

# ФИЛОСОФСКИЙ СМЫСЛ ПРОБЛЕМЫ ЕСТЕСТВЕННОГО И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА\*

*П.К.Анохин*

## **Вступление**

Трудно назвать более интересную научную проблему, чем проблема познания мозга, его фундаментальных механизмов и его молекулярной природы. Прямым следствием развития этой сферы знания должно быть разумное управление мозгом в будущем, а также использование законов его деятельности для конструирования различных механизмов, составляющих основу технического прогресса в нашу эпоху.

Когда один из корреспондентов спросил «отца кибернетики» Норберта Винера, допускает ли он возможность того, что высокоорганизованные «интеллектуальные машины» смогут в будущем поработить человека, Винер не без иронии ответил: «Если это и произойдет, то только по вине человека...»

И вопрос и ответ подчеркивают крайнюю злободневность проблемы естественного и искусственного интеллекта. Дело в том, что некоторые кибернетики полагают, будто можно все смоделировать и даже создать машины, которые по своим интеллектуальным качествам превзойдут человека. Такие надежды широко распространены среди физиков, математиков, электроников. Однако здесь допускается серьезная логическая ошибка, суть которой будет разъяснена в данной статье.

Несмотря на важность проблемы интеллекта, переход от изучения фундаментальных проблем мозговой деятельности на уровне интеллекта к использованию результатов исследований в технической кибернетике пока еще наталкивается на серьезные препятствия: отсутствует достаточно полная модель искусственного интеллекта, соответствующая современным представлениям о деятельности мозга в естественных условиях.

Проблема естественного и искусственного интеллекта ставит перед нами также и целый ряд вопросов философского характера. Действительно, тезис материалистической философии — «материя первична, сознание вторично» — устанавливает органическую историческую связь между этими феноменами, поскольку мы знаем, что неорганический мир существовал задолго до появления жизни на нашей планете, и, следовательно, интеллект должен был неизбежно отразить законы неорганического мира и «вписаться» в них. Но если это так, то все свойства интеллекта должны были развиваться на базе предшествующих органических форм и, естественно, должны быть приспособлены для оперирования объектами внешнего мира.

Иначе говоря, естественный интеллект (в примитивной форме — интеллект животных и в высшей форме — интеллект человека) неизбежно должен действовать на основе объективно познаваемых процессов и механизмов. Рассматривая этот вопрос в философском аспекте, мы можем сказать, что изучение искусственного интеллекта является одним из важнейших этапов познания материальной природы психических явлений и, следовательно, способствует дальнейшему развитию диалектического материализма.

Ясно, что надеяться на создание искусственного интеллекта можно только после создания достаточно солидного «концептуального моста», который даст возможность максимально использовать наши фактические знания о принципах работы мозга.

---

\* Опубликовано в журнале «Вопросы философии» (1973, № 6, с.83-97).

### **Важнейшие черты интеллекта и их характеристика**

В последние годы в связи с попытками конструирования искусственного интеллекта исследователи столкнулись с необходимостью определения самого интеллекта и выявления его характерных черт. Без этого невозможен контакт между нейрофизиологами, психологами и специалистами по технической реализации основных черт интеллекта в моделях и рабочих конструкциях.

Успех Маккаллока в создании искусственной нейронной сети объясняется как раз тем, что он наиболее четко выделил некоторые характерные логические черты мозговой деятельности и использовал их для конструкции распознающего и «мыслящего» устройства (Маккалок, Питтс, 1956; McCulloch, 1968; McCay, 1959; Минский, 1961). Благодаря этим исследованиям проблема искусственного интеллекта стала широко разрабатываться именно нейрокибернетиками, а не нейрофизиологами. Последние продолжали оставаться на позициях классической нейрофизиологии с господствующей в ней «рефлекторной» манерой мышления, не дающей возможности понять решающие свойства, характерные именно для интеллектуальной деятельности. Естественным следствием этого была неопределенность в понимании нейрофизиологических свойств интеллекта и отсутствие научно обоснованных формулировок. Это обстоятельство значительно затруднило контакт между психологами, нейрофизиологами и кибернетиками.

Конструкторы искусственного интеллекта все чаще и чаще подходили вплотную к изучению как раз тех свойств мозга, которые физиологи в своих исследованиях даже не затрагивали.

Пожалуй, Марон был первым из кибернетиков, кто пришел к выводу, что не может быть и речи о понимании интеллекта вообще и о конструировании «интеллектуальной машины», если эта система не будет обладать способностью к предсказанию. Сопоставляя человеческий мозг с наиболее совершенными машинами кибернетического характера, он особенно отчетливо сформулировал их различие: способность к предсказанию у мозга и отсутствие этой способности у машин.

Условный рефлекс, по И. П. Павлову, несомненно, основывается на предсказании, поскольку условная реакция имеет «предупредительный» характер. Как показал наш анализ, в составе условного рефлекса имеется аппарат, который создается в процессе формирования рефлекса для оценки предстоящей ситуации, т. е. акцептор результата реакции (Анохин, 1949).

Однако нейрофизиологи совсем не затрагивают в своих работах этой проблемы, поскольку господствующий принцип в понимании нервной деятельности — принцип рефлекторной дуги полностью исключает саму возможность предсказывать будущее.

