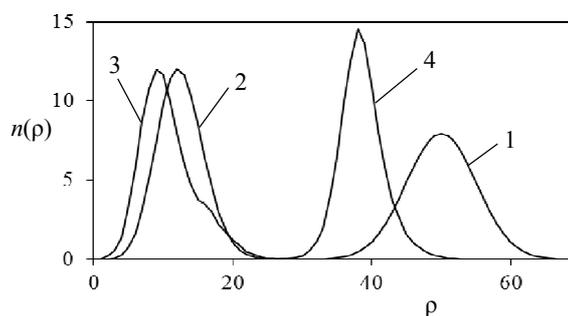


Рис. 4. Распределение агентов $n(\rho)$ по величинам ρ в первом поколении эволюции с учетом нагрузки на обучение; приспособленность агентов определяется выражением (2); 1 – распределение по $\rho = \rho(\mathbf{S}_{G_k}, \mathbf{S}_m)$ для исходных генотипов до обучения, 2 – распределение по $\rho = \rho(\mathbf{S}_{F_k}, \mathbf{S}_m)$ для агентов после обучения, но еще до отбора, 3 – распределение по $\rho = \rho(\mathbf{S}_{F_k}, \mathbf{S}_m)$ для отобранных агентов, 4 – распределение по $\rho = \rho(\mathbf{S}_{G_k}, \mathbf{S}_m)$ для генотипов отобранных агентов в конце поколения; смещение распределения 4 к меньшим значениям ρ существенно больше, чем на рис. 2 (усреднено по 10000 расчетам).

Подчеркнем, что генетическая ассимиляция для случаев приспособленностей, определяемых выражением (1) и выражением (2), имеет одну и ту же природу. В обоих случаях генотипы отобранных агентов \mathbf{S}_{G_k} приближаются к цепочкам \mathbf{S}_{F_k} обученных и отобранных особей. Т.е. на рис. 2 и рис. 4 кривые 2 и 4 приближаются к кривым 3. Существенное различие состоит только в том, что с учетом нагрузки на обучение это приближение становится более эффективным. В результате при использовании приспособленности, вычисляемой согласно (2), генетическая ассимиляция становится более четко выраженной.

Таким образом, учет нагрузки на обучение приводит к тому, что происходит более эффективная селекция особей с генотипами \mathbf{S}_{G_k} , близкими к \mathbf{S}_m , и процесс эволюции значительно ускоряется.

2.3. Выводы по формальной модели

Таким образом, в рамках формальной модели количественно охарактеризованы 1) механизм генетической ассимиляции приобретаемых навыков, 2) эффект экранирования и 3) влияние нагрузки на обучение на

исследованные процессы. Подробнее детали исследования формальной модели изложены в работе [1].

3. Модель формирования импринтинга

3.1. Описание модели

Предполагаем, что имеется эволюционирующая популяция агентов, аналогов цыплят. Популяция находится в двумерном клеточном мире. Также имеется агент-курица, которая может перемещаться по этому миру. Основная задача агентов-цыплят – следовать за курицей. Если цыпленок удалился от курицы, то он может погибнуть. Как и формальной модели предполагаем, что каждый агент-цыпленок имеет генотип и фенотип. Генотипы меняются в процессе эволюции. Фенотип меняется индивидуально в процессе обучения агента-цыпленка. Фенотип цыпленка определяет его систему управления. Численность популяции агентов-цыплят равна n .

Двумерный клеточный мир состоит из $N_x N_y$ клеток (координаты агентов могут быть равны $i_x = 1, 2, \dots, N_x$, $i_y = 1, 2, \dots, N_y$). Причем мир замкнут: например, если агент перемещается из клетки с координатами (N_x, i_y) вправо, то он попадает в клетку с координатами $(1, i_y)$. Время дискретно: $t = 1, 2, \dots$. Курица перемещается через каждые T_M тактов времени в случайном направлении на одну клетку.

Система управления агента-цыпленка определяет, какие действия должен цыпленок выполнять в той или иной ситуации. Ситуация определяется наличием или отсутствием курицы в той или иной клетке поля зрения цыпленка. Поле зрения включает в себя 5 клеток: ту клетку, в которой цыпленок находится, и 4-х ближайших к нему клеток (клеток впереди, сзади, направо и налево от цыпленка). Таким образом, имеется $N_S = 6$ возможных ситуаций: курица находится в одной из 5-ти клеток поля зрения цыпленка, либо цыпленок не видит курицы. Имеется $N_A = 5$ возможных действий цыпленка: не перемещаться или перемещаться в одну из ближайших 4-х клеток. В системе управления цыпленка имеется $N_R = N_A N_S = 30$ правил, определяющих соответствие между ситуациями и действиями. Каждое правило имеет вес W_i , $i = 1, 2, \dots, N_R$. Чтобы было удобно сопоставление с формальной моделью, предполагаем, что веса правил бинарны: веса W_i равны 0 либо 1. Веса правил оптимизируются путем обучения и эволюционного поиска. Совокупность этих N_R правил представляет собой систему управления цыпленка.

При рождении цыпленка он получает от своего родителя генотип S_G , т.е. начальные веса правил. Генотип агента-цыпленка в процессе его жизни остается неизменным. Фенотип цыпленка в каждый момент времени t равен текущему вектору весов его правил: $S_F(t) = \mathbf{W}(t)$. Как и для формальной модели полагаем, что в момент рождения цыпленка его фенотип равен генотипу: $S_F(t = 1) = S_G$. В процессе жизни цыпленка текущие веса правил оптимизируются путем обучения.

