

ПРОЕКТ «МОЗГ АНИМАТА»: РАЗРАБОТКА МОДЕЛИ АДАПТИВНОГО ПОВЕДЕНИЯ НА ОСНОВЕ ТЕОРИИ ФУНКЦИОНАЛЬНЫХ СИСТЕМ

К.В. Анохин, М.С. Бурцев, И.Ю. Зарайская, А.О. Лукашев, В.Г. Редько

НИИ НФ им. П.К. Анохина РАН

ИПМ им. М.В. Келдыша РАН

E-mail :redko keldysh.ru

Аннотация. В работе предложена общая схема управления адаптивным поведением анимата (искусственного или естественного организма). Схема предполагает, что у анимата есть несколько естественных потребностей (питания, безопасности, и т.п.). В соответствии с этими потребностями анимат имеет иерархическую структуру целей и подцелей. Схема управления строится на основе функциональных систем, которые содержат блоки нейросетевой ассоциативной памяти.

1. Введение. Цель и основные предположения модели

Цель настоящей работы – предложить проект компьютерной модели управления целостным адаптивным поведением анимата (естественного или искусственного организма). Главная задача нашего подхода – увязать реалистичным образом в единой модели процессы (а) адаптивного поведения, (б) его онтогенетического развития, (в) обучения и (г) эволюции поведения. В качестве концептуальной платформы такого моделирования мы избрали теорию функциональных систем П.К. Анохина [Анохин, 1975]. В этой теории в основу описания поведения, его индивидуального развития, обучения и эволюции положен единый критерий – адаптивный результат действий организма. Отдельная ветвь общей теории функциональных систем – теория системогенеза, рассматривает закономерности формирования функциональных систем в эволюции, индивидуальном развитии и обучении [Анохин, 1975, Судаков, 1997].

Одним из достоинств такого единого подхода является то, что моделирование комплекса интересующих нас процессов может быть разбито на отдельные этапы. Мы воспользовались этим преимуществом и на первом этапе проекта поставили задачу разработать модель поведения анимата с уже сформированной архитектурой управления. При этом мы пока не касаемся вопросов эволюции этой архитектуры (процессов эволюционного системогенеза) и ее формирования в индивидуальном развитии (процессов первичного системогенеза). На настоящем этапе проекта мы включаем в модель лишь некоторые, ограниченные возможности индивидуального обучения животного, заключающиеся в нахождении новых способов достижения адаптивных результатов (процессы вторичного системогенеза).

Мы также пока не привязываем систему управления анимата к реальным нейронным структурам мозга. Мы ее конструируем, синтезируем сами, в духе моделей Искусственной Жизни и Адаптивного Поведения [Редько, 2001]. В качестве элементов нашей конструкции мы используем отдельные функциональные системы. В построении функциональных систем нами применяются нейронные сети. Но важно подчеркнуть, что эти нейронные сети не связаны с реальными нейронами животного, роль нейронных сетей в нашей модели – реализация достаточно гибкой ассоциативной памяти. Эта память используется при обучении анимата. Основное же внимание мы обращаем на архитектуру системы управления и взаимоотношения между различными функциональными системами в этой архитектуре.

Поведение анимата нацелено на достижение полезных для анимата *результатов*. Поведение состоит из *действий*. Для обеспечения достижения результатов анимат делает *прогноз* ожидаемого результата и формирует *акцептор результатов действия* в виде модели параметров ожидаемого результата.

Важную роль в нашей модели играют *цели* поведения анимата. Мы считаем, что у анимата есть *потребности* (например, потребность питания и потребность безопасности). С

потребностями связаны цели поведения, которые количественно можно характеризовать мотивациями. Например, если анимат голоден, то у него есть выраженная мотивация к питанию и есть цель удовлетворить потребность питания. Потребности и цели организованы иерархическим образом. Потребности каждого из уровней вызывают цели и соответствующие функциональные системы более низкого ранга.

Организация иерархии целенаправленного поведения в нашем анимате близка к таковой в проекте «Животное» [Бонгард и др., 1975]. Как и проект «Животное», настоящая работа – именно проект, который предполагается в дальнейшем воплотить в виде компьютерной реализации модели.

Далее мы опишем общую архитектуру системы управления аниматом (раздел 2), базовый элемент системы управления (раздел 3) и работу этого элемента (раздел 4), работу всей системы управления (раздел 5) и обсудим модель в целом (раздел 6).

