

# Когнитивная архитектура работы мозга, основанная на теории функциональных систем П.К. Анохина

Е.Е. Витяев

Институт математики им. С.Л.Соболева СО РАН, Новосибирск

## Понятие цели и принцип опережающего отражения действительности

Рассмотрим целенаправленное поведение. Оно формируется *целью*. Цель нельзя достичь, не имея критерия её достижения. При постановке цели критерием её достижения является некоторый *акцептор достижения цели*. Цель без акцептора достижения цели (*акцептора результатов действия в целенаправленном поведении*) – не может существовать. Поэтому наличие акцептора является необходимым условием её существования. Таким образом, понятие акцептора результатов действия возникает сразу же из понятия цели. Понятие цели определяет также понятие *результата достижения цели*, когда критерий достижения цели удовлетворяется.

Понятие цели парадоксально – цель ничего не говорит о том, *как* её достичь и *как* надо организовать целенаправленное поведение. Для того, что бы знать – как достичь цели, нужен опыт. Если нет никаких знаний и опыта, то поведение организуется методом «проб и ошибок». В этом случае метод «проб и ошибок» является единственно возможным. Для организации поведения методом «проб и ошибок» существует специальная ориентировочно-исследовательская реакция.

Опыт складывается из полученных методом «проб и ошибок» случаев достижения цели. Здесь вступает в силу *принцип опережающего отражения действительности*. Он разрешает парадокс цели и определяет, *как* достичь цели. В соответствии с этим принципом на основании имеющегося опыта выбирается один из способов достижения цели.

Следующий вопрос, который возникает – как формируется способ достижения цели по принципу опережающего отражения действительности? На основании имеющегося опыта мы можем предвосхитить (с опережением по времени по внутреннему контуру мозга) достижение цели – удовлетворение акцептора результатов действия. Эти предвосхищения определяют способы (последовательностями действий), которыми можно достичь цели. Оценивая эмоциями различные способы, можно выбрать один из них – тот, который наиболее вероятно и с наименьшими затратами приводит к цели. Это осуществляется на стадии принятия решений. Принятие решений фиксирует не только способ достижения цели, но всю последовательность действий, которую надо осуществить, а также все промежуточные результаты и рецепторы этих результатов (акцептор результатов действия), которые нужно достичь.

Опережающее отражение действительности не может осуществляться линейно. Оно структурируется в соответствии со структурой внешнего мира. Эта структура является структурой предвос-

хищений стимулов внешнего мира. Она возникает автоматически в виде формирования значений промежуточных результатов действий и пространства целей и результатов (см. ниже).

Формирование этой структуры и промежуточных результатов происходит в два этапа:

1. Если некоторая последовательность действий хорошо прогнозируется, то такую последовательность действий назовём *стандартной последовательностью действий*.

Припишем этим последовательностям в качестве значения-результата ту совокупность стимулов, которые были предвосхищены и получены в результате осуществления этих последовательностей действий. Эти совокупности стимулов формируют *акцептор результатов последовательностей действий*.

2. Если действие переходит в новую ситуацию, стимулы которой плохо предвосхищаются из предыдущей ситуации, то возникает разрыв в структуре предвосхищений. Этот разрыв восполняется обнаружением единства последовательностей действий в рамках некоторой большей целостности (например, более масштабной или более общей когнитивной карты). Эта большая целостность, существующая в иерархической структуре внешнего мира, может взять на себя организацию (стыковку) последовательностей действий. Такая организация осуществляется уже на более высоком уровне, в котором более простые последовательности действий являются элементарными единицами действий и результатов.

Таким образом, опережающее отражение действительности автоматически формирует иерархию результатов деятельности (пространства целей и результатов – см. далее).

В своей работе Дж. Хокинс [11] также говорит о последовательностях действий, но не доводит свои рассуждения до объяснения иерархии предсказаний и пространств целей и результатов, а использует их для нахождения инвариантов восприятия.