Выбор действия k -м цыпленком определяется имеющимися у него в данный момент времени t весами правил $\mathbf{W}_k(t) = S_{Fk}(t)$ следующим образом. Рассмат-

риваются все правила, соответствующие текущей ситуации, среди этих правил выбираются те, для которых вес равен 1. Если такое правило одно, то выполняется действие, соответствующее этому правилу. Если таких правил оказалось несколько, то случайно выбирается произвольное из них, и действие, соответствующее этому выбранному правилу, выполняется. Если для данной ситуации в данный момент времени все веса правил равны 0, то выбирается действие, случайное из 5-ти возможных.

Обучение, т.е. изменение текущих весов правил $\mathbf{W}(t)$, происходит следующим образом. Если в определенной ситуации в результате действия цыпленка он приблизился к курице, то вес, соответствующий этой ситуации и действию, становится равным 1. Если, наоборот, цыпленок удалился от курицы, то соответствующий вес становится равным 0. Такое обучение предполагает, что цыпленок может положительно оценивать наличие курицы в поле зрения и свое приближение к курице. При компьютерном моделировании обучение обычно происходило каждый такт времени. Вводилась также возможность уменьшения интенсивности обучения, для этого случая каждый такт времени обучение производилась с определенной вероятностью p_l ($p_l < 1$).

Кроме этого, сделано дополнительное предположение относительно изменения весов правил. Считается, что вес правила, соответствующего текущей ситуации и выполненному действию, становится равным 1, если после действия цыпленок находится в одной клетке с курицей, и что вес правила становится равным 0, если после применения этого правила курицы нет в поле зрения цыпленка. Это дополнительное предположение способствует разумному формированию весов правил в начальные моменты жизни цыпленка. Это предположение модели будем называть дополнительным предположением А.

Так как цыпленок может оценивать свое приближение к курице или удаление от курицы, то предполагаем, что если в предыдущий момент времени цыпленок удалился от курицы, то в данный момент времени он возвращается обратно. Это предположение модели будем называть дополнительным предположением В.

В каждый такт времени оценивается уровень безопасности цыплят L . Считается, что этот уровень безопасности максимален ($L = L_{max}$), если цыпленок находится в одной клетке с курицей. Этот уровень средний ($L = L_{mean}$), если цыпленок находится в соседней клетке с курицей, и этот уровень минимальный ($L = L_{min}$), если цыпленок удален от курицы больше чем на одну клетку.

Популяция агентов-цыплят эволюционирует в течение ряда поколений. В начале первого поколения веса правил для цыплят различны и случайны. Каждое поколение длится T_g тактов времени, в течение этого времени цыплята обучаются. В начальный момент каждого поколения курица и все цыплята помещаются в одну клетку. Считается, что к концу поколения цыплята вырастают и рожают своих потомков. В конце поколения определяются приспособленности агентов-цыплят и производится отбор агентов в следующее поколение. Чем выше приспособленность агента, тем больше потомков он может родить.

Приспособленности определяются следующим образом. После того, как пройдет T_g тактов времени, выделяется еще определенное количество тактов T_{test} , в течение которых цыплята уже не обучаются, а только оценивается их уровень безопасности. То есть вычисляется сумма всех уровней безопасности для каждого агента-цыпленка в течение T_{test} тактов времени:

$$L_{sk} = \sum_{t=T_g+1}^{T_g+T_{test}} L_k(t), \quad (3)$$

где $L_k(t)$ – уровень безопасности k -го цыпленка в такт времени t , L_{sk} – суммарный уровень безопасности этого цыпленка. L_{sk} характеризует уровень безопасности цыпленка в конце поколения после того, как обучение закончилось. Приспособленность цыпленка определяется разностью между максимально возможным суммарным уровнем безопасности, равным $L_{sm} = L_{max}T_{test}$, и реально достигнутым уровнем L_{sk} , а именно, приспособленность k -го цыпленка равна:

$$f_k = \exp[-\beta(L_{sm} - L_{sk})] + \varepsilon, \quad (4)$$

где β – положительный параметр, характеризующий интенсивность отбора, $0 < \varepsilon \ll 1$. Как и в формальной модели роль слагаемого ε можно рассматривать как влияние случайных факторов окружающей среды на приспособленность агентов.

В конце поколения производится отбор агентов, родителей будущих агентов нового поколения, рулеточным методом. При каждом выборе вероятность селекции определенного агента-родителя пропорциональна его приспособленности. Каждый агент-родитель рождает одного потомка. Эти потомки составляют следующее поколение. Выбор агентов-родителей производится n раз, в результате численность популяции во всех поколениях равна n . Генотип S_g передается от родителя к потомку с небольшими мутациями.

Такая форма отбора подразумевает, что приспособленность неявно характеризует вероятность гибели цыплят при удалении их от курицы: чем больше суммарный уровень безопасности цыпленка L_{sk} , тем выше его приспособленность.