Суть дела в следующем: нервное возбуждение, вызванное раздражением какого-либо рецептора, согласно рефлекторной теории распространяется по «рефлекторной дуге» на основе линейно-поступательного принципа, т. е. от пункта к пункту. Между тем по своей сути предсказание предполагает «забегание вперед», опережение хода возбуждений: в самом начале появляются процессы и физиологические аппараты, которые должны проявить себя лишь в заключительной стадии рефлекторного действия (Анохин, 1935, 1949).

На этот вид нервной деятельности указывал О. Аттли. Описывая конструкцию управляющих механизмов, он отмечал: «Вычислитель может непрерывно подсчитывать для каждого управляющего движения, которое опробовалось раньше, вероятность того, что оно приведет к цели» (Аттли, 1964, с. 430). Говоря о цели, к которой ведут все движения, Аттли тем самым подчеркивает ее направляющее влияние на те действия, которые приближают будущие события.

Наиболее полно роль цели и предсказания рассмотрена в книге Л. Фогеля, А. Оуэнса и М. Уолша (1969).

Определяя понятие «искусственный интеллект», авторы стремятся найти те характерные признаки, которые могли бы быть общими и для естественного и для искусственного интеллекта. Они совершенно правильно, на наш взгляд, сосредоточивают внимание не на тонкости, четкости и скорости выполнения отдельных операций, а на *логике механизмов*, составляющих интеллект. Среди этих механизмов на первый план они ставят механизмы «принятия решения» и «предсказания», т. е. формирования цели. Авторы считают, что будет «...более содержательным определять интеллект в терминах поведения некоего стремящегося к цели существа и измерять степень его интеллекта по адекватности принимаемых им решений» (Фогель и др., 1969, с. 19). Таким образом, мы видим, что определение интеллекта охватывает наиболее сложные формы поведенческой деятельности: цель, «принятие решения», предсказание. По сути дела это верно. Однако недостаток такого определения состоит в том, что существенные и характерные для интеллекта факторы просто перечисляются, а не даются в той логической связи и последовательности, которая соединила бы их прочной нитью системного детерминизма.

И действительно, в приведенных выше определениях, как, впрочем, и во многих других, *«цель» выступает как нечто наперед данное*. А дальше начинает рассматриваться цепь поведенческих актов, направленных на достижение этой цели. Но как возникла сама цель? Какие факторы и какие материальные процессы предшествовали ее появлению и создали из нее материальный аппарат, направляющий специфическое стремление организма? Упомянутые выше авторы, а также многие другие (Месарович, 1970; Уотерман, 1970; Sadovsky, 1971) эту фазу «предцели» совершенно не рассматривают.

То же самое можно сказать и о принятии решения. Какие факторы толкают организм на принятие именно этого, а не другого решения? Ясно, что в процессе принятия решения происходит непрерывный подбор наиболее адекватного для данной ситуации решения. Но как это происходит? На основе каких конкретных нейрофизиологических механизмов выбирается одна-единственная поведенческая степень свободы из миллионов возможных степеней?

Обычно все эти вопросы исследуются отдельно, вне их логической связи в масштабе целого поведенческого акта, а потому иногда и правильно выделенные факторы естественного интеллекта, как, например, «предсказание» (Фогель и др.), цель и принятие решения (Аттли), остаются изолированными фрагментами интеллекта, не связанными между собой логикой функционирования.

Оценивая нынешний этап в исследовании основных специфических характеристик естественного и искусственного интеллекта, мы можем сказать, что главным недостатком этих исследований является отсутствие универсальной модели, которая логически связала бы все этапы формирования интеллектуальных актов. Естественно, что эта модель должна достаточно полно отражать и нейрофизиологические механизмы каждого фрагмента интеллектуальных процессов.

Анализируя значение и содержание всех попыток моделирования интеллектуальных процессов на базе перцептронов, Ф. Розенблатт очень ярко отразил стремления тех, кто занят исследованием процессов «принятия решения». «В ближайшем будущем,— писал он,— потребуется, по-видимому, разработать целую программу психологических экспериментов с животными или людьми для пополнения наших сведений о характеристиках моделей. Когда это произойдет, модели фактически станут использоваться как *«предсказывающие»* устройства, способные вырабатывать определенные данные (вначале, возможно, довольно грубые), которые у людей до сих пор не наблюдались. Конечным использованием модели мозга с точки зрения психологической ценности как раз и являются эксперименты такого рода, в которых модель правильно предсказывает явления, еще не открытые в биологических системах» (Rosenblatt, 1962).

Как видим, Розенблатт возлагает весьма большие надежды на модели мозга типа «предсказывающих» устройств. Однако, чтобы эти модели правильно предсказывали и правильно ориентировали в будущих явлениях, они должны заимствовать у мозга те

его свойства и механизмы в форме акцептора результатов действия, которые дадут им возможность формировать цель поведения, предсказывать сам результат поведения и ежесекундно контролировать и сличать полученный результат с поставленной заранее целью. Именно этими-то свойствами не обладает ни одна из существующих моделей мозга, т. е. искусственного интеллекта. Ясно, что такая удовлетворяющая исследователя модель может быть построена только при непременном использовании данных нейрофизиологии для постоянной коррекции работы этой модели.

