



**Ордена Ленина
ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
имени М. В. Келдыша
Российской академии наук**

Бурцев М. С.

**Модель эволюционного
возникновения целенаправленного
адаптивного поведения.
2. Исследование развития иерархии
целей**

Препринт №

Москва

Ордена Ленина ИНСТИТУТ ПРИКЛАДНОЙ МАТЕМАТИКИ
им. М.В. Келдыша
Российской академии наук

Бурцев М. С.

**Модель эволюционного возникновения целенаправленного
адаптивного поведения.
2. Исследование развития иерархии целей**

Москва
2002

АННОТАЦИЯ

Важным свойством адаптивного поведения является целенаправленность. Целенаправленность проявляется на различных уровнях организации живой материи, можно выделить цели на уровне вида, популяции, отдельного организма. В этой иерархии цели нижних уровней непосредственно связаны с взаимодействием организма со средой в текущий момент, а цели высших уровней определяют его поведение в перспективе. Можно ли смоделировать возникновение подобной иерархии целей? Каковы механизмы целенаправленного поведения? Какую роль в нем играют мотивации? Эта работа является попыткой разобраться в этих вопросах.

ABSTRACT

The main objective of this investigation is a study of mechanisms of simulated evolution of primitive agents in simple environment. There is an artificial world inhabited by some population of agents. Agents can receive information from their external and internal environment and perform some actions. Behavior of every agent is governed by modular artificial neural network. Weights of this network are coded in agent's genome. The population evolves through modifications of agents' genomes. As a result of simulated evolution it has been formed a hierarchy of goals in the population of simple software agents. There also has occurred the differentiation in the behavior repertoire of the population.

Введение

Одним из современных направлений компьютерного моделирования, позволяющим изучать сложные системы является "Искусственная жизнь" (Artificial Life) [1]. В рамках "Искусственной жизни" проводятся исследования мультиагентных эволюционных систем с использованием таких инструментов, как искусственные нейронные сети, эволюционные алгоритмы, нечеткие логики и т.д. Это позволяет создавать интересные модели биологических [2], робототехнических [3] и социальных [4] систем.

Данная работа является развитием работы [5], в которой методами "Искусственной жизни" изучалось влияние мотивации на адаптивное поведение. Ниже приведено краткое описание предыдущей модели и основные результаты экспериментов с ней.

В предыдущей работе рассматривалась популяция агентов, эволюционирующая во времени. Каждый агент мог воспринимать изменения окружающей его среды и совершать некоторые действия. Управление агентом (выбор действий) осуществлялось его нейронной сетью. Параметры нейросети (веса синапсов) модифицировались в процессе эволюции популяции агентов. Геном агента определял веса синапсов нейронной сети.

Каждый агент имел внутренний энергетический ресурс.

В клетках с некоторой вероятностью появлялась трава. Если агент находился в клетке, где растет трава, агент мог съесть траву в этой клетке, тем самым, пополнив свой энергетический ресурс.

Любой агент видел ситуацию только в трех клетках: в той, которой находился сам агент, и в двух соседних клетках.

Агенты могли выполнять следующие действия: 1) отдых, 2) питание, 3) движение на одну клетку вправо или влево, 4) прыжок (на несколько клеток в случайном направлении), 5) скрещивание.

Если два агента, находящиеся в соседних клетках, вырабатывали действие "скрещиваться", то становилось возможным появление нового агента – "потомка". При появлении нового агента осуществлялась равномерная рекомбинация генов родителей, затем геном потомка подвергается малым мутациям.

Выполнение какого-либо действия приводило к изменению внутреннего энергетического ресурса агента. Если энергетический ресурс агента уменьшался до нуля, агент погибал.

У агентов были заданы два мотивационных параметра, соответствующие потребностям:

1) мотивация к поиску пищи, которая соответствует пищевой потребности (потребности пополнения ресурса энергии);

2) мотивация к скрещиванию, которая соответствует половой потребности.

Мотивация к поиску пищи уменьшалась с ростом энергетического ресурса агента, а мотивация к скрещиванию увеличивалась.

Поведение агентов исследовалось с помощью компьютерного моделирования.

Целью работы [5], было исследование влияния мотиваций на поведение отдельного агента и популяции в целом. Для этого проводилось моделирование эволюции популяции агентов с "работающими" мотивациями и с "выключенными" мотивациями (сигналы на мотивационных входах не меняются в течение "жизни" агента).

В каждом эксперименте начальная популяция состояла из агентов, которые имели один и тот же набор синаптических весов, задававших некие "начальные инстинкты". В частности, мы требовали, чтобы агенты исходной популяции имели инстинкты: 1) пополнения ресурса и 2) размножения.

Результаты экспериментов показали, что популяция агентов, имеющих мотивации, лучше приспосабливается к окружающей среде, чем популяция агентов без мотиваций.

Анализ результатов показал, что наличие мотивационных входов позволяет эволюции сформировать иерархию в системе управления поведением агента. В отсутствие мотиваций в дополнение к двум заложенным нами программам поведения ("ВИЖУ ЕДУ" -> "ЕСТЬ" и "ВИЖУ СОСЕДА" -> "СКРЕЩИВАТЬСЯ") вырабатывается третья - "НЕ ВИЖУ НИЧЕГО" -> "ОТДЫХАТЬ" (рис. 1).