Таким образом, как и формальной модели, обученные агенты-цыплята отбираются по конечным фенотипам $S_{pk}(t = T_g)$, т.е. по весам правил $W_k(t = T_g) = S_{pk}(t = T_g)$, полученным в результате обучения, а по наследству передаются потомкам (с малыми мутациями) генотипы S_{Gk} , т.е. начальные веса правил агентов.

Поскольку потомкам передаются именно генотипы S_{Gk} , полученные от родителей, а не конечные фенотипы $S_{pk}(t = T_g)$, полученные в результате обучения, то эволюционный процесс имеет дарвиновский характер.

Как и в формальной модели, учтем то, что процесс обучения может иметь определенную нагрузку для агента-цыпленка и приспособленность агента может уменьшаться под влиянием этой нагрузки. С этой целью, помимо приспособленностей (4), будем рассматривать и модифицированные следующим образом приспособленности агентов:

$$f_{mk} = \exp(-\alpha d_k) \{ \exp[-\beta(L_{sm} - L_{sk})] + \varepsilon \}, \quad (5)$$

где α – положительный параметр, учитывающий нагрузку, связанную с обучением, $d_k = \rho(W_k(t=1), W_k(t=T_g))$ – расстояние по Хеммингу между начальной и конечной цепочками весов правил W_k k -го

агента-цыпленка, т.е. величина, характеризующая интенсивность всего процесса обучения этого агента. Фактор $\exp(-\alpha d_k)$ приводит к уменьшению приспособленности агента, это уменьшение тем больше, чем больше было изменение весов правил W_k в процессе обучения, т.е. чем больше была нагрузка на обучение.

3.2. Результаты моделирования

3.2.1. Схема и параметры моделирования

В основном будем использовать подход, примененный при исследовании формальной модели. Сопоставим режим эволюции с обучением и режим чистой эволюции, проанализируем механизмы генетической ассимиляции, эффект экранирования и влияние нагрузки на обучение на исследуемые процессы. Дополнительно рассмотрим еще некоторые особенности модели агентов-цыплят, в частности, проанализируем возможную роль случайной гибели агентов-цыплят при удалении их от курицы.

Генотип и фенотип агента-цыпленка определяются набором весов правил. Генотип S_g – начальный набор весов правил, передаваемый по наследству, фенотип S_p – текущий набор весов правил. Длина цепочек N генотипа и фенотипа равна числу правил, т.е. $N = N_R = 30$.

Как и формальной модели, введем понятие оптимальной цепочки S_m (длина этой цепочки также равна N), которая определяет веса правил следующим образом:

1) Если ситуация такова, что агент-цыпленок находится в той же клетке, что и курица, то ему не нужно перемещаться, т.е. цыпленку целесообразно осуществлять действие «покой». Считаем, что для оптимальной цепочки S_m вес правила, соответствующего указанным ситуации и действию, равен 1, веса правил, соответствующих этой ситуации и всем остальным действиям, для цепочки S_m полагаем равными 0.

2-5) Если курица находится в одной из ближайших к цыпленку клеток (таких клеток четыре), то цыпленку целесообразно перемещаться в эту ближайшую клетку (например, если курица находится в клетке справа от цыпленка, то цыпленку нужно переместиться на одну клетку вправо). Считаем, что для оптимальной цепочки S_m веса правил, соответствующих таким ситуациям и целесообразным в этих ситуациях действиям, равны 1.

Веса всех остальных не указанных правил для цепочки S_m полагаем равными 0.

Таким образом, оптимальная цепочка S_m определяет веса правил, при которых цыпленок четко следует за курицей. Только 5 символов этой цепочки равны 1, все остальные равны 0.

Используя оптимальную цепочку S_m , можно определить расстояние по Хеммингу ρ между генотипом S_g или фенотипом S_p рассматриваемого цыпленка и S_m . При вычислении этого расстояния одну ситуацию, при которой курица находится вне поля зрения цыпленка, можно не учитывать, так как в этой ситуации оптимальное действие цыпленка не определено. Таким образом, определяем расстояние до оптимума $\rho = \rho(S_g, S_m)$ (или $\rho = \rho(S_p, S_m)$), как число несовпадающих символов в цепочках S_g и S_m (или в цепочках S_p и S_m)

для 5-ти указанных ситуаций (курица в той же клетке, что и цыпленок, или в соседней с цыпленком клетке) и для 5-ти возможных действий цыпленка. Это расстояние может меняться от 0 до 25.

Аналогично формальной модели, основные параметры эволюционных процессов определяем следующим образом: $N = 30$, $\beta = 1$, $p_m = 0.03 \approx N^{-1}$, $n = N = 30$. Здесь p_m – вероятность случайной замены символа в генотипе при мутациях в процессе наследования генотипов (т.е. при передаче генотипа родителя к потомку при переходе от старого поколения к новому).

Новые параметры модели формирования импринтинга выбираем следующим образом. Длительность поколения считаем достаточно большой $T_g = 100$, время тестирования при оценке сформировавшегося к концу поколения уровня безопасности считаем равным $T_{test} = 100$, параметр ϵ считаем достаточно малым: $\epsilon = 0.001$. Эта величина ϵ выбрана таким образом, чтобы стохастичность внешней среды существенно влияла на приспособленности агентов-цыплят, а обучение приводило к существенному различию приспособленностей агентов. Уровни безопасности считались равными: $L_{max} = 1$ (если цыпленок находится в одной клетке с курицей), $L_{mean} = 0.5$ (если цыпленок находится в соседней клетке с курицей) и $L_{min} = 0$ (если цыпленок удален от курицы больше, чем на одну клетку). Размеры клеточного мира считаем равными: $N_x = N_y = 10$. Длительность остановок курицы, т.е. время между случайными перемещениями курицы в большинстве расчетов было равно: $T_M = 10$, в отдельных указанных ниже расчетах величина T_M уменьшалась. Большинство расчетов было выполнено при вероятности обучения $p_l = 1$, только в одном случае считалось $p_l = 0.1$.