2. Архитектура системы управления

Мы предполагаем, что система управления аниматом имеет иерархическую архитектуру (рис.1). Базовым элементом системы управления является отдельная функциональная система (ФС). ФС осуществляет выбор действия в соответствии с заданной целью и текущей ситуацией, формирует прогноз результата действия, а также инициирует команду на выполнение действия (если согласно прогнозу оно должно привести к положительному результату).

Рис. 1. Архитектура системы управления аниматом.

Эта архитектура в основном соответствует работам [Анохин, 1975, Бонгард и др., 1975, Жданов, 1999, Максимов и др., 1997, Умрюхин и др., 2000, Швырков, 1978]. Система верхнего уровня соответствует главной цели организма – выживанию. Следующий уровень – целям достижения основных потребностей (питания, размножения, безопасности, накопления знаний). Более низкие уровни системы управления соответствуют тактическим целям поведения. Управление с верхних уровней может передаваться на нижние уровни (от «суперсистем» к «субсистемам») и возвращаться назад. «Суперсистемы» ставят цели «субсистемам». Для определенности в данной работе предполагаем, что каждая ФС (кроме ФС0) подчинена только одной «суперсистеме».

Предполагаем, что система управления аниматом функционирует в дискретном времени: $t = 0, 1, \dots$ Также предполагаем, что каждый такт времени активна только одна ФС.

3. Функциональная система – базовый элемент управления

Схема отдельной ФС представлена на рис. 2. *Задача* ФС такова:

- существует цель, поставленная "суперсистемой", и известна ситуация, заданная входным вектором $X(t)$, характеризующим состояние внешней и внутренней среды анимата,
- осуществить поиск наиболее оптимальных для данной ФС способов достижения цели,
- если, согласно прогнозу, есть действия, обеспечивающие достижение цели, то дать команду на выполнение действия, максимально приближающего к цели,

– и осуществить контроль правильности выполнения действия – соответствия обратной афферентации от достигнутого результата и параметров запланированного результата, заложенных и хранящихся в аппарате акцептора результатов действия.

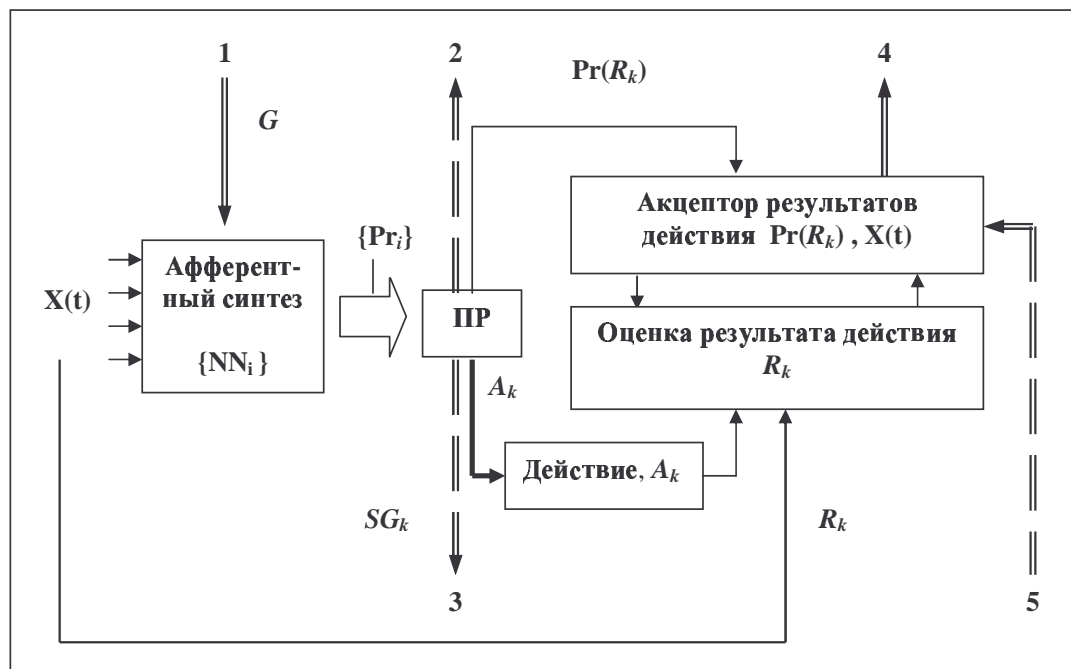


Рис. 2. Схема ФС.