Рис. 1. Схема управления агента без мотиваций.

Возможность получения информации о мотивациях позволяет эволюционному процессу найти более сложную стратегию поведения.

В дополнение к трем программам, рассмотренным выше, появляется следующий уровень управления – метапрограмма, управляющая выбором выработанных программ в зависимости от мотиваций.

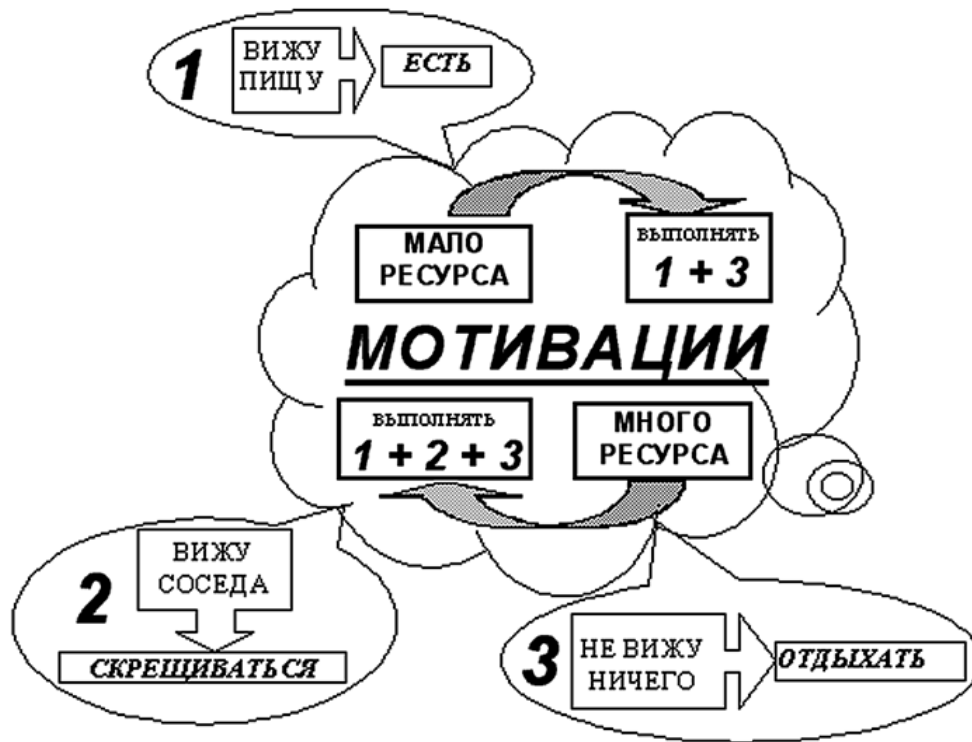


Рис. 2. Схема управления агента с мотивациями.

Возникшая система управления является иерархической – при помощи мотиваций формируется управление более высокого уровня, управляющее программами, направленными на удовлетворение потребностей пополнения ресурса энергии и размножения. В случае малого ресурса мотивация к пополнению ресурса тормозит потребность в скрещивании, выключая программу 2. В случае большого ресурса включается мотивация к размножению, и общую стратегию можно представить следующим образом (рис. 2).

Переход от схемы управления без мотиваций (рис. 1) к схеме с мотивациями (рис. 2) можно рассматривать, как метасистемный переход, в соответствии с теорией, предложенной В. Ф. Турчиным [6].

Итак, результаты предыдущих экспериментов показали, что поведение агента, управляемого иерархической системой, состоящей из двух уровней – уровня "простых инстинктов" и метауровня, обусловленного мотивациями, более эффективно, чем поведение, определяемое одними лишь простыми инстинктами. Мотивации формируют действия агента в соответствии с естественными целями – целями удовлетворения потребностей питания и размножения.

Для того, чтобы исследовать эволюцию более общего и более сложного поведения модель была модифицирована. Мир, в котором "живут" агенты, был сделан двумерным. Скрещивание было заменено на деление. Добавились новые действия, позволяющие агентам бороться друг с другом. Описание модифицированной модели и результаты экспериментов с ней приведены ниже.

Модель целенаправленного адаптивного поведения

Модель представляет собой искусственный мир в двухмерном замкнутом пространстве (тор). Пространство разбито на клетки, в которых могут находиться агенты и их пища. В одной клетке не может быть больше одного агента. Время дискретно, $t = 0, 1, 2, \dots$. В каждый такт времени каждый агент совершает строго одно действие. Эти действия таковы:

1. есть;
2. двигаться на одну клетку вперед;
3. повернуть налево;
4. повернуть направо;
5. отдыхать;
6. делиться;
7. ударить агента в клетке спереди (напасть);
8. защищаться.

В клетках с некоторой вероятностью появляется пища, потребляемая агентами. Агенты могут получать информацию о состоянии окружающего их мира и своем внутреннем состоянии.

Каждый агент имеет внутренний энергетический ресурс. Ресурс пополняется при питании и расходуется при выполнении действий.

Одно из действий, которое совершаемых агентами – деление. При делении потомок наследует систему управления родителя измененную мутациями и часть его энергии. Если в течение жизни агента его энергетический ресурс уменьшается до 0, то агент погибает.