Отсюда мы получаем общую когнитивную архитектуру работы мозга, где прогноз достижения цели – главная функция работы мозга. Более подробные схемы этой архитектуры приведены ниже, а статьи, описывающие эти схемы, приведены на сайте [13]. Память, внимание, воображение, интуиция, восприятие, язык и мышление являются производными в данной архитектуре. Некоторые из этих понятий уже описаны в предыдущих работах [13]. Данная архитектура была успешно применена для разработки аниматов [7].

## Модель работы функциональной системы

На рис. 1 приведена модель работы функциональной системы [2-6]. Пусть функциональной

системе мотивацией ставиться цель  $P_0$ . Представим цель, как запрос к функциональной системе – достичь цели  $P_0$ . На вход функциональной системы подаётся также информация об окружающей среде в виде описания ситуации  $P_{i1}, \dots, P_{im}$  и обратная афферентация от осуществлённых действий. Афферентным синтезом из памяти извлекается вся информация, связанная с достижением цели  $P_0$ , которая храниться в памяти в виде множества закономерностей вида:

- $\langle P_{i1}, \dots, P_{ik}, A_i \rangle \rightarrow P_0$  - для закономерностей нижнего уровня иерархии, когда осуществляются действия  $\langle A_i \rangle$  (последовательности действий);
- $\langle P_{i1} \& \dots \& P_{ik} \& P_{j1} \& \dots \& P_{jn} \rangle \rightarrow P_0$  - для закономерностей более высокого уровня иерархии, когда надо организовать выполнение последовательностей действий более низкого уровня путём достижения подцелей  $P_{j1}, \dots, P_{jn}$ .

Все закономерности содержат свойства  $P_{i1}, \dots, P_{ik}$  текущей ситуации  $P_{i1}, \dots, P_{im}$  и информацию (обратную аффертацию) о результатах уже осуществлённых действий. Эта информация говорит о том, какие условия внешней среды должны быть выполнены и какие результаты от предыдущих действий должны быть достигнуты, чтобы достижение цели в данной ситуации было возможным. Закономерности означают что, если ситуация включает в себя условия, зафиксированные в закономерности, то после осуществления действия  $A_i$  (цепочки действий) или достижения подцелей  $P_{j1}, \dots, P_{jn}$  мы достигнем цель  $P_0$  с определенной вероятностью. Достижение подцелей осуществляется отправкой запросов на их достижение вниз по иерархии подцелей, что обозначено на рис.1 двойной стрелкой вниз. Достижение этих

подцелей может потребовать достижение ещё более низких по иерархии целей и т.д. Если какая-то из подцелей не может быть выполнена в данной ситуации (нет закономерностей предсказывающих достижение подцели в данной ситуации), то ответом на запрос является отказ и соответствующая закономерность исключается из рассмотрения.

Таким образом, активация закономерностей в блоке афферентного синтеза автоматически извлекает из памяти набор действий  $A_i$  или подцелей, которые могут привести к достижению цели  $P_0$  в данной ситуации. Этот набор вместе с оценками условных вероятностей достижения цели передается в блок принятия решений. В случае передачи действий блок принятия решений выбирает действия  $A_i$ , которые с максимальной вероятностью приводят к достижению цели. В случае же передачи подцелей блок принятия решений выбирает такие подцели, которые с максимальной вероятностью приводят к достижению цели. При этом учитывается вероятность достижения подцелей, оцениваемая в подсистемах и передающаяся как прогноз назад в блок принятия решений (см. двойную пунктирную стрелку внизу блока принятия решений). Общая вероятность достижения цели вычисляется таким же способом, как вычисление ответа на запрос в логическом программировании – путем иерархического развертывания вычисления вниз по иерархии всех подцелей. Рекурсивное вычисление вероятностей определяется следующим образом: если вероятности  $v_1, \dots, v_n$  достижения подцелей  $P_{j1}, \dots, P_{jn}$  уже вычислены, то вероятность достижения цели  $P_0$  равна  $v \cdot v_1 \cdot \dots \cdot v_n$ , где  $v$  - вероятность правила  $\langle P_{i1} \& \dots \& P_{ik} \& P_{j1} \& \dots \& P_{jn} \rangle \rightarrow P_0$ , в которое входят эти подцели.