Схема ФС состоит из блока афферентного синтеза, блока принятия решения (ПР), эффекторного устройства, выполняющего действия (блок "Действие"), блока акцептора результатов действия и блока оценки результатов действия.

Двойными стрелками показаны каналы передачи управления между данной ФС и другими ФС (обозначены цифрами 1-5). Сплошными двойными стрелками показаны каналы 1,4, работающие в обычном режиме (при выполнении действия и без передачи управления подсистеме), штриховыми стрелками – каналы 2,3,5, работающие только при определенных условиях.

Блок афферентного синтеза включает в себя массив нейронных сетей $\{NN_i\}$, осуществляющих прогноз результатов возможных действий A_i ($i = 1, \dots, n$). Считаем, что ФС может выбрать одно из n возможных действий.

4. Работа отдельной ФС

Рассмотрим работу отдельной ФС. В данный такт времени t рассматриваемая ФС может быть активна либо неактивна. Неактивные ФС находятся в ждущем режиме. Активная ФС выполняет указанную выше задачу. Рассмотрим работу активной ФС.

Активация ФС производится по команде суперсистемы (по каналу 1). При этом суперсистема определяет текущую цель G для данной ФС. Считаем, что цель задается в виде вектора, компоненты которого нормированы и представляют собой действительные числа в интервале $[0,1]$. На вход активированной ФС поступает также входной сигнал $X(t)$, характеризующий состояние внешней и внутренней среды анимата. По входному сигналу $X(t)$ определяется прогноз результата каждого из возможных действий A_i . Прогноз результата i -го действия A_i осуществляет i -я нейронная сеть массива прогнозирующих нейронных сетей $\{NN_i\}$ блока афферентного синтеза. Таким образом, определяются прогнозы $Pr_i(t)$ для всех действий A_i ($i = 1, \dots, n$). Прогнозы представляют собой векторы $Pr_i(t)$, нормированные таким же образом, как и цели G .

Отметим, что каждая нейронная сеть выполняет функцию гетероассоциативной памяти, т.е. реализует отображение $X(t) \rightarrow Pr_i(t)$.

Далее прогнозы поступают в блок принятия решения ПР, в котором прогнозы $Pr_i(t)$ сравниваются с целью G и определяется: есть ли действия, для которых рассогласование между целью и прогнозом $delta_i = \|Pr_i(t) - G\|$ меньше заданного порога ($delta_i < th$). $\| \|$ означает некоторую меру близости, например, Евклидову.

Если таких действий нет, то считается, что данная ФС с заданием не справилась, управление передается назад "суперсистеме" (по каналу 2), а во всех нейронных сетях блока афферентного синтеза данной ФС производится разобучение (веса синапсов нейронных сетей немного случайно варьируются).

Если есть действия, для которых $delta_i < th$, то определяется номер выполняемого действия A_k по минимальной удаленности прогноза и цели: $k = \arg \min \{delta_i\}$. При этом прогноз ожидаемого результата действия Pr_k передается в акцептор результатов действия. Далее действие A_k выполняется. Особые действия соответствуют передаче управления вниз по уровням иерархии (по каналу 3). При этом данная ФС задает и подцель SG_k для подсистемы. Специфику таких особых действий рассмотрим ниже. Пока остановимся на рассмотрении обычных действий.

После выполнения действия A_k определяется состояние внешней и внутренней среды анимата в следующий такт времени $t + 1$ и оценивается результат действия R_k (в блоке оценки результатов действия). Данные о результате R_k поступают в блок акцептора результатов действий. Далее определяется, насколько прогноз отличается от результата, т.е. вычисляется величина $Delta = \|R_k - Pr(R_k)\|$, где $Delta$ – величина ошибки прогноза.

Если величина ошибки прогноза $Delta$ меньше/больше некоторого заданного значения Th , то соответствующая нейронная сеть NN_k массива прогнозирующих нейронных сетей $\{NN_i\}$ блока афферентного синтеза дообучается/разобучается, соответственно. При *дообучении* закрепляется отображение между входом $X(t)$ и достигнутым результатом R_k , а при *разобучении* проводится случайная модификация весов синапсов данной нейронной сети NN_k .

После выполнения действия и обучения происходит передача информации о достигнутом результате R_k на верхний уровень, "суперсистеме" и возвращается управление суперсистеме (по каналу 4).