Анализ модели был проведен путем компьютерного моделирования. Полученные зависимости усреднялись по 1000 или 100000 расчетам с разными генераторами случайных чисел.

Сначала изложим результаты, аналогичные формальной модели.

3.2.2. Сопоставление режимов с обучением и без обучения

На рис. 5 представлены зависимости среднего по популяции расстояния по Хеммингу $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ между генотипом агента \mathbf{S}_{Gk} и оптимальной цепочкой \mathbf{S}_m от номера поколения G для режимов эволюции с обучением (кривая 1) и режима чистой эволюции, без обучения (кривая 2). Приспособленности агентов определялись выражением (4). Видно, что оптимизация геномов для чисто эволюционного процесса (кривая 2) происходит значительно медленней, чем при наличии обучения (кривая 1). Это аналогично явному преимуществу режима с обучением по сравнению со случаем чистой эволюции для формальной модели (рис. 1). Хотя имеется определенное отличие от формальной модели: в настоящей модели в режиме чистой эволюции хоть медленно, но оптимизация генотипов происходит. Это обусловлено тем, что в данном случае размерность N цепочек \mathbf{S}_{Gk} существенно меньше, чем для расчетов в формальной модели.

Зависимость среднего по популяции расстояния по Хеммингу $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ между генотипом агента \mathbf{S}_{Gk} и

оптимальной цепочкой \mathbf{S}_m от номера поколения G для режима чистой эволюции для больших значений G представлена на рис. 6. Видно, что без обучения эволюционный процесс поиска оптимума сходится в течение нескольких тысяч поколений. Анализ расчетов для чистой эволюции показывает, что при $G = 15000$ генотипы \mathbf{S}_{Gk} , равные оптимальной цепочке \mathbf{S}_m , в популяции присутствует, т.е. эволюционный поиск обеспечивает нахождение оптимального генотипа.

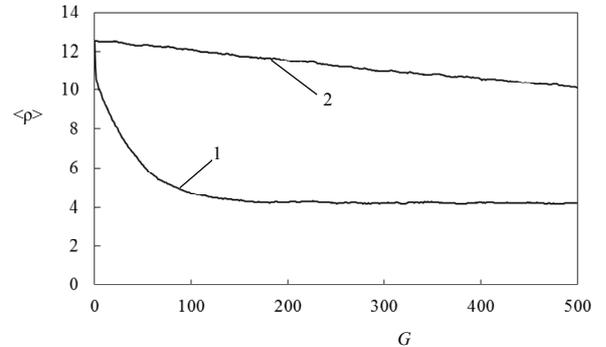


Рис. 5. Зависимость среднего по популяции расстояния между генотипами агентов \mathbf{S}_{Gk} и оптимальной цепочкой $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ от номера поколения G ; 1 – режим с обучением, 2 – режим чистой эволюции (усреднено по 1000 расчетам).

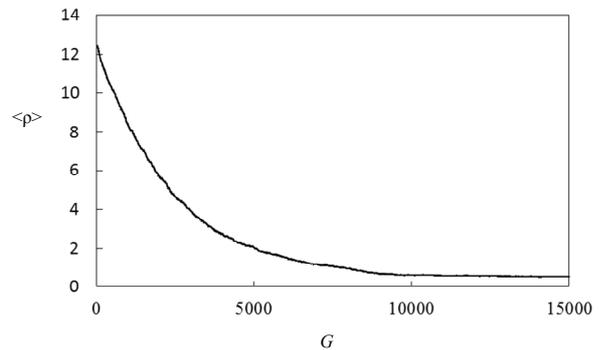


Рис. 6. Зависимость среднего по популяции расстояния между генотипами агентов \mathbf{S}_{Gk} и оптимальной цепочкой $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ от номера поколения G для режима чистой эволюции (усреднено по 1000 расчетам).

Как и в случае формальной модели, при наличии обучения происходит генетическая ассимиляция приобретаемого в процессе обучения навыка приближения к оптимуму. Механизм генетической ассимиляции представлен на рис. 7, на котором показано распределение агентов $n(\rho)$ по ρ в разные моменты первого поколения. Кривая 1 показывает распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для исходных генотипов цыплят-агентов в начале поколения. Кривая 2 показывает распределение по $\rho = \rho(\mathbf{S}_{Pk}, \mathbf{S}_m)$ для фенотипов агентов после обучения, но еще до отбора. Кривая 3 показывает распределение по $\rho = \rho(\mathbf{S}_{Pk}, \mathbf{S}_m)$ для агентов, отобранных в соответствии с приспособленностями (4). Кривая 4 показывает распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для генотипов отобранных агентов в конце поколения. Генотипы отобранных агентов \mathbf{S}_{Gk} достаточно близки к фенотипам обученных и отобранных агентов \mathbf{S}_{Pk} , поэтому распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для генотипов (кривая 4) перемещается в сторону распределения для конечных

фенотипов S_{P_k} (кривая 3). В итоге после отбора в конце поколения формируется распределение генотипов по ρ (кривая 4), более близкое к распределению, представленному кривой 3, чем начальное распределение генотипов (кривая 1).

