

М. С. БУРЦЕВ

Институт прикладной математики им. М. В. Келдыша РАН, Москва.
mbur@narod.ru

ФОРМИРОВАНИЕ ИЕРАРХИИ ЦЕЛЕЙ В МОДЕЛИ ИСКУССТВЕННОЙ ЭВОЛЮЦИИ

Аннотация

Изучены возможные процессы формирования целенаправленного адаптивного поведения. Моделируется эволюция популяции простейших адаптивных систем в искусственном мире. Полученные результаты демонстрируют возникновение иерархии целей в системе управления агента.

Abstract

The possible processes of purposeful adaptive behavior formation were studied. An evolutionary dynamics of simple control systems in an artificial world was simulated. The results show emergence of goals hierarchy in the control system of an agent.

Введение

Важным свойством адаптивного поведения является целенаправленность. Целенаправленность проявляется на различных уровнях организации живой материи, можно выделить цели вида, популяции, отдельного организма. В этой иерархии, цели нижних уровней непосредственно связаны с взаимодействием организма и среды в текущий момент, а цели высших уровней определяют 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.

Описание модели

Рассматриваемая модель является развитием работы [1], в которой изучалось влияние мотивации на адаптивное поведение. При построении модели использовались методы «Искусственной жизни» [2-4].

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. В каждый такт времени агент должен совершить какое-либо действие. Через определенные промежутки времени в клетках с некоторой вероятностью появляется пища, потребляемая агентами. Каждый агент имеет внутренний энергетический ресурс, который пополняется при питании и расходуется на выполнение действий. Если энергетический ресурс агента уменьшается до 0, то агент погибает. Агенты могут получать информацию о состоянии окружающего их мира и своем внутреннем состоянии, совершать определенные действия. Популяция агентов эволюционирует во времени за счет изменения структуры и весов управляющей нейронной сети.

Агент имеет направление («вперед»), относительно которого ориентировано его поле зрения и действия совершаемые им. Поле зрения состоит из 4-х клеток расположенных – впереди, справа, слева и под агентом (рис. 1).

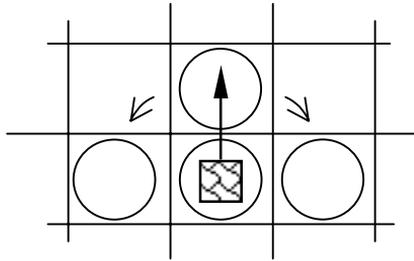


Рис. 1. Ориентация поля зрения относительно направления «вперед» агента (окружностями отмечены клетки, из которых агент может получать информацию, квадратом обозначен агент, стрелочкой – направление «вперед»).

Система управления агента состоит из однослойной модульной нейронной сети, состоящей из рецепторных (входных) и эффекторных (выходных) модулей (см. рис. 2). Набор модулей, а так же параметры связи между ними определяются геномом агента.

Агент может иметь следующие входы нейронной сети: 1) наличие пищи рядом (в клетке, в которой расположен сам агент), 2) наличие пищи слева, 3) наличие пищи впереди 4) наличие пищи справа, 5) наличие другого агента слева, 6) наличие другого агента впереди 7) наличие другого агента справа, 8) состояние внутреннего энергетического ресурса, 9) изменение ресурса за последний такт времени.

Каждому входу нейронной сети соответствует один рецептор.

В каждый такт времени агент совершает одно из следующих действий: 1) отдыхать, 2) есть, 2) двигаться (вперед), 3) поворачиваться влево, 4) поворачиваться вправо, 5) ударять (другого агента), делиться.

Каждому действию соответствует один эффектор (в нашем случае – один нейрон). Нейронная сеть осуществляет связи между рецепторами и эффекторами.

Значения на выходах нейронной сети y_j вычисляются 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.

Агент выполняет действие, соответствующее нейрону с наибольшим выходным сигналом y_j . Веса матрицы связей определяются геномом агента (рис. 2а).

Количество модулей (рецепторов и эффекторов) нейронной сети может изменяться в процессе эволюции.