Полученный прогноз достижения цели  $P_0$  отправляется в акцептор результатов действий (двойная пунктирная стрелка) и вверх по иерархии

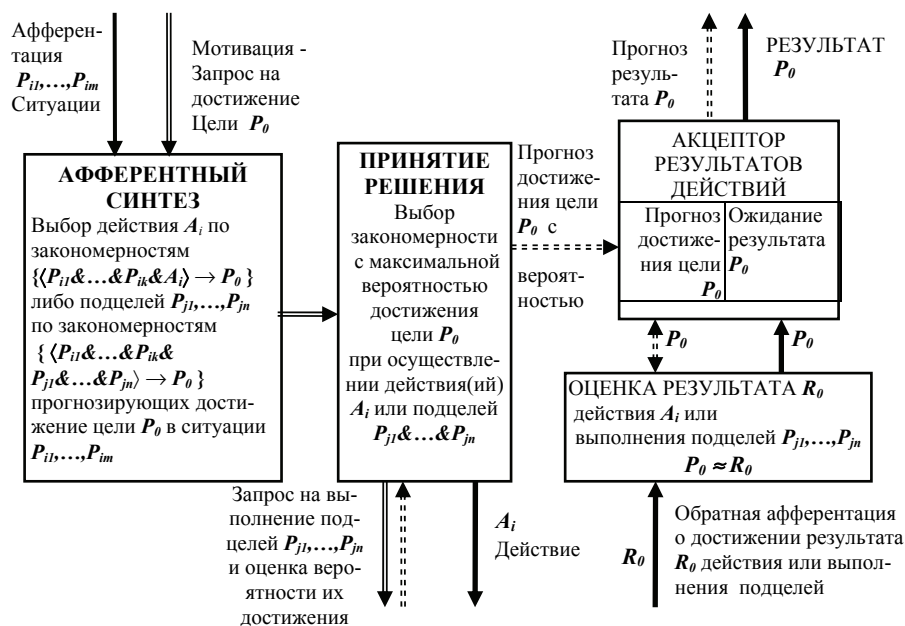


Рис. 1

функциональных систем. Если эмоциями будет выбран данный способ достижения цели, то будет принято решение о данном способе действия и будет сформирован акцептор результатов действий, контролирующий правильность выполнения всех последовательностей действий и достижение всех подцелей во иерархии. После этого запускается последовательное выполнение действий и получения всех результатов действий.

Данные о полученном результате  $R_i$  поступают в акцептор результатов действий в блок оценки результата. Проводится сравнение спрогнозированного и полученного результата. Если прогноз и результат совпадают с заданной степенью точностью  $P_0 \approx R_0$ , то акцептор результатов действий фиксирует достижение цели и получение результата  $P_0$ . Сообщение о достижении цели передается вверх по иерархии функциональных систем.

### Принятие решений. Переключающая и подкрепляющая функции эмоций.

Принятие решений осуществляется с помощью переключающей функции эмоций, определённой в информационной теории эмоций П.В.Симонова [8-9]. Она основана на вероятностном прогнозе достижения цели, который осуществляется ещё до всяких действий. Переключающей функцией эмоций осуществляется анализ всех возможных способов достижения цели. Различные способы достижения цели (например, при движении по некоторой местности), имеют различную вероятность, различные энергетические затраты и различные возможные опасности, связанные с отрицательными эмоциями, и т.д. В этом случае задача становится как минимум трёх параметрической - вероятность достижения цели, суммарное значение отрицательных эмоций (от энергетических затрат, опасностей, риска, трудностей и т.д.) и значение положительных эмоций (от достижения цели). Для эффективного принятия решений необходим синтез всех этих показателей в один параметр, что и осуществляется эмоциями. Они

интегрируют как вероятность достижения цели, так и положительные и отрицательные эмоции, связанные с достижением цели. На основе эмоций и принимается решение: “Зависимость эмоций не только от величины потребности, но и от вероятности её удовлетворения чрезвычайно усложняет конкуренцию сосуществующих мотивов, в результате чего поведение нередко оказывается переориентированным на менее важную, но легко достижимую Цель: “синица в руках” побеждает “журавля в небе” ... [8, разд. 2.2].