Агент ориентирован. Он имеет направление «вперед», относительно которого ориентировано его поле зрения и действия совершаемые им. Поле зрения агента состоит из 4-х клеток расположенных – перед, справа, слева и под агентом (рис. 3). Агент может видеть пищу и других агентов, расположенных в поле его зрения.

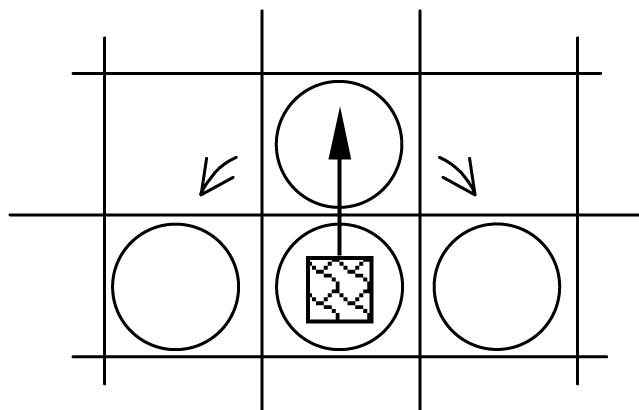


Рис.3. Поле зрения агента (Стрелкой показано направление «вперед», окружностями отмечены узлы, из которых агент может получать информацию, квадратом обозначен узел, в котором находится агент).

Опишем физиологию агентов, характеризуя взаимосвязи между действиями агента и его внутренним состоянием.

Используем следующие обозначения: R – энергетический ресурс, ΔR – изменение ресурса за один такт времени; k_i – параметры, характеризующие изменение внутреннего энергетического ресурса. Емкость внутреннего энергетического ресурса агента ограничена, она не может быть больше R_{max} .

Выполнение какого-либо действия приводит к изменению внутреннего ресурса агента по следующим правилам:

- 1) Отдых (нахождение в состоянии покоя): $\Delta R = -k_1$.
- 2) Движение: $\Delta R = -k_2$.
- 3) Поворот (на право или на лево): $\Delta R = -k_3$.
- 4) Питание: $\Delta R = -k_4 + k_5$.
- 5) Действие деления: $\Delta R = -k_6$.

При появлении потомка родитель передает ему часть своей энергии: $\Delta R_p = -0.5R_p$, здесь индекс p обозначает родителя. Начальный ресурс потомка R_o равен полученной от родителя энергии $R_o = 0.5R_p$.

6) При нападении на другого агента: $\Delta R = 0.5k_7$, если жертва не защищается, и $\Delta R = -k_7$ в противном случае.

- 7) При защите от другого агента: $\Delta R = -k_8$.

Система управления агента состоит из однослойной модульной нейронной сети, состоящей из рецепторных (входных) и эффекторных (выходных) модулей (рис. 4). Набор модулей, а так же параметры связи между ними определяются геномом агента.

Нейронная сеть агента имеет следующие входы:

- 1) сигналы о наличии пищи в поле зрения агента (рядом, впереди, слева, справа);
- 2) сигналы о наличии другого агента в поле зрения (впереди, слева, справа);
- 3) текущее значение внутреннего энергетического ресурса;
- 4) разность между максимальным и текущим значениями внутреннего энергетического ресурса;
- 5) изменение энергетического ресурса по сравнению с предыдущим тактом времени;
- 6) вспомогательный постоянный входной сигнал.

Значения, поступающие на входы нейронной сети, нормированы к R_{max} .

Значения на выходах нейронной сети y_j вычисляется умножением вектора входных сигналов на матрицу весов связей входов с выходами, где входы соответствуют рецепторам, а выходы – эффекторам:

$$y_j = \sum_i w_{ij} x_i,$$

где x_i – входы j -го нейрона, а w_{ij} – его синаптические веса.

Действие, соответствующее выходу, имеющему максимальное значение, выполняется агентом.

Веса нейронной сети агента в течение жизни не изменяются.

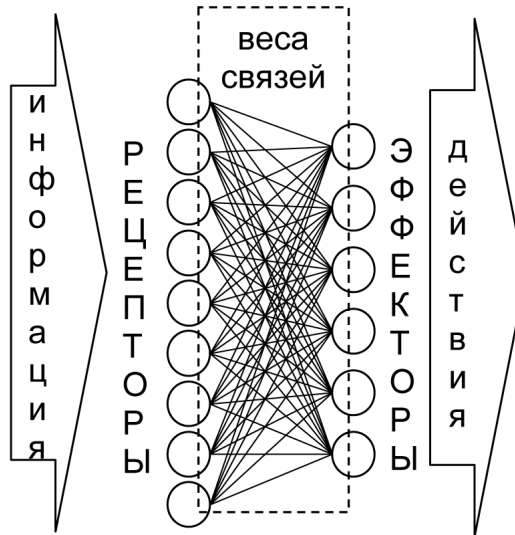


Рис. 4. Схема полной структуры системы управления (нейронной сети) агента.

Популяция агентов эволюционирует во времени за счет изменения структуры и весов синапсов управляющей нейронной сети.

Геном агента S состоит из двух хромосом $S = (W, M)$. Первая хромосома содержит веса синапсов нейронной сети, представленные целыми числами. Наличие или отсутствие модуля в структуре сети определяется значением соответствующего двоичного числа во второй хромосоме.

