

Модель памяти, обеспечивающей вероятностное прогнозирование.

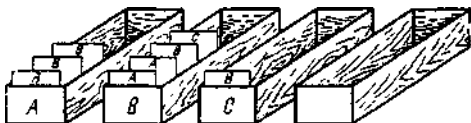
И. М. Фейгенберг

В вероятностно организованной среде информация о предстоящих событиях и о вероятности их наступления может быть извлечена из прошлого опыта. Память хранит информацию о событиях прошлого и о вероятностном (частотном) соотношении между событиями. Биологически одно из важнейших назначений памяти -- реализация вероятностного прогнозирования и преднастройки к действиям в ситуации, адекватной наиболее вероятной предстоящей ситуации.

Как можно представить себе логическую структуру памяти, способной из вероятностной организации прошлого опыта строить вероятностный прогноз будущего? Модель такой памяти была описана нами^{1 2}. Представим себе среду, в которой функционирует память, в виде последовательности событий А, В, С, ..., следующих в случайном порядке с определенными вероятностями. Модель памяти можно условно изобразить в виде каталожного шкафа с выдвигаемыми ящиками.

Пока система не приобрела опыта, память ее пуста — это шкаф с пустыми каталожными ящиками. Обозначим буквой А первое событие (явление, сигнал), с которым встретилась система, и соответственно присвоим индекс А одному из ящиков модели. Тем самым в память записано, что событие А имело место в опыте системы. Пусть и в дальнейшем особый ящик с соответствующим индексом заводится для каждого впервые встретившегося события.

Если, например, после события А наступило событие В, то индекс В присваивается следующему (пустому) ящику. Кроме того, карточка В ставится в ящик А — фиксация в памяти того, что В произошло после А. И в дальнейшем карточка с записью каждого происшедшего события (в какой бы раз оно ни произошло) ставится в ящик того события, которое непосредственно предшествовало ему. При этом каждая новая карточка (т. е. карточка о самом последнем событии) помещается впереди ранее поставленных. Сформированная таким образом память уже может обеспечить системе возможность вероятностного прогнозирования предстоящих событий и, следовательно, возможность преднастройки — подготовки к действиям, соответствующим прогнозируемым событиям. В простейшем случае прогноз осуществляется следующим образом. После того как произошло некоторое событие (например, В), в соответствующем ему ящике (В) подсчитывается доля различных находящихся в нем карточек (А, В, С, ...). Доля карточек А (по отношению ко всему числу карточек в ящике) характеризует вероятность, с которой прогнозируется наступление события А; доля карточек В — вероятностный прогноз наступления события В и т. д. Таким образом, для прогнозирования используется вся память, накопленная за всю «жизнь» системы. В частном случае, если непосредственно после события А всегда следовало только определенное событие (например, В), ящик А будет заполнен только карточками В. И тогда событие В будет прогнозироваться после А с вероятностью, равной единице, т. е. система будет «абсолютно уверена», что за событием А последует именно В.



¹ Фейгенберг И.М. Мозг, психика, здоровье. Изд. «Наука», Москва, 1972.

² Feigenberg J. Wahrscheinlichkeitsprognostizierung im System der zielgerichteten Aktivität. AFRA-Verlag, Butzbach-Griedel, 2000.

Фейгенберг И.М. Вероятностное прогнозирование в деятельности человека и поведении животных. Изд. НЬЮДИАМЕД. Москва 2008.

Рис. 1. В памяти зафиксирована последовательность событий *ABCBBABAB*

Мы рассмотрели прогноз, основанный на использовании опыта, накопленного в памяти за всю «жизнь» системы (в нашей модели — на привлечении всех карточек данного ящика). Однако «память всей жизни» оказывается весьма ненадежной, если система живет в среде, вероятностные характеристики которой со временем изменяются. Чтобы сделать прогнозы хоть сколько-нибудь соответствующими изменившейся среде, система должна «прожить» в этой среде отрезок времени, соизмеримый с уже прожитой ранее «жизнью», т. е. накопить достаточный новый опыт. Система оказывается «косной», плохо адаптирующейся к изменяющимся внешним условиям.