На рис. 2 мотивация (неудовлетворенная потребность) показана двойной стрелкой подходящей слева к блоку эмоций, а прогноз достижения цели двойной пунктирной стрелкой, подходящей справа. Рассогласование между прогнозом и «наличной действительностью» можно измерить как  $1-v_1 \dots v_4$ , где  $v_i$  – вероятность достижения цели блоком  $i$ , т.е. вероятность закономерности выбранной в блоке  $i$ .

На рис. 2 переключающая функция эмоций реализуется тем, что:

- по всем извлеченным из памяти способам достижения цели получается прогноз достижения цели и передается в блок эмоций. На рис. 2 путь прогноза показан двойной пунктирной стрелкой, проходящей через блоки принятия решений и акцепторы результатов действий;
- принимается решение о выборе того или иного целенаправленного поведения, которое обладает максимальной эмоциональной оценкой и, значит, с максимальной вероятностью при минимальных затратах и отрицательных эмоциях приводит к достижению цели;
- формирует план достижения цели и акцептор результатов действий.

В процессе достижения цели в соответствии с планом действий проявляется подкрепляющая функция эмоций. П.В.Симонов считает, что необходимым условием подкрепления является не сам подкрепляющий раздражитель (санкционирующая

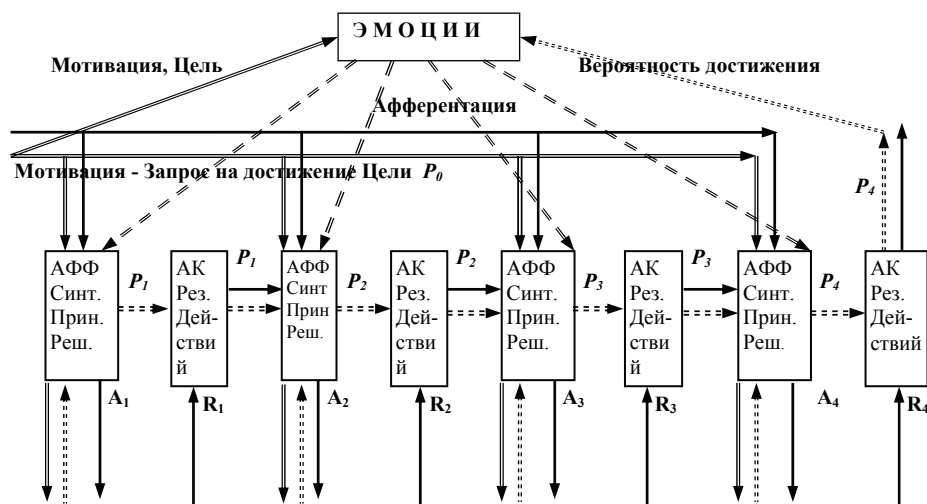


Рис 2.

афферентация), а действие положительных эмоций при наличии мотивации: “Однако ни афферентация из полости рта (санкционирующая афферентация), ни голодовое возбуждение (мотивация) сами по себе не могут играть роль подкрепления ... Только интеграция голодового возбуждения от фактора, способного удовлетворить данную потребность, т.е. механизм, генерирующий положительную эмоцию, обеспечивает выработку условного рефлекса” [8. с.34].

Участие оценки вероятности в формировании эмоций сразу же делает подкрепление более точным [4-5]: любая последовательность действий, приближающая к цели и увеличивающая прогноз достижения цели, вызывает положительную эмоцию и подкрепляет те «мозговые структуры», которые осуществили эту последовательность действий. Следовательно, эмоции, основанные на вероятностном прогнозировании, осуществляют подкрепление каждой успешной последовательности действий, увеличивающей вероятность достижения цели.