Изменение генома происходит от родителя к потомку. В результате выработки агентом действия «делиться» появляется потомок. Геном потомка задается при помощи следующего генетического алгоритма:

1. добавить к каждому гену W_i хромосомы, определяющей вес связи, случайную величину x , равномерно распределенную на интервале $[-p_m, p_m]$;
2. изменить число M_j , определяющее наличие того или иного модуля, на противоположное с малой вероятностью p_s .

Моделирование

Модель была реализована в виде компьютерной программы.

Цель компьютерных экспериментов заключалась в том, чтобы показать возможность возникновения иерархии целей у агентов в процессе искусственной эволюции. Так как в нашей модели поведение отдельного агента не изменялось за время его существования, то адаптация к среде проходила на уровне популяции. Естественно рассматривать основную цель популяции как цель выживания в данной среде. Для отдельного агента она будет являться *метацелью*, определяющей цели его существования – получения энергии и размножения. Но и эти цели в свою очередь могут быть разбиты на *подцели*, которые позволяют оптимизировать процесс достижения целей высших уровней. Иерархия целей, возникающая в нашей модели, схематично изображена на рисунке 5.

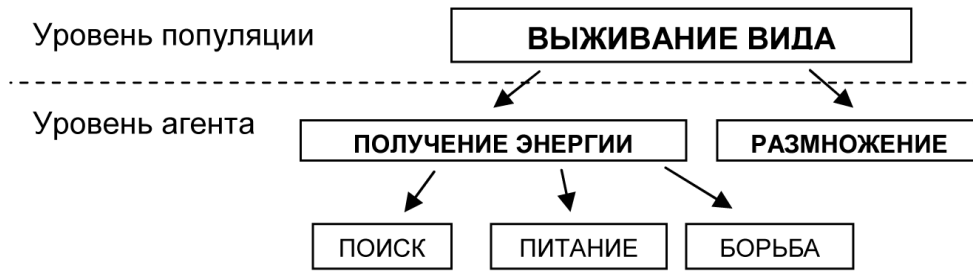


Рис. 5. Иерархия целей, возникающая у агентов в процессе эксперимента

В начале каждого эксперимента мир заселялся популяцией агентов, имеющих минимальный набор рецепторов и эффекторов. Агент начальной популяции мог видеть пищу в той клетке, где он находился, и в клетке находящейся перед ним. А также совершать следующие действия – есть, двигаться и делиться. Веса синапсов были заданы таким образом, чтобы обеспечить агенту два начальных инстинкта – питания и размножения. Если агент видел пищу рядом с собой (в своей клетке), то он должен был выполнять действие «есть», если в соседней клетке, то двигаться в соседнюю клетку; если он не видел ничего, то вызывалось действие «делиться». Очевидно, что агенты начальной популяции обладали лишь примитивной стратегией.

Был проведен ряд компьютерных экспериментов. Результаты одного эксперимента изложены ниже (параметры эксперимента приведены в приложении 1).

На рисунке 6 приведена динамика количества агентов, выполняющих то или иное действие. Этот график позволяет нам провести предварительный анализ эволюции поведения в модели. Так увеличение числа агентов, выполняющих действие "отдыхать", в период $t = 12 \cdot 10^6 - 18 \cdot 10^6$ приводит к росту общей численности популяции. Это позволяет нам предположить, что эволюция нашла квазиоптимальную стратегию, связанную с действием "отдыхать". После $t = 18 \cdot 10^6$ численность популяции начинает снижаться, мы можем предположить, что это связано с взаимодействиями между агентами. При увеличении числа агентов, растет вероятность встречи двух агентов, и, следовательно, агенты, нападающие на соседей, быстрее пополняют свою энергию. Это в свою очередь приводит к увеличению их числа. Количество агрессивных агентов растет, и к моменту $t = 20 \cdot 10^6$ они практически уничтожают "отдыхающих" агентов.