Несмотря на то, что проблемам принятия решения, формулировки цели и предсказания в последнее десятилетие уделяется очень много внимания, все попытки создания искусственного интеллекта находятся лишь в самой начальной стадии.

Как известно, при изучении биологических систем в силу устоявшихся традиций принято считать, что любой поведенческий акт заканчивается действием. При этом полезный результат действия фактически никогда не включается в процесс как самостоятельная физиологическая категория. А между тем именно в этом пункте лежит причина трагического взаимонепонимания между представителями весьма близких областей науки — нейрофизиологии и психологии. Для последней, как известно, цель и принятие решения стали необходимыми факторами в изучении интеллектуальных процессов.

Кибернетика внесла в психологию целый ряд смелых идей, которые заставили ее принять такие синтетические понятия, как цель, польза, предсказание и т. д.

Можно было бы назвать большое число исследований последних лет, в которых детальнейшим образом разрабатывается проблема «принятия решения» в самых разнообразных ситуациях. Наиболее полные обзоры результатов изучения этой темы были сделаны на специальных симпозиумах<sup>1</sup>.

Одним из значительных событий в исследовании характерных свойств интеллекта явилась организация во Франции Института «высшего синтеза», занимающегося изучением «принятия решения» и пути построения «искусственного интеллекта» (Institut des Hautes Syntheses Nice, France). На одной из последних сессий института (1971 г.) специально обсуждался вопрос о соотношении естественного и искусственного интеллекта.

XX Международный психологический конгресс (Токио, 1972) провел особый симпозиум по проблематике «принятия решений» (Dynamic Aspect of «Decision Making») <sup>2</sup>. В докладах У. Эдварса, Г. Экеля, М. Тода, М. Месика были еще раз поставлены вопросы об основных признаках принятия и выполнения решений. Экель акцентировал внимание также на физиологических коррелятах принятия решения и соответствующих ему положительных и отрицательных эмоциональных состояниях.

Все сказанное выше можно резюмировать следующим образом.

а) Большое внимание к проблемам искусственного и естественного интеллекта привело специалистов различных областей науки к актуальным задачам по изучению характерных черт естественного интеллекта и применению результатов этой работы к построению искусственного интеллекта. Сама возможность удачного решения этого вопроса сулит широкие перспективы прогрессивного развития многих областей экономики и промышленности.

Использование результатов таких исследований в электронике, медицине, педагогике и в других областях может привести к революционным сдвигам в этих науках. Поэтому весьма важно правильное понимание сути такого прогресса.

б) Несмотря на значительность проблемы, ее нынешнее состояние нельзя считать удовлетворительным. Нет четкого определения самого понятия интеллекта, его состава и решающих механизмов его отдельных операций. Наиболее синтетические и характерные для интеллекта узловые механизмы, такие, как «принятие решения», «цель», «предсказание», не только не изучены в их глубоком нейрофизиологическом содержании, но даже не вскрыты их операциональные взаимодействия в момент осуществления интеллектуальных актов.

<sup>1</sup> См., например: Decision Processes. Ed. by R. M. Thrall. N. Y., 1954.

<sup>2</sup> См.: XX Congress International de Psychologie. Guide Resume. Tokyo, 1972.

Пожалуй, одним из самых существенных пробелов в изучении искусственного и естественного интеллекта является то, что не выявлено логическое единство в самой архитектуре интеллектуальных актов, не установлена детерминистическая связь между указанными выше узловыми механизмами интеллекта. Каждый из них берется для исследования как нечто отдельное, независимое от других свойств и механизмов интеллекта.

В последующих разделах статьи мы попытаемся применить для исследования этой проблемы системный подход в виде *теории функциональных систем*, разрабатываемой в нашей лаборатории на протяжении последних сорока лет.

### **Функциональная система**

#### **как логическая модель искусственного интеллекта**

Как мы уже видели, одной из существенных тенденций современной нейрофизиологии является *изоляция* отдельных механизмов мозга для удобства экспериментирования над ними в целях изучения их свойств. Этот аналитический прием, общий для многих биологических наук, уже дал значительные результаты. Однако прием этот полезен и хорош только на определенной стадии научно-исследовательского процесса: при сборе первичных материалов и в период подготовки к широким обобщениям.

Все функции организма и особенно функции его нервной системы по своей природе являются логически целостными, и потому понимание их биологического смысла зависит от того «высшего синтеза», в котором выявится реальная роль каждого механизма в образовании целого. Теория функциональной системы как раз и имеет своей целью выявить органическое единство механизмов, которые обычно исследуются в отдельности.

Много лет назад, изучая процесс компенсации нарушенных функций, мы увидели, что все сложные факторы деятельности, такие, как память, эмоция и цель, выступают в органическом единстве и только это единство способно восстановить нарушенную функцию (Анохин, 1935). Такое функциональное единство в силу его системного характера было названо нами *функциональной системой*. Она представляет собой законченную единицу деятельности любого живого организма и состоит из целого ряда узловых механизмов, обеспечивающих логическое и физиологическое формирование поведенческого акта.

Так как характеристика теории функциональной системы неоднократно приводилась нами в различных публикациях, здесь будет дана лишь краткая характеристика ее узловых механизмов с точки зрения их важности для построения «искусственного интеллекта».