На рис. 2 от блоков акцепторов результатов действия идет две стрелки – одна пунктирная обозначающая прогноз достижения цели, вторая сплошная, обозначающая достижение цели и получение результата и преобразующая прогноз в факт, имеющий вероятность 1. Поэтому, как только, например, в блоке 1 достигнут результат, то вероятность прогноза увеличится с  $v_1 \cdot v_2 \cdot v_3 \cdot v_4$  до  $1 \cdot v_2 \cdot v_3 \cdot v_4$ . Тогда положительное рассогласование, вызывающее положительную эмоцию, равно  $1 \cdot v_2 \cdot v_3 \cdot v_4 - v_1 \cdot v_2 \cdot v_3 \cdot v_4 = (1 - v_1) \cdot v_2 \cdot v_3 \cdot v_4$ . Возникшая положительная эмоция действует генерализованно и сразу же подкрепляет те структуры, которые осуществили действие, т.е. блок 1 и выбранную там закономерность. Поэтому каждая закономерность, действие которой приблизили к достижению цели, будет подкреплена на величину увеличения вероятности приближения к цели.

### Пространства целей и результатов

Как говорилось ранее, опережающее отражение действительности не может осуществляться линейно.

Объясним, почему цепочки действий выстраиваются в пространство целей и результатов некоторого уровня (по Н.А.Бернштейну [3]). Результаты цепочек действий некоторого уровня определяют язык, отражающий единицы действий этого уровня. Этот язык замкнут, так как «смыслы» должны быть стыкуемы между собой для осуществления сложных последовательностей действий соответствующего уровня. Язык каждого уровня дает пространство целей и результатов соответствующего уровня.

Когда возникает разрыв в структуре предвосхищений соответствующего уровня, то необходимо перейти на более высокий (масштабный) уровень рассмотрения. Последовательности действий более высокого уровня (их цели и результаты) имеют более широкий (масштабный) смысл и,

соответственно, обеспечиваются более широкой афферентацией. Например, на уровне синергий [3] пространством целей и результатов являются движения тела, а обратной афферентацией – импульсация от мышц; на уровне С «пространственного поля» пространством целей и результатов является внешнее пространство, а обратной афферентацией восприятие окружающего мира; на уровне D - предметного действия пространством целей и результатов является оперирование с предметами, а обратной афферентацией – смыслом оперирования предметами. Без адекватной обратной афферентации, содержащейся среди признаков  $P_{11}, \dots, P_{ik}$  в условии закономерностей невозможно правильно организовать иерархию предсказаний в пространстве целей и результатов.

### Благодарности

Эта работа поддержана интеграционными проектами СО РАН № 47, 111, 119 и программой президента Российской Федерации поддержки научных школ НШ-3606.2010.1

### Литература

1. Анохин П.К. Эмоции // Большая медицинская энциклопедия т. 35, М. 1964.
2. Анохин П.К. Биология и нейрофизиология условного рефлекса. М., 1968.
3. Бернштейн Н.А. Биомеханика и физиология движений. // Избранные психологические труды, Москва-Воронеж, 1997, с.605
4. Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова. Нейроинформатика, 2008, том 3, № 1, стр. 25-78.
5. Витяев Е.Е. Принятие решений. Переключающая и подкрепляющая функции эмоций // VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006», Сборник научных трудов, Москва, 2006, с.24-30.
6. Витяев Е.Е. Объяснение Теории Движений Н.А.Бернштейна. VII Всероссийская научно-техническая конференция «Нейроинформатика-2005», Сборник научных трудов, часть 1., Москва, 2005, стр.234-240
7. Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления // Нейроинформатика, 2008, том 3, № 1, стр. 79-107.
8. Симонов П.В. Эмоциональный мозг. М.: Наука, 1981. с. 140.
9. Симонов П.В. Высшая нервная деятельность человека (мотивационно-эмоциональные аспекты). М.: Наука, 1975. с. 173.
10. Михуенко Е.В., Витяев Е.Е. Моделирование работы функциональной системы // VI Всероссийская научно-техническая конференция «Нейроинформатика-2004». Сборник научных трудов. В 2-х частях. Ч.2., М.: МИФИ, 2004, 124-129.
11. Джефф Хокинс, Сандра Блейкли. Об интеллекте. Вильямс, 2007 г., 240 стр.
12. Evgenii Vityaev. The logic of prediction. In: Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference (August 16-19, 2005, Novosibirsk, Russia), World Scientific, Singapore, 2006, pp.263-276
13. Scientific Discovery: <http://www.math.nsc.ru/AP/ScientificDiscovery>