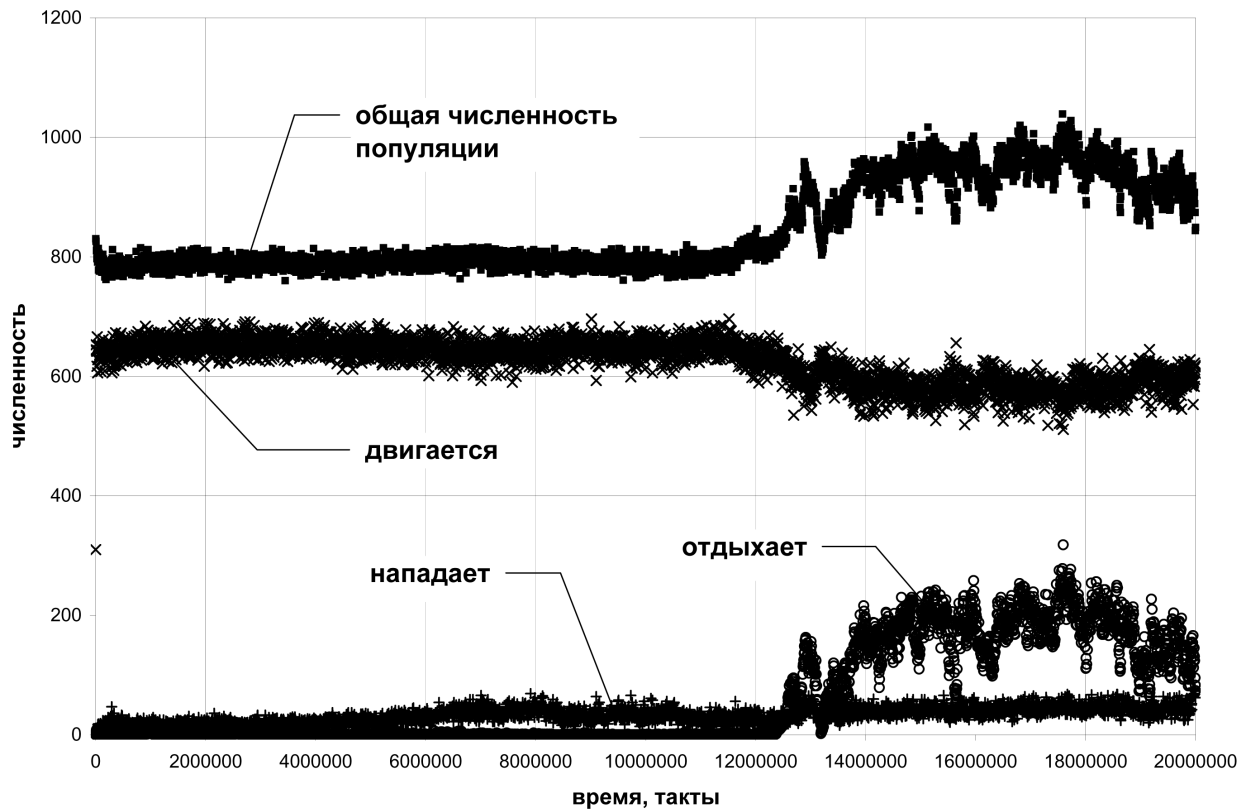


Рис.6. Динамика численности популяции и некоторых действий, выполняемых агентами.

Для более детального анализа воспользуемся картой генома популяции, которая отображает значения весов в системе управления для каждого агента популяции в заданный момент времени (рис. 7). Такая карта позволяет определить различные стратегии поведения присутствующие в популяции. Агентов, имеющих сходные веса, будем объединять в виды.

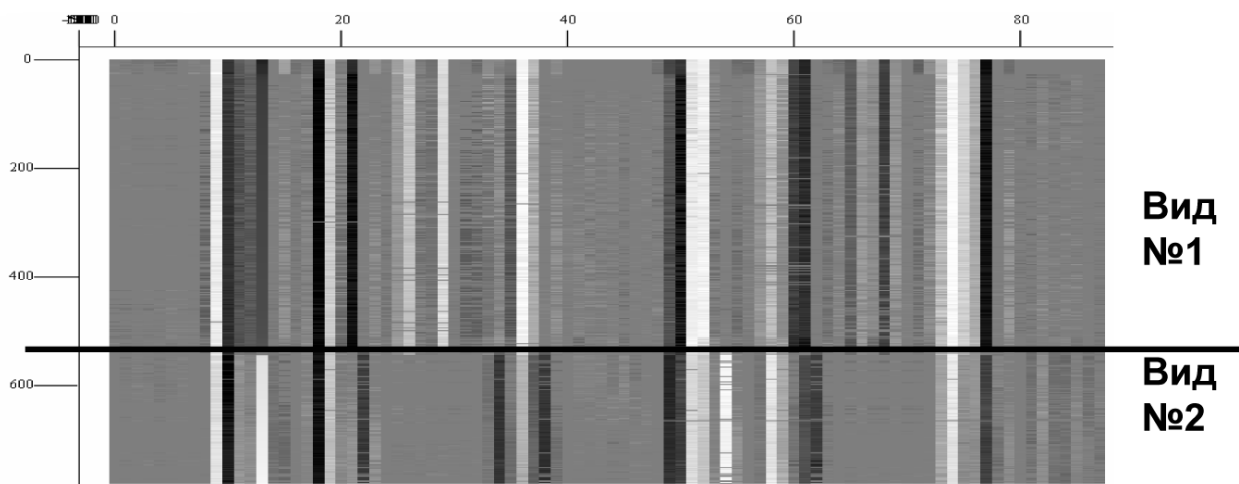


Рис.7. Карта генома популяции: каждая горизонтальная линия представляет собой набор весов нейронной сети отдельного агента, белый цвет соответствует максимальным значениям весов, а черный минимальным.

Имея значения весов нейронных сетей агентов в популяции для различных моментов времени, мы можем выделить виды и построить их динамику в популяции (рис. 8).