Функциональная система устраняет дефект имеющихся схем интеллекта. Как мы уже отмечали, механизм «принятия решения» большинством авторов рассматривается как нечто первичное и исходное для всех других процессов интеллектуального акта. Такой подход не может удовлетворить объективно мыслящего исследователя, поскольку принятию решения должен предшествовать весьма сложный процесс обработки многообразной информации.

Эта стадия интеллектуального акта была названа нами «афферентным синтезом» в связи с тем, что в процессе этого синтеза происходит *одновременная* обработка самой разнообразной информации, поступающей в центральную нервную систему из внешнего и внутреннего мира. На этой стадии «предрешения» синтезируется целый ряд возбуждений. Как и весь поведенческий акт в целом, стадия «предрешения» формируется на основе доминирующей в данный момент эмоции или мотивации. Последняя, говоря психологическим языком, представлена желанием или потребностью. Такое доминирующее возбуждение, как показывают эксперименты на простых формах потребности (голод, жажда, половая потребность и т. д.), обладает способностью извлекать из многочисленных синаптических образований мозга все то, что было связано в прошлом с удовлетворением или разрешением именно этой, доминирующей в данный момент, потребности (исследования наших сотрудников Судакова, Котова, Журавлева и др.).

В процессе распространения возбуждения по нейронам мозга неизбежно привносятся и другие возбуждения — от совокупности факторов внешней обстановки.

Таким образом, и это было показано в эксперименте, на каждом нейроне коры головного мозга одновременно обрабатываются возбуждения трех различных источников: внутреннее возбуждение, связанное с формированием той или иной доминирующей мотивации, внешние возбуждения, представленные содействием данной обстановки, и возбуждения памяти, извлеченные как мотивацией, так и данной обстановочной афферентацией. Только одновременная обработка этих возбуждений и сопоставление всех комбинаций возбуждений с прошлым опытом дают возможность организму принимать то или иное решение для получения полезного результата.

Экспериментальные исследования свидетельствуют о том, что все упомянутые возбуждения, а иногда дополнительно и специальный пусковой фактор (например, условный сигнал) должны одновременно встретиться на одном и том же нейроне, или, правильнее говоря, на каждом из миллионов нейронов.

В стадии «предрешения», т. е. *афферентного синтеза*, во всех случаях формирования поведенческого акта решается главнейший вопрос: какой полезный результат должен быть получен в данной ситуации и при данной комбинации составных возбуждений этой стадии?

Мы видим, что только строго научное исследование стадии «предрешения» может привести к совершенно четкому детерминистическому объяснению и самого процесса принятия решения. Действительно, микроэлектродное исследование отдельных нейронов коры головного мозга показало, что этот процесс обработки всей исходной информации совершается при помощи многих динамических механизмов, биологический смысл которых состоит в том, чтобы выработать наиболее адекватное решение для данной ситуации и обеспечить наиболее точное его исполнение.

Так, например, активизирующие аппараты подкорковой области (гипоталамус, ретикулярная формация) обеспечивают образование ассоциаций и извлечение информации из памяти. Эти же активизирующие возбуждения значительно повышают различные способности нервных элементов коры мозга, и в частности способность к конвергенции на них разнородных возбуждений. К этому надо добавить еще усиление реверберации возбуждений между корой и подкорковыми областями, благодаря чему осуществляется поиск наиболее продуктивного синтеза для предстоящего принятия решения (опыты Шумиловой).

Таким образом, нам необходимо представить себе все то, что могут дать нам нейрофизиологические эксперименты для понимания механизмов принятия решения как одного из главнейших факторов формирования интеллекта.

### **Нейрофизиологические предпосылки принятия решения**

Для понимания этого ответственного синтетического процесса в интеллектуальной деятельности мы должны представить себе отдельный нейрон и миллионы нейронов как образования, обладающие бесчисленным количеством степеней свободы, которые обусловлены способностью нейрона производить самые разнообразные конфигурации нервных разрядов.

Простой математический расчет показывает, что количество степеней свободы в масштабе целого мозга с трудом может быть записано цифрой длиной в 9,5 млн. км!.. Именно это количество степеней свободы головного мозга и представляет собой ту бесконечную клавиатуру, на которой разыгрываются сотни миллионов различных мелодий—поведенческих и интеллектуальных актов.

Итак, мозг и организм в каждый данный момент обладают необъятным количеством степеней свободы, одновременный запуск которых привел бы к чудовищному хаосу в поведении организма. Организованное поведение человека и животных предполагает неизбежное ограничение этого огромного разнообразия. Следовательно, принятие решения по своей сути представляет собой выбор одной степени свободы, наиболее адекватно удовлетворяющей требованиям данной ситуации. Суть проблемы заключает-

ся в том, каким образом мозг осуществляет выбор из миллиардов возможных одной-единственной степени свободы, дающей полезный эффект именно в данной ситуации.

Здесь необходимо коснуться вопроса, который обычно ускользает от исследователя, но который неизбежно возникает, если придерживаться системной точки зрения, т. е. рассматривать весь процесс построения поведенческого акта в аспекте функциональной системы.