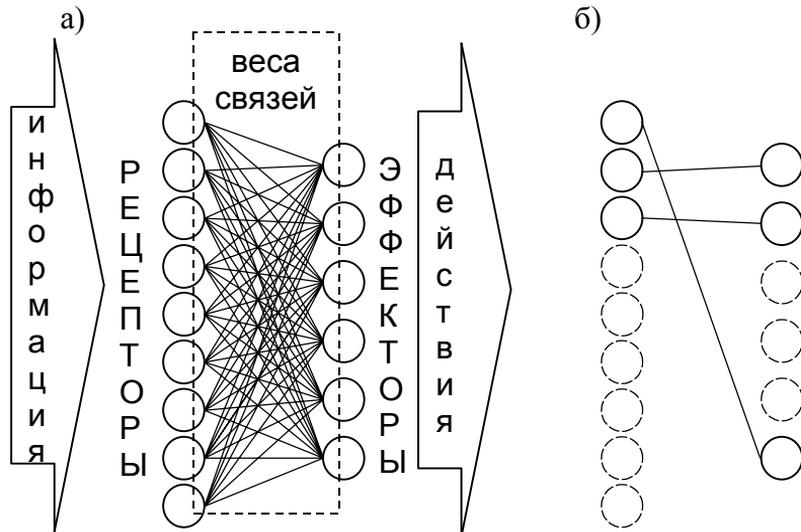


Рис. 2. Схема полной структуры системы управления (нейронной сети) агента (а), и агента начальной популяции (б), пунктирными окружностями обозначены отсутствующие модули.

Геном агента \mathbf{S} состоит из двух хромосом $\mathbf{S} = (\mathbf{W}, \mathbf{M})$. Первая хромосома \mathbf{W} содержит веса синапсов нейронной сети w_{ij} . Вторая хромосома \mathbf{M} определяет структуру нейронной сети: наличие или отсутствие того или иного модуля в структуре сети определяется значением соответствующей булевой переменной.

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 программы на языке Java, с использованием которой был проведен ряд экспериментов.

В начале каждого эксперимента мир заселялся популяцией агентов, имеющих минимальный набор рецепторов и эффекторов (рис. 2б). Такая структура нейронной сети позволяла агенту начальной популяции видеть пищу в той клетке, где он находился, и в клетке, находящейся впереди. Агенты могли 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, то должен был выполнять действие «есть», если в соседней клетке, то двигаться в соседнюю клетку; если не видел ничего, то вызывалось действие «делиться». Очевидно, что агенты начальной популяции обладали лишь примитивной стратегией.

Изменение энергетического ресурса агента при совершении действий было задано следующим образом. При отдыхе агент затрачивал наименьшее количество энергии, при повороте направо/налево в два раза больше, при движении, питании и делении еще в два раза больше, причем после деления половина ресурса родителя передавалась потомку. Если агент совершал действие «есть», и в его клетке была еда, то его энергетический ресурс увеличивался. Агент мог совершить действие «ударить» при этом, если перед агентом находился сосед, то у последнего отнималась часть энергии и передавалась нападающему.

Цель экспериментов заключалась в том, чтобы показать возможность возникновения иерархии целей у агентов в процессе искусственной эволюции, своего рода *метасистемный* переход [5]. 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 – выживания 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, показана на рисунке 3.

Как происходит формирование иерархии целей при экспериментах с моделью? Сначала заметим, что структура целей, обеспечивающая наиболее эффективное поведение агентов, определяется внешней по отношению к ним средой, а сама эволюционирующая популяция представляет собой стохастический генератор стратегий. Каждая конкретная стратегия определяет поведение отдельного агента. Агенты с удачной стратегией выживают и дают потомство, а плохо приспособленные гибнут.



Рис. 3. Иерархия целей, возникающая у агентов в процессе эксперимента

Стратегию агента можно представить, как набор действий направленный на достижение некоторых целей. Усредняя по всей популяции веса матрицы связей нейронной сети, выработанные в процессе искусственной эволюции, мы можем представить себе, каким образом осуществляется управление поведением агентов в популяции. Для типичных экспериментов это поведение можно описать в виде следующего алгоритма:

```

функция ЧтоМнеДелать? {
  если (какие-либо объекты в поле зрения отсутствуют)
    если (мало энергии) двигаться;
    иначе размножаться;
  если (еда рядом) есть;
  если (еда впереди) двигаться;
  если (еда справа/слева) поворачивать направо/налево;
  если (агент впереди) ударить;
  если ((агент впереди) И (еда впереди /справа/слева))
    если (энергия увеличивалась) ударить;
    иначе повернуть направо/налево;
},

```

или представить в виде блок-схемы (рис. 4).