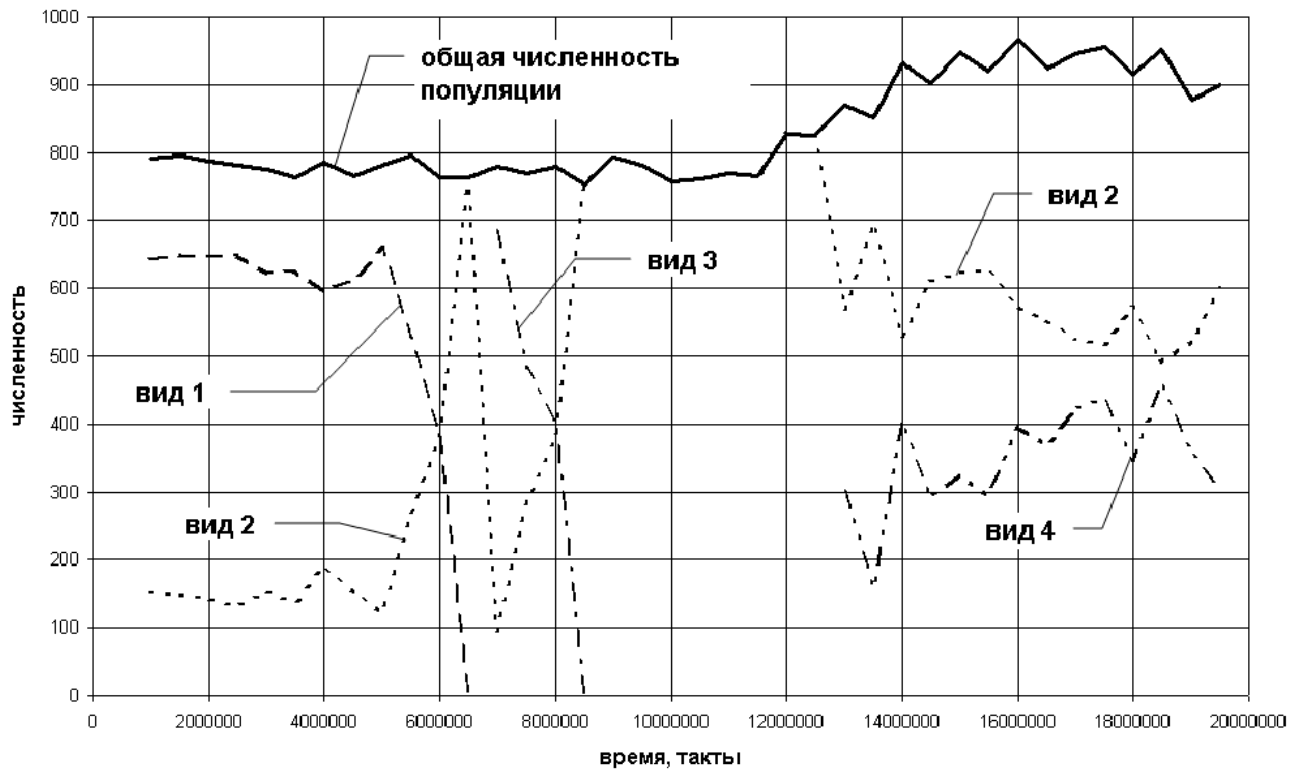


Рис.8. Динамика видов в популяции.

Рассмотрим динамику видов в популяции. На интервале $t = 1 \cdot 10^6 - 6 \cdot 10^6$ в популяции существует два вида агентов. Анализируя веса для каждого из видов можно определить типы стратегий. Агенты, относящиеся к виду №1, обладают "мирным" поведением, они не нападают на других агентов. Единственным источником энергии для них является пища, случайно появляющаяся в клетках. Агенты вида №2 "всеядны", они могут есть пищу, а также нападать на других агентов. Отметим, что в модели наблюдается колебание отношения этих двух видов. Подобные колебания присутствуют в природных экосистемах, в которых существуют отношения "хищник - жертва". Для описания такой динамики математическая экология использует аппарат дифференциальных уравнений (модель Лотки-Вольтерра), в исследуемой модели появление колебаний является следствием эволюционной самоорганизации.

При $t = 6 \cdot 10^6 - 8 \cdot 10^6$ вид №2 последовательно вытесняет сначала вид №1, а затем промежуточный вид №3. Затем примерно до $12 \cdot 10^6$ такта в популяции доминирует вид №2. Результаты моделирования показывают, что поведение агентов в каком-либо из видов не остается стабильным, а постепенно эволюционирует. Так после $8,5 \cdot 10^6$ такта агрессивность агентов, относящихся к виду №2, падает и постепенно начинается формирование вида №4. К моменту времени $t = 12 \cdot 10^6$ мы уже можем четко выделить вид №4. Агенты вида №4 обладают агрессивной стратегией и способностью экономить энергию, выполняя

действие "отдыхать". В период $t = 12 \cdot 10^6 - 20 \cdot 10^6$ снова наблюдаются колебания, только теперь в качестве "жертв" выступают агенты вида №2, а "хищников" вида №4.

Рассмотрим, как происходит формирование иерархии целей в экспериментах с моделью.