При внимательном изучении общей схемы последовательного действия узловых механизмов функциональной системы можно увидеть, что принятие решения ориентировано на тот результат, который соответствует доминирующей в данный момент мотивации. Наблюдения последних лет свидетельствуют, однако, о том, что в стадии афферентного синтеза из памяти извлекаются не только общие афферентные черты той или иной внешней ситуации, но и признаки тех результатов, которые когда-то получались при подобных мотивационных и эмоциональных состояниях.

Иначе говоря, наш мозг обладает поразительной способностью к системной генерализации возбуждения, способностью охватывать не только частные признаки каких-либо событий, но и степень успешности и полезности тех результатов, которые были получены в аналогичных ситуациях в прошлом. Результаты прошлого могут последовательно извлекаться из памяти и сопоставляться с потребностью данной ситуации до тех пор, пока доминирующая ныне мотивация не станет вполне соответствовать одному из результатов прошлого. Пожалуй, это одна из самых замечательных способностей нашего мозга, которую можно было бы назвать *перебором мнимых результатов прошлого* и сопоставлением их с потребностью данного момента.

Возникнув в эмоциональных структурах мозга (гипоталамус, лимбическая система и ретикулярная формация), мотивационное возбуждение иррадирует даже в те структуры мозга, которые хранят в памяти *результаты* различных удовлетворений именно данной мотивации в прошлом.

Например, состояние аппетита зависит от того, что латеральное ядро гипоталамуса непрерывно раздражается «голодной» кровью. Это возбуждение, поднимаясь в кору мозга в восходящем направлении, мобилизует здесь элементы прошлого опыта, относящегося именно к данной мотивации. Мы начинаем перебирать возможность ее удовлетворения в соответствии с данной ситуацией. Попросту говоря, ищем, где бы мы могли закусить. При этом часто говорим, что в такой-то ресторан мы не пойдем, «поскольку там плохо кормят».

Что значит такое решение с нейрофизиологической точки зрения? Оно означает, что, перебирая многие возможности удовлетворения пищевой мотивации, мы не только извлекаем из памяти сведения о посещении этого ресторана в прошлом, но также и результат посещения, т. е. вспоминаем саму еду и вкусовые ощущения от пищи, полученные когда-то именно в этом ресторане.

Под влиянием доминирующей мотивации в процесс воспоминания включается практически вся функциональная система со всеми ее механизмами, в том числе и механизм оценки полученного результата.

Поразительная вещь! Интеллект оперирует гармоническим сочетанием главнейших факторов нейрофизиологической основы, необходимых для принятия решения,— всей настоящей ситуацией (голод, обстановка) и всем многообразием опыта прошлого, также связанного с удовлетворением пищевой мотивации. Казалось бы, эта система взаимодействия так далека от реальной мозговой структуры. Однако мы видим, что каждый элемент нашей интеллектуальной деятельности имеет вполне определенную нейрофизиологическую основу. Составляющие ее механизмы функциональной системы детально изучаются в нашей лаборатории (Анохин, Судаков, 1971).

Возвращаясь к принятию решения, которое согласно нашей схеме, является результатом предшествующего афферентного синтеза, мы должны на основании всего предыдущего признать, что на стадии афферентного синтеза осуществляется универсальный перебор извлеченных из памяти всех прежних результатов действия и всех прежних оценок этих результатов в соответствии с наличной доминирующей мотивацией. Именно для этой ответственной стадии, очевидно, и нужен процесс реверберации, мо-

билизирующий все сокровища кладовых нашей памяти. Следовательно, «принятие решения» является той процедурой, результат которой был признан после перебора «мнимых» результатов наиболее адекватным для данной обстановки.

С нейрофизиологической точки зрения этот процесс выбора единственной степени свободы состоит, очевидно, в непрерывном сканировании различных результатов, а эталоном для этого сканирования служит наличная в данный момент доминирующая мотивация. Экспериментами нашей лаборатории показано, что в ряде специфических случаев кортико-гипоталамическая реверберация может быть весьма отчетливой (Каграманов, 1965).

### **Акцептор результатов действия**

В этом разделе мы подходим к анализу такого нейрофизиологического аппарата, в котором скрестились главнейшие пути исторических поисков отгадки тайн человеческой психики («цель», «предсказание», «ошибка», «память», «ожидание» и многое другое). Как оказалось, все эти факторы имеют единый нейрофизиологический стержень, совершенно четко формирующийся в момент (или несколько позднее) принятия решения. Прежде всего я имею в виду нейрофизиологический аппарат предсказания, названный нами «акцептором результатов действия». Что это за аппарат? Какова его природа и каковы функции? Благодаря перебору всех признаков прошлых результатов и сличению их с данной доминирующей мотивацией этот аппарат сосредоточивает в себе все афферентные признаки того *конечного результата, по поводу которого было принято решение*.

Разберем пример. Если принято решение взять со стола стакан, то в появляющемся акцепторе результатов будут сконденсированы все относящиеся к действию признаки стакана: его внешний вид, вес, тактильные особенности, воспринимаемые кожными рецепторами, температура и т. д. Смысл этого аппарата, опережающего и предсказывающего свойства будущего результата, состоит в том, что в конце действия, т. е. после взятия стакана, должна быть получена вся информация о параметрах этого действия. Именно в этот момент в центральной нервной системе и происходит сличение результата, который прогнозировался в акцепторе результатов действия (взять стакан), с параметрами реально полученного результата (Анохин, 1949).