Опишем *процедуру обучения нейронных сетей*.

Считаем, что нейронные сети NN_i имеют структуру, соответствующую обычному методу обратного распространения ошибки [Rumelhart et al., 1986]. Т.е., мы предполагаем, нейронная сеть многослойная, без обратных связей, нейроны имеют логистическую активационную функцию:

$$Y_j = F(\sum_i w_{ij} X_i), \quad F(b) = [1 + \exp(-b)]^{-1},$$

где Y_j – выход j -го нейрона, X_i – входы нейрона, w_{ij} – синаптические веса j -го нейрона. В процедуре дообучения (при достоверном прогнозе, когда $Delta < Th$) мы требуем, что нам надо уточнить тот прогноз, который сделала данная нейронная сеть, и полагаем, что полученный результат R_k есть целевой вектор T , который нужно подкрепить. Далее применяется метод обратного распространения ошибок и закрепляется соответствие $X \rightarrow R_k$ между входом и полученным результатом. За счет этого, если сеть работает правильно, т.е. она дает достаточно достоверный прогноз, то соответствие между входом X и прогнозом Pr_k будет уточняться. А если сеть срабатывает неправильно (ее прогноз недостоверен), то, как сказано выше, данная нейронная сеть будет наказываться – синаптические веса в ней будут случайно варьироваться.

Итак, рассматриваемая процедура обучения формирует отображения $X(t) \rightarrow Pr_i$, обеспечивающие правильные прогнозы результатов действий.

5. Работа системы управления аниматом в целом

Выше мы рассмотрели работу отдельной ФС. ФС получает задание (достичь определенную цель) от "ФС-суперсистемы", и если она компетентна выполнить задание, она его выполняет. Если ФС некомпетентна выполнить задание, то она только сообщает суперсистеме, что она не смогла выполнить задание.

Часть действий ФС состоит в передаче управления на нижние уровни. Такие действия мы выше обозначили как особые. Рассмотрим процессы, происходящие в системе управления аниматом при этих действиях.

При передаче управления на нижний уровень ФС верхнего уровня (обозначим эту ФС как ФС_В) задает подцель управления SG_k для ФС нижнего уровня (обозначим ее ФС_Н). Цель подается на блок афферентного синтеза ФС_Н («с точки зрения» ФС_Н подцель SG_k может рассматриваться как основная цель G). При этом ФС_В верхнего уровня запоминает свой входной сигнал $X_B(t)$, который был в момент передачи управления, и прогноз $Pr(R_k)$, выполненный в соответствии с этим сигналом и намеченным действием передачи управления на нижний уровень; эти векторы запоминаются до времени возврата управления с нижнего уровня. Для определенности считаем, что запоминание векторов $X_B(t)$ и $Pr(R_k)$ происходит в блоке акцептора результатов действия системы ФС_В. Как только ФС_Н нижнего уровня в первый раз после момента времени t возвращает управление ФС_В (считаем, что это происходит в такт времени $t + t_1$; управление возвращается по каналу 5), то в ФС_В происходит обучение. ФС_В оценивает, насколько правильным был ее прогноз в момент t . Обучение происходит точно так же, как это было описано выше, за тем исключением, что прогноз $Pr(R_k)$, сделанный в момент t , сравнивается с результатом R_k , полученным в момент $t + t_1$, а не с непосредственно полученным в следующий момент $t + 1$ результатом.

Итак, предложенная иерархическая схема позволяет промоделировать достаточно нетривиальное управление адаптивным поведением.

Отметим, что в данную схему управления поведением анимата несложно включить процедуру прерывания верхними уровнями работы нижних уровней иерархии функциональных систем, с помощью специальных связей между ФС. Например, если в ФС1, отвечающую за безопасность, поступил сигнал, характеризующий серьезную опасность для жизни анимата, а анимат занимался поиском "пищи" в дереве решений, "возглавляемом" ФС2, то ФС1 имеет право прервать работу ФС2 и дать команду на избежание опасности.

Также нетрудно включить в данную схему процесс формирования новых ФС. Например, если некоторая ФС_Н не находит нужного решения, то блок акцептора результатов действий ФС_В ("ФС-суперсистемы" для данной «ФС-субсистемы») может дать (с определенной малой вероятностью) спец-команду на формирование новой случайной ФС, работающей на том же уровне, что и данная ФС_Н, которой в дальнейшем ФС_В сможет передавать аналогичные задания. Т.е. при невыполнении задания может быть создан "дублер" несправившейся ФС.