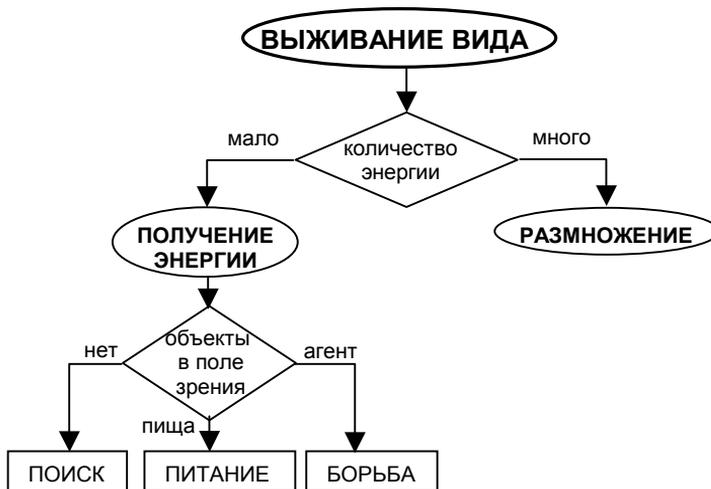


Рис. 4. Дерево условий для управления выбором подцелей

Результаты эксперимента показывают, что из примитивной стратегии заданной нами для агентов начальной популяции, развивается достаточно сложное поведение, которое можно назвать целенаправленным. Так первоначальный «инстинкт» агента, направленный на получение энергии, (напомним, агент ест, когда пища рядом, и двигается, когда пища впереди) оптимизируется за счет появления еще одного уровня подцелей направленных соответственно: на само питание, на поиск пищи, борьбу.

Для эффективного управления поведением, имеющем 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 нашей 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 [6,7] к выбору того, или иного типа поведения. Так значение внутреннего ресурса является мотивацией 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 мотивация положительна, т.е. предыдущее действие привело к увеличению энергии, то агент нападает на соседа. Если мотивация отрицательна, то атака прекращается.

Заключение

Эксперименты с моделью показывают, 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.

Благодарности

Я хотел бы поблагодарить Владимира Георгиевича Редько за идеи послужившие основой для данной работы и полезные советы во время ее выполнения. Также хотел бы сказать спасибо Ерику Берентаеву за помощь при написании программы.

Список литературы

1. *Бурцев М.С., Гусарев Р.В., Редько В.Г.* Модель эволюционного возникновения целенаправленного адаптивного поведения 1. Случай двух потребностей. // Препринт ИПМ РАН, 2000, N 43. <http://www.keldysh.ru/BioCyber/PrPrint/PrPrint.htm>
2. *Yaeger L.* Computational Genetics, Physiology, Learning, Vision, and Behavior or PolyWord: Life in a New Context. // Langton, C. G. (ed).

- Artificial Life III. Addison-Wesley, 1994, pp. 263-298.
<http://www.beanblossom.in.us/larryy/PolyWorld.html>
3. *Ackley, D. Littman, M.* Interactions between learning and evolution. // Langton, C. G., Taylor, C., Farmer, J. D., and Rasmussen, S. (Eds.) *Artificial Life II*. Reading, MA: Addison-Wesley, 1992. pp.487-509.
 4. *Nolfi S., Parisi D.* Learning to adapt to changing environments in evolving neural networks // *Adaptive Behavior* V.5, N.1, 1997, pp. 75-98.
<http://kant.irmkant.rm.cnr.it/nolfipub.html>
 5. *Турчин В.Ф.* Феномен науки. Кибернетический подход к эволюции. // М.: ЭТС, 2000, 368с. (2-е изд).
 6. *Анохин П.К.* Системные механизмы высшей нервной деятельности. // М.: Наука, 1979, 453 с.
 7. *Balkenius C.* The roots of motivations. // In J.-A. Mayer, H. L. Roitblat and S. W. Wilson (eds.), *From Animals to Animats II*, MA: MIT Press., 1993.
<http://www.lucs.lu.se/People/Christian.Balkenius/Abstracts/ROM.html>