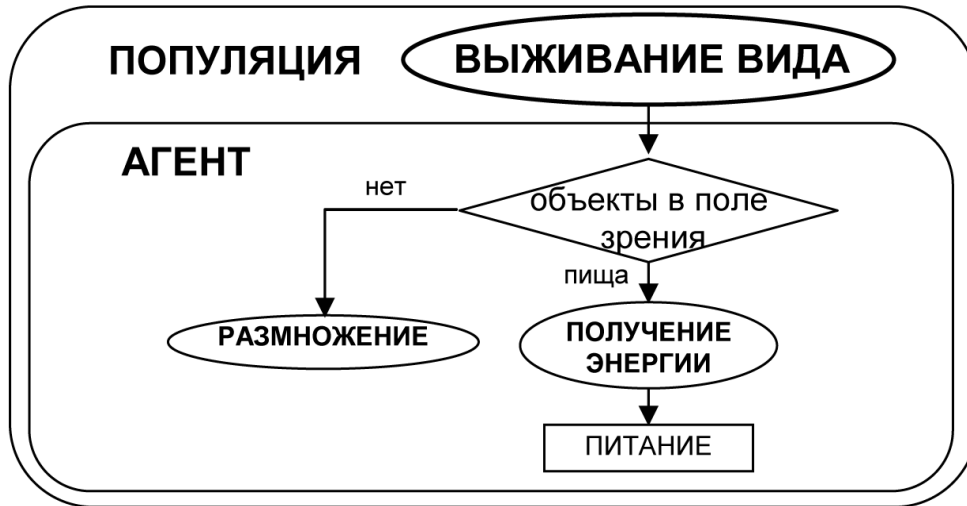


Рис.9. Дерево условий для управления выбором подцелей агента начальной популяции.

Поведение агентов начальной популяции (имеющих минимальный набор рецепторов и эффекторов) можно схематично представить в виде блок-схемы (рис. 9). Агенты начальной популяции обладают примитивной стратегией, в которой учитывается только наличие пищи в поле зрения.

В ходе эволюции поведение агентов усложняется. Стратегия агентов на $5 \cdot 10^6$ такте (вид №2) может быть представлена в виде схемы, показанной на рис. 10 (см. также таблицу приложения 2). Видно, что из примитивной стратегии, заданной нами для агентов начальной популяции (рис. 9), развивается достаточно сложное поведение, которое можно назвать целенаправленным. Так первоначальный «инстинкт» агента, направленный на получение энергии оптимизируется за счет появления еще одного уровня подцелей, направленных, соответственно: на само питание, на поиск пищи, борьбу. Для эффективного управления поведением, имеющим подобную многоуровневую иерархию целей, необходимо иметь информацию о том, какие цели являются предпочтительными для агента в данный момент времени. Для этого агенту необходимо знать не только состояние окружающей его среды, но и своего энергетического ресурса. В модели эта информация может быть получена от входов, связанных со значением и изменением ресурса за последний такт. Значения на этих входах могут быть интерпретированы как мотивации к выбору того или иного типа поведения. Так, большое значение внутреннего ресурса может рассматриваться как мотивация к размножению. Если ресурса много, то отсутствие раздражителей во внешней среде приводит к делению агента, а если мало, то к поиску пищи. Изменение внутреннего ресурса используется агентами как

мотивация к нападению на впереди стоящего агента. Если эта мотивация положительна, т.е. предыдущее действие привело к увеличению энергии, то агент нападает на соседа. Если данная мотивация отрицательна, то атака прекращается.

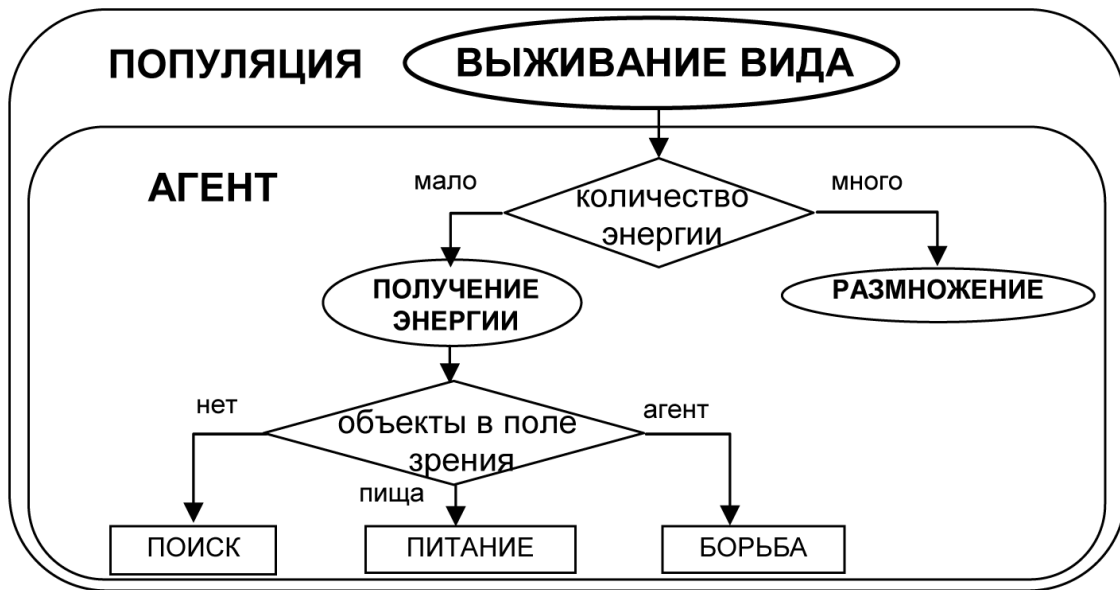


Рис.10. Дерево условий для управления выбором подцелей

Заключение

Проведенное моделирование показывает, что в процессе искусственной эволюции в популяции агентов вырабатывается поведение, которое можно рассматривать как поведение, управляемое в соответствии с некоторой иерархией целей. Причем эта иерархия постепенно усложняется в процессе эволюции. Формирование иерархии целей приводит к потребности в механизме выбора текущих подцелей. В качестве такого механизма выступают мотивации.