Сопоставление рис. 2 и рис. 7 показывает, что механизм генетической ассимиляции для агентов-цыплят такой же, как и для формальной модели. Генетическая ассимиляция происходит за счет того, что при обучении распределение фенотипов агентов смещается в сторону оптимума, а в дальнейшем при отборе выбираются те агенты, генотипы которых также приближаются к оптимуму.

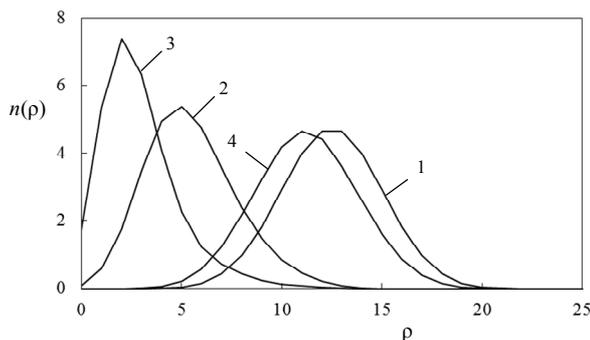


Рис. 7. Распределение агентов-цыплят $n(\rho)$ по величинам ρ в первом поколении эволюции: 1 – распределение по $\rho = \rho(S_{G_k}, S_m)$ для исходных генотипов до обучения, 2 – распределение по $\rho = \rho(S_{P_k}, S_m)$ для фенотипов агентов после обучения, но еще до отбора, 3 – распределение по $\rho = \rho(S_{P_k}, S_m)$ для отобранных агентов, 4 – распределение по $\rho = \rho(S_{G_k}, S_m)$ для генотипов отобранных агентов в конце поколения (усреднено по 100000 расчетам).

Рис. 5-7 показывают, что, как и в формальной модели, в данном случае обучение радикально ускоряет эволюционный процесс.

3.2.3. Эффект экранирования

Для того чтобы изложенный механизм генетической ассимиляции работал, обучение должно быть достаточно сильным. Однако сильное обучение может не только способствовать эволюции, но и препятствовать нахождению оптимального генотипа. Как показывает кривая 1 на рис. 5, при больших G уменьшение $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ ограничено: конечное значение $\langle \rho \rangle$ остается достаточно большим, приблизительно равным 4.2. Это связано с тем, что для сильного обучения ($p_l = 1$) при $G \sim 500$ нахождение фенотипов S_{P_k} , близких к оптимальной цепочке S_m , происходит практически независимо от генотипа S_{G_k} . Поэтому на конечных стадиях эволюции не происходит приближения генотипов к S_m . Наблюдается эффект экранирования [3].

Этот эффект иллюстрируется рис. 8, на котором представлены распределения по ρ для генотипов и фенотипов агентов-цыплят в конце эволюционного процесса (при $G = 500$) для режима эволюции с обучением в разные моменты последнего поколения. Видно, что распределение по $\rho = \rho(S_{P_k}, S_m)$ для фенотипов после обучения включает в себя агентов, для которых $\rho = 0$, т.е. после обучения находятся фенотипы S_{P_k} , равные оптимальной цепочке S_m . В результате в конце эволю-

ционного процесса дальнейшего уменьшения $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ не происходит. Это подтверждается тем, что в конце эволюции показанные на рис. 8 кривые от поколения к поколению практически не смещаются, а также согласуется с тем, что при больших G величина $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ становится постоянной (кривая 1 на рис. 5). То, что распределения по ρ в генотипах в начале поколения и после отбора (кривые 1 и 4 на рис. 8) немного различаются, связано с наличием мутаций, которые приводят к небольшому увеличению ρ в начале поколения по сравнению с распределением после отбора. Итак, при $G \sim 500$ за счет обучения находятся оптимальные фенотипы, и в результате этого дальнейшего уменьшения $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ в популяции не происходит.

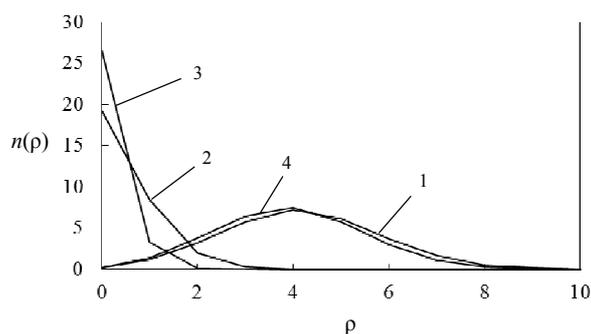


Рис. 8. Распределение агентов $n(\rho)$ по величинам ρ при $G = 500$: 1 – распределение по $\rho = \rho(S_{G_k}, S_m)$ для исходных генотипов до обучения, 2 – распределение по $\rho = \rho(S_{P_k}, S_m)$ для фенотипов агентов после обучения, но еще до отбора, 3 – распределение по $\rho = \rho(S_{P_k}, S_m)$ для отобранных агентов, 4 – распределение по $\rho = \rho(S_{G_k}, S_m)$ для генотипов отобранных агентов в конце поколения (усреднено по 1000 расчетам).