6. Обсуждение

Выше изложена только основная структура схемы управления. Ее нужно дополнить 1) схемой задания целей, 2) моделью среды, в которой происходит жизнь анимата. Рассмотрим эти аспекты.

В простейшем случае цели можно задать "руками". Т.е., конструктор модели продумывает, каковы могли бы быть цели анимата, и задает все цели и подцели.

Второй вариант задания целей и подцелей таков. Конструктор задает основные потребности анимата: питание, безопасность, размножение, потребность накопления знания [Жданов, 1999] и цели, соответствующие удовлетворению этих потребностей. Помимо этого, конструктор задает некоторые исходные подцели. Но эти подцели могут варьироваться в процессе жизни анимата. Например, если ФС_В ("суперсистема") передает управление ФС_Н («субсистема»), и ФС_Н успешно выполняет задание по поставленной ей подцели, но при этом происходит удаление от той цели, которая поставлена перед ФС_В, то происходит случайная вариация подцели, передаваемой от ФС_В к ФС_Н. В следующий раз ФС_В формирует перед ФС_Н несколько модифицированную подцель.

В качестве примера моделей среды можно предложить уже известные модели, например, такие как обучение анимата эффективно обходить препятствия в заданной ограниченной области пространства [Donnart et al., 1996].

Данный проект не рассматривает процесс формирования иерархической нейросетевой архитектуры управления. Здесь мы только отметим, что подходы к такому процессу формирования есть. Например, в работе [Бурцев, 2002] была продемонстрирована возможность эволюционного формирования иерархической нейросетевой архитектуры управления поведением анимата. Более того, там же продемонстрировано, что одна и та же нейронная сеть может выполнять множество тех функций, которые в данном проекте распределены в явном виде по различным нейронным сетям. Тем не менее, реализация предлагаемого здесь проекта в компьютерной модели важна, так как она могла бы позволить максимально четко и явно представить работу системы управления целостным адаптивным поведением.

Список литературы

- [Анохин, 1975] Анохин П.К. Очерки по физиологии функциональных систем. – М.: Медицина, 1975. См. также: <http://www.keldysh.ru/pages/BioCyber/RT/Functional.pdf>
- [Бонгард и др., 1975] Бонгард М.М., Лосев И.С., Смирнов М.С. Проект модели организации поведения – «Животное» // Моделирование обучения и поведения. – М.: Наука, 1975.
- [Бурцев, 2002] Бурцев М. С. Формирование иерархии целей в модели искусственной эволюции // IV Всероссийская научно-техническая конференция «Нейроинформатика-2002». Сборник научных трудов. В 2-х частях. Ч.1. М.: МИФИ, 2002. См. также: <http://wsni2003.narod.ru/Papers/Burtsev.htm>
- [Жданов, 1999] Жданов А. А. Метод автономного адаптивного управления // Известия Академии Наук. Теория и системы управления. 1999. N. 5.
- [Максимов и др., 1997] Максимов П.В., Вайнцвайг М.Н., Максимов В.В. Проект системы, обучающейся целесообразному поведению // Международная конференция ММ-РО 8, Тезисы докладов. – Тверь, 1997.
- [Редько, 2001] Редько В.Г. Эволюционная кибернетика. – М.: Наука, 2001.
- [Судаков, 1997] Судаков К.В. (ред.). Теория системогенеза. – М.: Горизонт, 1997.
- [Умрюхин и др., 2000] Умрюхин Е.А., Судаков К.В. Информационная модель системной организации психической деятельности человека («детектор интеллекта») // Моделирование функциональных систем (под ред. Судакова К.В. и Викторова В.А.). – М.: РАМН, РСМАН, 2000.
- [Швырков, 1978] Швырков В.Б. Теория функциональной системы как методологическая основа нейрофизиологии поведения // Успехи физиологических наук. 1978. Т. 9. №1.
- [Donnart et al., 1996] Donnart J.Y. and Meyer J.A. Learning Reactive and Planning Rules in a Motivationally Autonomous Animat // IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics. 1996. V. 26. N. 3.
See also: <http://www-poleia.lip6.fr/ANIMATLAB/#Publications>
- [Rumelhart et al., 1986] Rumelhart D.E., Hinton G.E., Williams R.G. Learning representation by back-propagating error // Nature. 1986. V.323. N. 6088.