В момент сличения двух комплексов возбуждений наша нервная система осуществляет контроль результатов произведенного действия. Если сличение показало, что прогнозируемые параметры (предсказание) в акцепторе будущего результата полностью совпадают с параметрами реально полученного результата, то данное действие заканчивается, а его результаты получают «санкцию» и используются для формирования следующего этапа поведения. Если же выявляется несовпадение параметров реально полученного результата действия с запрогнозированными, то это рассогласование стимулирует построение и подбор новой программы действия, более точно обеспечивающей получение запрогнозированных результатов.

Поскольку все наше поведение представляет собой подлинный континуум результатов, больших и маленьких (Анохин, 1971), то практически такого рода сличения происходят в нервной системе непрерывно. Так, например, даже результаты таких незначительных действий, как открывание двери на лестницу, спуск по лестнице, посадка в автобус и т. д., оцениваются и формируют цепи получения последующих результатов. Однако и они могут быть дискретизированы на еще более мелкие результаты: например, постановка ноги на первую ступеньку автобуса, на вторую и т. д. Поразительно, что от каждого такого «маленького» результата наша нервная система должна непременно получить информацию, которая обрабатывается в соответствующем акцепторе результатов действия. Малейшее несоответствие результата прогнозу (мы оступились) — и мозг немедленно подбирает новое движение.

Из этой краткой характеристики функций аппарата, прогнозирующего результаты, становится понятным и его название «акцептор результатов действия». Латинское слово *accipere* содержит в себе два смысла — «принимаю» и «одобряю», которые представлены в функциях акцептора результатов действия (Анохин, 1949; Анохин, 1955).

В нашей лаборатории были проведены многочисленные эксперименты на клеточном и нейрохимическом уровнях для исследования того, как создается этот аппарат предсказания в каждом отдельном случае и каковы его функции в масштабе целой функциональной системы.

Теперь мы можем оценить значение этого аппарата для интеллектуальной функции человека и животных и определить его роль в изучении искусственного интеллекта. Прежде всего он является аппаратом предсказания, поскольку в нем прогнозируются свойства будущего, еще не полученного результата. Так как во всех наших действиях получение того или иного результата связано с заранее поставленной целью, то совершенно очевидно, что аппарат акцептора результатов действия практически является и аппаратом цели. Из этого положения вытекает, что цель в нашем понимании и в наших экспериментах не является чем-то изначальным, а подготавливается сложной работой нервной системы в стадии афферентного синтеза. Именно это обстоятельство позволяет выразить *цель как психологическое понятие* на языке нейрофизиологических механизмов и объективных причинных связей между *процессами, происходящими в головном мозге*.

Необходимо подчеркнуть, что при рассмотрении вопросов о предсказании и цели перед нами особенно четко вырисовывается философская сторона проблемы интеллекта и ее решения на основе концепции функциональной системы.

Действительно, совсем недавно одно лишь произнесение таких слов, как «предсказание», «цель», «целесообразное поведение», грозило физиологу обвинением в идеализме, в отрыве от материалистических принципов. Такая ситуация в физиологии исторически вполне объяснима, поскольку еще не был подготовлен соответствующий «научный климат» для материалистического разрешения проблемы цели и предсказания. Такой климат был создан в физиологии мозга в основном рефлекторной теорией.

И потому, естественно, мы мирились с поразительным парадоксом: каждый мыслящий человек прекрасно осознавал, что он ставит цель сделать «что-то» значительно раньше, чем реализует это «что-то». Но вместе с тем физиология мозга не имела средств для объяснения тех механизмов, с помощью которых мозг ставит перед человеком «цель» и с помощью которых он может предсказывать реализацию этой «цели».

С выявлением объективных нейрофизиологических закономерностей, обеспечивающих высшие функции интеллекта, отношение к этой проблеме радикально изменилось. В настоящее время, как можно было видеть, она успешно разрабатывается на основе принципов и категорий диалектического материализма, что приближает нас к реальному моделированию интеллекта.

Это становится возможным благодаря тому, что концепция функциональной системы позволяет охватить все те основные механизмы, которые помогают понять естественный интеллект.

Как показывают электрофизиологические исследования формирования подобного аппарата у человека, мы можем искусственно вводить и выводить из акцептора результатов действия любые новые компоненты (так называемое обогащение акцептора результатов действия), что значительно расширяет нашу власть, нашу способность воздействовать на интеллектуальную деятельность, в частности на процессы обучения.

### **Эволюция основных свойств интеллекта**

Возникает, однако, один вопрос: являются ли все описанные выше принципиальные механизмы интеллекта специфическими только для высших уровней развития животных или даже только для человеческого мозга? Это вопрос краеугольный, поскольку он очень тесно связан с другими вопросами, например, с такими: есть ли интеллект у животных, когда и у каких животных он появляется в процессе эволюции?

Интересную попытку сделали Фогель и другие, попробовав построить эволюционную модель «искусственного интеллекта», которая должна совершенствовать свои основные свойства путем разрастания примитивных механизмов в последующих «поколениях» (Фогель и др., 1969). Следует отметить, что в эволюционном моделировании накопление опыта осуществляется несколько иным способом, чем в интеллекте живот-

ных в процессе эволюции.