Эффект экранирования можно существенно ослабить, если уменьшить вероятность обучения p_l . Зависимость $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ от номера поколения G для ослабленного обучения ($p_l = 0.1$, прочие параметры расчета остались прежними, расчет проведен для режима эволюции с обучением) представлена на рис. 9.

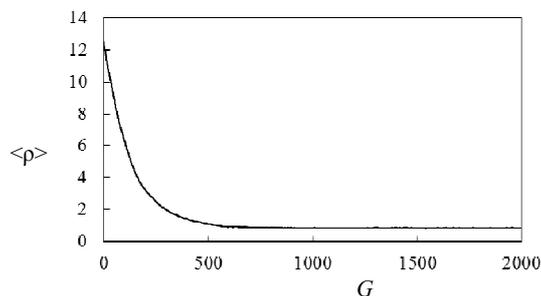


Рис. 9. Зависимость среднего по популяции расстояния генотипов до оптимума $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ от номера поколения G при ослабленном обучении: $p_l = 0.1$ (усреднено по 1000 расчетам); по сравнению со случаем $p_l = 1$ скорость обучения уменьшается, но находятся генотипы агентов, значительно более близкие к S_m .

Для этого случая скорость уменьшения $\langle \rho \rangle = \langle \rho(S_{G_k}, S_m) \rangle$ в процессе эволюции падает, однако конечное значение $\langle \rho \rangle$ значительно уменьшается и становится приблизительно равным 0.85. Итак, значение

$\langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ в конце эволюции значительно уменьшается по сравнению со случаем $p_l = 1$. Следовательно, ослабление обучения приводит к тому, что происходит отбор агентов, генотипы которых достаточно близки к \mathbf{S}_m . Более того, анализ результатов расчетов показывает, что в конце эволюционного процесса (при $G = 2000$) агенты, для которых $\rho = 0$, в популяции присутствуют в начале поколения, после обучения и после отбора.

Итак, эффект экранирования приводит к тому, что сильное обучение препятствует эволюционной оптимизации генотипов, так как оно увеличивает шансы нахождения хорошего фенотипа независимо от генотипа агента. В рассмотренном случае экранирование наблюдается в конце эволюционного процесса. Ослабление обучения приводит к уменьшению экранирования.

3.2.4. Влияние нагрузки на обучение на моделируемые процессы

Был проведен анализ влияния нагрузки на обучение на моделируемые процессы эволюции с обучением. Для этого случая приспособленность определялась выражением (5). Расчет проводился для указанных выше параметров ($N = n = 30$, $\beta = 1$, $p_m = 0.03$, $p_l = 1$, $T_g = 100$, $\varepsilon = 0.001$, $T_M = 10$), величина α полагалась равной 1. Результаты расчета представлены на рис. 10, 11. На рис. 10 показана зависимость среднего по популяции расстояния $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ между генотипами \mathbf{S}_{Gk} и оптимальной цепочкой \mathbf{S}_m от номера поколения G . На рис. 11 представлена динамика распределения агентов $n(\rho)$ по величинам ρ в первом поколении эволюции в разные моменты поколения.

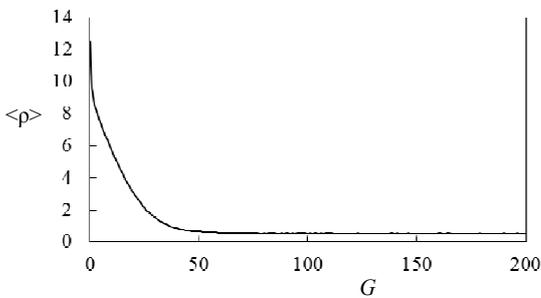


Рис. 10. Зависимость среднего по популяции расстояния до оптимума $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ для генотипов агентов от номера поколения G ; приспособленность агентов определяется выражением (5), уменьшение $\langle \rho \rangle$ происходит значительно быстрее, чем на рис. 5 (усреднено по 1000 расчетам).

Сравнение рис. 5, 7 и рис. 10, 11 показывает, что учет нагрузки на обучение приводит к дополнительному значительному ускорению эволюции. Это ускорение связано с тем, что с учетом нагрузки при отборе происходит более сильная селекция агентов, для которых мало расстояние между начальным и конечным фенотипами (т.е. между генотипами и конечными фенотипами), чем для случая приспособленности (4). Отметим, что распределение 3 на рис. 11 имеет немного «растянутый хвост» справа, это можно интерпретировать как то, что в соответствии с приспособленностями (5) преимущественно отбираются агенты, для которых

мало изменение фенотипов \mathbf{S}_{P_k} при обучении. В конце эволюционного процесса (при $G = 200$) агенты, для которых $\rho = 0$, в популяции присутствуют в начале поколения, после обучения и после отбора. Эффекта экранирования в данном случае не наблюдается.

Таким образом, учет нагрузки на обучение приводит к тому, что происходит более эффективная селекция агентов с генотипами \mathbf{S}_{Gk} , близкими к \mathbf{S}_m , генетическая ассимиляция усиливается и процесс эволюции значительно ускоряется.