Результаты экспериментов также демонстрируют, что в рассматриваемой модели возможно возникновение и одновременное существование агентов использующих различные стратегии поведения, притом, что изначально популяция однородна. Можно интерпретировать это как возникновение и взаимодействие видов.

Текущая модель позволяет исследовать влияние неоднородности распределения пищи на поведение и видообразование в популяции. Так предварительные результаты показывают, что неоднородное распределение приводит к увеличению числа подвидов. Также при помощи модели можно рассмотреть вопросы, связанные с эволюционным обучением и сохранением знания на уровне популяции в нестационарных условиях (при изменении количества пищи во времени).

На следующем этапе исследований в модель планируется ввести обучение агента в течение жизни и возможность усложнения нейронной сети агентов в процессе эволюции. Эти усовершенствования позволят изучить зависимость эволюционных процессов от обучения, а также дадут возможность возникновения сложных иерархий в поведении.

Благодарности

Автор благодарен В.Г. Редько за ряд полезных консультаций.

Список литературы

1. Artificial Life // Langton C., The Proceedings of an Interdisciplinary Workshop on the Synthesis and Simulation of Living Systems, Redwood City CA: Addison-Wesley, 1989.
2. Artificial Life II // Langton C., Taylor C., Farmer D., and Rasmussen S., Artificial Life II, Santa Fe Institute Studies in the Sciences of Complexity, X, Reading, MA: Addison-Wesley, 1992.
3. Nolfi S., Parisi D. Learning to adapt to changing environments in evolving neural networks // Adaptive Behavior, 5, 1, 1997, pp. 75-98
[<http://gral.ip.rm.cnr.it/nolfi/papers/nolfi.changing.pdf>]
4. Tesfatsion L. How Economists Can Get Alive // W.Brian Arthur, Steven Durlauf, and David Lane, Santa Fe Institute Studies in the Sciences of Complexity, XXVII, Redwood City CA: Addison-Wesley, 1997.
[<http://www.econ.iastate.edu/tesfatsi/surveys.htm>]
5. Бурцев М.С., Гусарев Р.В., Редько В.Г. Модель эволюционного возникновения целенаправленного адаптивного поведения 1. Случай двух потребностей. М.: ИПМ РАН, (2000).
[<http://www.keldysh.ru/pages/BioCyber/PrPrint/PrPrint.htm>]
6. Турчин В.Ф. Феномен науки. Кибернетический подход к эволюции. М.: ЭТС, (2000). [<http://refal.net/turchin/phenomenon>]

Приложение 1

Параметры эксперимента

Размер мира: 100 на 100 клеток

Численность начальной популяции: 1000

Вероятность возникновения пищи в клетке: $3 \cdot 10^{-4}$

Максимальное значение веса нейронной сети: 1000

Максимальное изменение веса при мутации: $p_m = 20$

Вероятность мутации модуля: $p_s = 0,01$

Максимальное значение внутреннего энергетического ресурса: $R_{max} = 5000$

Энергетическая ценность пищи: $k_5 = 500$

Энергетические затраты на деление: $k_6 = 20$

Энергетические затраты на движение: $k_2 = 20$

Энергетические затраты на питание: $k_4 = 20$

Энергетические затраты на поворот: $k_3 = 10$

Энергетические затраты на отдых: $k_1 = 5$

Энергетические затраты на нападение: $k_7 = 30$

Энергетические затраты на защиту: $k_8 = 20$

Приложение 2

Таблица. Стратегии агентов на $5 \cdot 10^6$ такте

	Вид 1 (82% популяции)	Вид 2 (18% популяции)
Мало энергии ($R = 0.02 R_{max}$)		
Ничего не видно	двигаться вперед	двигаться вперед
Еда рядом	есть	есть
Еда впереди	двигаться вперед	двигаться вперед
Агент впереди	поворачивать налево	поворачивать налево
Еда слева/справа	поворачивать налево/направо	поворачивать налево/направо
Агент слева/справа	двигаться вперед	двигаться вперед
Среднее количество энергии ($R = 0.05 R_{max}$)		
Ничего не видно	двигаться вперед	двигаться вперед
Еда рядом	есть	есть
Еда впереди	двигаться вперед	двигаться вперед
Агент впереди	поворачивать налево	ударить/поворачивать направо
Еда слева/справа	поворачивать налево/направо	поворачивать налево/направо
Агент слева/справа	двигаться вперед	двигаться вперед
Много энергии ($R = 0.96 R_{max}$)		
Ничего не видно	двигаться вперед	двигаться вперед
Еда рядом	есть	делиться
Еда впереди	делиться	двигаться вперед
Агент впереди	поворачивать налево	ударить
Еда слева/справа	поворачивать налево/направо	поворачивать налево/направо
Агент слева/справа	двигаться вперед	двигаться вперед