Отвечая на поставленные выше вопросы, мы должны прежде всего высказать основное положение, сложившееся у нас в результате многих лет работы над описанными выше свойствами интеллекта: ни одно из тех свойств мозговой деятельности, которые мы выше рассматривали как характерные признаки интеллекта, не появлялось внезапно, на каком-то «рубиконе», до которого этого свойства не было и после которого оно появилось.

Все эти свойства возникли уже на заре зарождения жизни, и все они являлись уже тогда частью динамической физиологической архитектуры. Более того, они явились *conditio sine qua non* самого развития живых существ.

Это может показаться странным, поскольку мы всегда описываем интеллект как выдающееся свойство живого, присущее по крайней мере самым высшим и совершенным представителям животного царства.

Однако наше недоумение немедленно рассеется, если мы представим себе, как происходило формирование интеллекта. Возьмем для примера *предсказание* будущих событий или результатов какой-либо деятельности, выполняемой некоторой четко очерченной функциональной системой.

При каких условиях внешнего и внутреннего мира животного возможно предсказание? Главным условием предсказания является то, что цепь событий, по поводу которой осуществляется предсказание, неоднократно повторялась в прошлом в определенных местах пространства в определенные моменты времени. Наш интеллект может предсказать, что после дня последует вечер, а после вечера ночь, только потому, что этот нерушимый ход внешних событий повторялся миллионы лет, и тогда, когда были живые существа, включая человека, и даже тогда, когда не было намека на их появление.

Здесь потребуется несколько отойти от конкретных нейрофизиологических процессов и обратиться к обобщениям более широкого характера.

Пространственно-временной континуум движения материи, как справедливо отмечал Планк, является абсолютным законом мира. Но этот закон действовал задолго до появления жизни на Земле. Иначе говоря, жизнь, т. е. живые существа, должна была *volens nolens* «вписаться» в рамки, задаваемые этим фундаментальным законом, и только при этом условии им было обеспечено выживание. Этот факт и привел к тому, что именно «вписанность», или *отражение* живыми существами пространственно-временного континуума, стала совершенно неизбежной предпосылкой предсказания.

Рассмотрев этот вопрос применительно к целому ряду биологических явлений, мы в свое время сформулировали *принцип опережающего отражения мозгом действительного ряда событий* во внешнем мире. Именно это свойство является изначальным свойством протоплазматических процессов даже у низших животных, для которых смена, например, сезонных явлений (лето — осень — зима — весна — лето) была неизменным условием жизни на протяжении миллионов лет.

Можно привести десятки примеров поразительной точности и целесообразности приспособления к этим абсолютным законам неорганического мира.

В сущности, открытие И. П. Павловым условного рефлекса было открытием *опережающих отражений* внешнего мира в высокоспециализированном субстрате — в нервной системе. В самом деле, когда в ответ на звонок у собаки выделяется слюна, то это происходит совсем не потому, что слюна должна «переваривать звонок», а потому, что в *будущем* появится пища, которую надо переваривать. Следовательно, в силу повторения последовательности определенных воздействий внешнего мира мы создали линию облегченных реакций, в которых достаточно первого толчка, чтобы химическая реакция протоплазмы, подобно реакции бикфордова шнура, распространилась по нервной системе в будущее, опережая последовательное развитие внешних событий.

Из разнообразных примеров и приведенных рассуждений становится ясным, что «предсказание» как феномен изолированного интеллектуального акта имеет глубокие исторические корни. На высшем этапе эволюции органом этого опережающего процесса стала нервная система. Именно она в сотни раз обострила и ускорила опережающие

процессы, и именно благодаря этому мы можем совершить почти фантастическое путешествие в будущее в ответ на какой-либо толчок или сигнал из внешнего мира.

Эволюция, начавшаяся с примитивного протоплазматического «предсказания», усовершенствовала этот процесс в материальных явлениях мозга до такой степени, что мозг стал органом, который в каждый данный момент своей деятельности сочетает в себе прошлое, настоящее и будущее.

Все это не является фантазией нейрофизиолога. Микроэлектродный метод дает нам возможность установить, что некоторые нейроны, испытывая *настоящее* раздражение, включают накопленный в *прошлом* опыт и одновременно с этим формируют процессы, содержащие в себе качества того результата действия, который будет получен только в *будущем*. В нашей лаборатории эти нервные клетки мозга мы назвали «нейронами трех времен».

Резюмируя обсуждение проблемы «предсказания», мы должны подчеркнуть, что «предсказание», выявляемое на высших этапах интеллектуальной деятельности, есть продукт наиболее совершенного развития того прототипного процесса, который проявляется уже в опережающем протоплазматическом отражении действительности.

Однако возвратимся к естественному развитию событий при формировании поведенческих актов на высшем уровне. Процесс афферентного синтеза, как мы уже говорили, заканчивается принятием решения, являющегося итогом перебора возможных результатов, органически связанных в прошлом с данной мотивацией. Таким путем осуществляется одно из самых замечательных явлений в активности мозга: *формирование на уровне нервной системы модели всех признаков и свойств будущего полезного результата, в связи с которым и ради которого развивались процессы афферентного синтеза*. Это и есть цель.