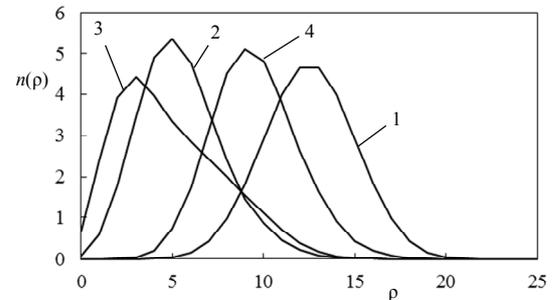


Рис. 11. Распределение агентов-цыплят $n(\rho)$ по величинам ρ в первом поколении эволюции с учетом нагрузки на обучение; приспособленность агентов определяется выражением (5); 1 – распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для исходных генотипов до обучения, 2 – распределение по $\rho = \rho(\mathbf{S}_{P_k}, \mathbf{S}_m)$ для фенотипов агентов после обучения, но еще до отбора, 3 – распределение по $\rho = \rho(\mathbf{S}_{P_k}, \mathbf{S}_m)$ для отобранных агентов, 4 – распределение по $\rho = \rho(\mathbf{S}_{Gk}, \mathbf{S}_m)$ для генотипов отобранных агентов в конце поколения; смещение распределения 4 к меньшим значениям ρ существенно больше, чем на рис. 7 (усреднено по 100000 расчетам).

3.2.5. Анализ особенностей модели формирования импринтинга

Итак, в модели формирования импринтинга наблюдаются те же эффекты взаимодействия между обучением и эволюцией, что и в формальной модели: генетическая ассимиляция приобретаемых в процессе индивидуального обучения навыков, эффект экранирования и существенное влияние нагрузки на обучение на процессы поиска оптимума. Основное отличие состоит в том, что в модели формирования импринтинга чисто эволюционная оптимизация (без обучения) весов правил, хоть и медленно, но происходит (см. рис. 1, 5, 6). Это различие связано с тем, что размерность пространства поиска N в данной модели значительно меньше, чем для представленных результатов в формальной модели.

Кроме приведенных результатов, для модели формирования импринтинга были проанализированы следующие ее особенности: 1) была проанализирована возможная роль явной гибели агентов-цыплят при удалении их от курицы, рассматривались 2) возможность более частого перемещения курицы, чем предполагалось выше, 3) возможность случайного выбора действия агентом-цыпленком независимо от весов правил, 4) возможность исключения из модели дополнительных предположений А и В (см. выше, в описании модели) о режиме обучения и дополнительном переме-

щении цыплят обратно в случае понижения их уровня безопасности.

Явная гибель цыплят предполагала, что агент-цыпленок с определенной вероятностью P_D в рассматриваемый такт времени погибает, если его уровень безопасности в этот такт равен минимальному значению L_{min} . Расчет с учетом такой явной гибели цыплят для режима эволюции с обучением иллюстрируется рис. 12. На рисунке представлена зависимость среднего по популяции расстояния $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ между генотипами \mathbf{S}_{Gk} и оптимальной цепочкой \mathbf{S}_m от номера поколения G для случаев явного учета гибели (кривая 2) и для обычного расчета (кривая 1). Напомним, что в обычном расчете гибель цыплят учитывалась неявно: предполагалось, что приспособленность агентов учитывает такую возможность в величине приспособленности (4). Кривая 2 на рис. 12, помимо этого, учитывает и явную гибель цыплят, вероятность P_D для этого расчета составляла величину: $P_D = 0.1$.

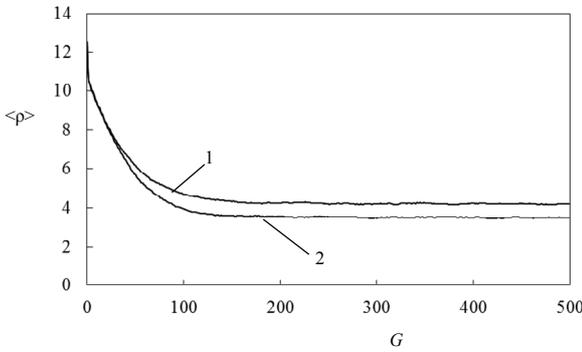


Рис. 12. Зависимость среднего по популяции расстояния между генотипами агентов \mathbf{S}_{Gk} и оптимальной цепочкой $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ от номера поколения G ; 1 – обычный режим, 2 – расчет с учетом явной гибели цыплят при $P_D = 0.1$ (усреднено по 1000 расчетам).

Видно, что учет явной гибели цыплят не существенно меняет процесс поиска оптимальных весов правил. Скорость уменьшения $\langle \rho \rangle$ для обеих зависимостей на рис. 12 приблизительно одинакова. Ослабевает только эффект экранирования: для обычного расчета конечное значение $\langle \rho \rangle \approx 4.2$, с учетом гибели цыплят эта величина уменьшается до значения $\langle \rho \rangle \approx 3.5$. Варьирование вероятности гибели цыплят P_D приводило к аналогичным результатам.

При рассмотрении более частого перемещения курицы полагалось, что время покоя курицы между случайными перемещениями на одну клетку равно не $T_M = 10$, как выше, а несколько меньше, например, $T_M = 4$ или $T_M = 5$. Результаты расчетов для таких величин T_M показали, что уменьшение T_M приводит примерно к тем же результатам, что ослабление обучения: скорость уменьшения $\langle \rho \rangle = \langle \rho(\mathbf{S}_{Gk}, \mathbf{S}_m) \rangle$ немного падает, конечное значение $\langle \rho \rangle$ также уменьшается. Например, при $T_M = 4$ конечное значение $\langle \rho \rangle$ составило: $\langle \rho \rangle \approx 2.2$.

При случайном выборе действий цыплятами агентами независимо от весов правил поиска оптимальных весов правил, естественно, не происходило.

В компьютерной программе была предусмотрена возможность исключения одного или двух дополни-

тельных предположений А и В о режиме обучения и дополнительном перемещении цыплят обратно в случае понижения их уровня безопасности. Расчеты показали, что такие исключения не приводили с существенным изменениям в поиске оптимума. Хотя наблюдались небольшие особенности, характерные для расчетов с ослабленным обучением. Например, при исключении двух предположений А и Б наблюдалось уменьшение скорости уменьшения $\langle \rho \rangle$ примерно в 2 раза по сравнению с обычным расчетом в режиме эволюции с обучением, конечная величина $\langle \rho \rangle$ составила: $\langle \rho \rangle \approx 2.6$.

4. Обсуждение, заключение

В настоящей работе построена и исследована модель формирования импринтинга для агентов-цыплят, которым целесообразно следовать за курицей. Модель предполагает, что цыпленок способен отличать благоприятную ситуацию (быть рядом с курицей) от неблагоприятной (быть далеко от курицы). Эта способность лежит в основе процесса обучения цыплят. Если мать-курица вообще отсутствует (например, в случае, когда цыплят вырастили в инкубаторе), то аналогичная способность отличать благоприятную ситуацию от неблагоприятной также возможна, если вместо курицы имеется другой, подобный курице защитник цыплят. Такую возможность несложно рассмотреть в рамках настоящей модели, видоизменяя величины уровней безопасности L_{max} , L_{mean} и L_{min} .

Таким образом, модель формирования импринтинга предполагает, что у цыплят имеется врожденное понятие «защитник» и врожденное представление о том, что приближение к защитнику полезно, а удаление от защитника вредно для цыплят. Если такого врожденного чувства у цыплят нет, то рассмотренное в модели обучение невозможно, в этом случае формирование импринтинга возможно только в результате эволюционной оптимизации поведения агентов-цыплят. В нашей модели эволюционная оптимизация иллюстрируется рис. 6. При такой оптимизации отбор агентов-цыплят осуществляется по их суммарному уровню безопасности, т.е. импринтинг в этом случае формируется, хотя и медленней, чем при наличии обучения.

Подводя итог, отметим, что построена и исследована модель формирования импринтинга и изучены механизмы взаимодействия между обучением и эволюционной оптимизацией в процессах формирования импринтинга.

Благодарности

Работа выполнена при поддержке программы Президиума РАН № 15, проект 2.10 и РФФИ, проект 13-01-00399. Авторы благодарны Л.Е. Цитоловскому за предложение моделировать явление импринтинга.

Литература

1. В.Г. Редько Модель взаимодействия между обучением и эволюционной оптимизацией // *Математи-*

ческая биология и биоинформатика (электронный журнал). – Т.7. № 2. 2012. С. 676-691.

URL: http://www.matbio.org/2012/Redko_7_676.pdf

2. G.E. Hinton, S.J. Nowlan How learning can guide evolution // *Complex Systems*. – V. 1. 1987. P. 495-502.

3. G. Mayley Guiding or hiding: Explorations into the effects of learning on the rate of evolution // In: *Proceedings of the Fourth European Conference on Artificial Life (ECAL 97)*. Eds. Husbands P. and Harvey I. Cambridge, Massachusetts: MIT Press. 1997. P. 135-144.

4. М. Эйген *Самоорганизация материи и эволюция биологических макромолекул*. – М.: Мир, 1973.

5. М. Эйген, П. Шустер *Гиперцикл. Принципы самоорганизации макромолекул*. – М.: Мир, 1982.

6. В.Г. Редько *Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики*. – М.: УРСС. 2005.

7. В.Г. Редько, Ю.Р. Цой Оценка эффективности эволюционных алгоритмов // *Доклады АН*. – Т. 404. № 3. 2005. С. 312-315.

8. В.Г. Редько, Ю.Р. Цой Оценка скорости и эффективности эволюционных алгоритмов // В: *Бионические информационные системы и их практические применения*. Под ред. Зинченко Л.А., Курейчика В.М., Редько В.Г. М.: Физматлит, 2011. С. 109-126.

9. М. Кимура *Молекулярная эволюция: теория нейтральности*. – М.: Мир, 1985.

The model of formation of the simple form of imprinting in the process of learning and evolution of autonomous agents

V.G. Red'ko, O.V. Red'ko

Abstract. The computer model of the origin a simple form of imprinting in the process of learning and evolution of autonomous agents is constructed and investigated. Imprinting is a well-known biological phenomenon of impression in memory of animals of certain objects. The current paper models the formation of the simple form of imprinting: imprinting the image of the mother by newborn chicks and subsequent use of this image. The formation of this kind of imprinting is demonstrated by means of computer simulation. The mechanisms of interaction between learning and evolutionary optimization of autonomous agents in the process of formation of imprinting are analyzed. The model is constructed on the basis of previous formal model of generation of simple skills of autonomous agents during their learning and evolution.