### Заключение

Приведенные в этой статье соображения убеждают нас в том, что для познания основных свойств естественного и искусственного интеллекта необходимо идейное перевооружение всей современной нейрофизиологии, выработка новых методических и методологических подходов. И действительно, поиски наиболее характерных черт искусственного интеллекта показали, что нейрофизиология, строящаяся только на традиционной, по преимуществу аналитической, основе, не может надеяться на успех в решении данной проблемы.

В свое время в связи с возникновением кибернетики крупнейший французский физиолог А. Фессар писал, что наука вступила в эпоху обратных связей (Fessard, 1953-1954). Это было совершенно верно. Но к этому надо добавить вступление нейрофизиологии в эпоху более широкого и глубокого синтеза нейрофизиологических и поведенческих исследовательских задач, нового синтеза данных, добытых в отдельных биологических науках. Прежде всего здесь следует сказать об усовершенствовании системного подхода, который содержал в себе все возможности для изучения самых высших форм деятельности мозга: «принятия решения», «цели», «предсказания», «интеллекта» и др.

Предлагаемый в данной статье подход, опирающийся на теорию функциональной системы, как нам кажется, приближает нас к решению все еще загадочных проблем интеллекта. Во всяком случае, нам стали ясны некоторые общие аспекты проблемы, открылся доступ к конкретному научному и экспериментальному исследованию тех аспектов, которые еще совсем недавно были прерогативой психологии, а часто оказывались основой для идеалистических интерпретаций.

И здесь при изучении исторических предшественников интеллекта были сделаны серьезные обобщения, позволившие понять непрерывную эволюцию интеллекта. Главнейшие свойства интеллекта — афферентный синтез, постановка цели, принятие решения и оценка полученного результата, предсказание и обратная афферентация, санкционирующая получение полезного результата. Эти синтетические процессы имели большую предысторию: они развивались из тех примитивных форм, которые сложились уже на заре жизни на нашей планете.

Эта общность архитектуры поведенческого акта и есть тот исторический фактор, благодаря которому жизнь и мозг развились до высшего этапа — человеческого интеллекта.

Исследования последних лет еще более убеждают нас в успешном применении теории функциональной системы к решению вопросов интеллекта. Ее конкретные синтетические узловые механизмы, выросшие на системной основе, дают возможность перекинуть «концептуальный мост» между нейрофизиологией, психологией и теми проблемами, которые возникают на путях познания тайн интеллекта.

Итак, диалектический материализм еще раз получает доказательство того, что интеллект, сознательная деятельность и активное преобразование самого приспособления к внешним факторам являются истинным продуктом исторического развития от материи к сознанию на основе фундаментальных законов неорганического и живого мира на нашей планете.

#### ЛИТЕРАТУРА

- Анохин П. К.* Проблема центра и периферии в современной физиологии нервной деятельности. В кн.: Проблема центра и периферии в физиологии нервной деятельности. Горький, 1935, с. 9.
- Анохин П. К.* Узловые вопросы в изучении высшей нервной деятельности. В кн.: Проблемы высшей нервной деятельности. М., 1949, с. 9.
- Анохин П. К.* Особенности афферентного аппарата условного рефлекса и их значение для психологии. — Вопр. психол., 1955, № 6, с. 16.
- Анохин П. К.* Принципиальные вопросы общей теории функциональных систем. М., 1971.
- Анохин П. К., Судаков К. В.* Нейрофизиологические механизмы голода и насыщения. — Успехи физиол. наук, 1971 т. 1, № 11, с. 3.
- Аттли О. М.* Механизмы процессов мышления (заключительная речь). В кн.: Самоорганизующиеся системы. М., 1964, с. 430.
- Каграманов К. М.* О холинергических и адренергических механизмах деятельности головного мозга. Автореф. канд. дис., М., 1965.
- Маккалок У., Питтс У.* Логическое исчисление идей, относящихся к нервной активности. В кн.: Автоматы. М., 1956.
- Месарович М.* Теория систем и биология: точка зрения теоретика. В кн.: Системные исследования. М., 1970, с. 137.
- Минский М.* На пути к искусственному интеллекту., 1961, 49.
- Павлов И. П.* Лекции о работе больших полушарий головного мозга. — Полн. собр. трудов, т. 4. М.—Л., 1947.
- Уотерман Т. Х.* Теория систем и биология: точка зрения биолога. В кн.: Системные исследования. М., 1970, с. 164.
- Фогель Л., Оуэнс А., Уолли М.* Искусственный интеллект и эволюционное моделирование. М., 1969.
- Fessard A.* Points de contact entre neurophysiologie et cybernetique. — In: Structure et evolution des techniques, Paris, 1953-1954, N 35-36.
- McCay D. M.* Operational aspect of intellect. — In: Mechanization of thought processes, 1959, vol. 1.
- McCulloch W. S.* Logic and closed loops for a computer junket to mars. — In: Neural Networks. Symposium. Ed. by E. R. Caianiello. N. Y., 1968.
- Rosenblatt F.* Principles of neurodynamics. Washington, 1962.
- Sadovsky V. N.* The history and perspectives of the systems approach development and general systems theory. — In: International congress of the history of sciences (18—24 August, 1971).