При этом чем «старше» система, т. е. чем больше ее старый опыт, тем труднее она приспосабливается к изменяющейся среде. Опыт, приобретенный за последнее время, занимает все меньшее место по сравнению с накопленным за долгое время старым опытом.

Чтобы устранить этот дефект, можно попытаться осуществить вероятностное прогнозирование, опираясь не на «память всей жизни», а лишь на опыт последнего периода.

Если произошло событие *A*, то модель просматривает первые *N* карточек в ящике *A* и подсчитывает, какую часть их составляют карточки событий *A*, *B*, *C* и т. д. В соответствии с полученными величинами и прогнозируются вероятности возникновения событий *A*, *B*, *C*,... и осуществляется преднастройка к соответствующим действиям. Весьма важен вопрос о рациональном выборе числа *N*. При $N = 1$ прогноз носит не вероятностный, а жестко predetermined характер: всегда предсказывается с вероятностью $P = 1$ (однозначно) то событие, которое в последний раз следовало за событием *A*. При $N = 1$ хорошее прогнозирование будет обеспечено, только если вслед за *A* всегда следует одно и то же событие.

Система, использующая для прогнозирования лишь недавний опыт, способна адаптироваться к изменяющимся условиям. В частности, ее прогнозирование достаточно для выработки классического условного рефлекса и его угашения.

Но в прогнозировании такой системы есть еще весьма существенные недостатки.

При небольшом *N* система будет очень «доверчивой» в своем прогнозе; влияние на прогноз случайного, но недавно встретившегося события будет значительным; модель будет быстро менять прогноз даже под влиянием случайных изменений среды.

При слишком большом *N* модель окажется слишком «косной» в своем прогнозе, недостаточно чуткой к изменениям вероятностной характеристики среды.

При *N*, равном числу всех карточек в ящике (т. е. если модель обладает «бесконечно большой» памятью — в пределах всей ее «жизни»), вероятностный прогноз (как уже говорилось выше) будет достаточно хорошим лишь до тех пор, пока будут оставаться стабильными вероятностные характеристики «среды обитания» модели. Если же вероятностная структура среды изменится, модель начнет выдавать неверные прогнозы и будет лишь медленно приспосабливаться к новой среде.

Итак, вероятностный прогноз оказывается неточным как при слишком малом *N* («легкомысленная», «доверчивая» модель), так и при слишком большом *N* («косная», «скептическая» модель).

Рациональный выбор *N* зависит от того, в какой среде работает модель — как часто меняются вероятностные характеристики этой среды.

Таким образом, вероятность, с которой рассматриваемая система (опирающаяся только на опыт последнего периода) прогнозирует событие *B* в случае наступления события *A*, равна n_B/N , где n_B — число карточек *B* среди первых *N* карточек в ящике *A*. Система подготавливается к действиям, соответствующим событиям *A*, *B*, *C*,..., в соответствии с величиной вероятностного прогноза n_A/N , n_B/N , n_C/N , ...

Система будет работать гораздо лучше, если, в большей мере опираясь на опыт последнего периода, она вместе с тем не будет забывать и о событиях далекого прошлого.

Пусть карточки в картотеке имеют некоторый «весовой коэффициент давности» события: лучше помнится то, что произошло недавно. Первые N_1 карточек, стоящие в ящике (недавние события), имеют коэффициент α_1 , следующие N_2 карточек (более давние события) имеют коэффициент α_2 , меньший, чем α_1 . Дальнейшие N_3 , карточек (еще более давние события) — еще меньший коэффициент α_3 и т. д. (рис. 2).

При такой структуре модели в величину вероятностного прогноза события В после наступления А входят доля карточек В среди первых N_1 карточек в ящике А с коэффициентом α_1 , доля карточек В среди следующих N_2 карточек с коэффициентом α_2 , доля карточек В среди следующих N_3 карточек с коэффициентом α_3 и т. д. Вероятность того, что наступит событие В при условии, что уже наступило событие А, будет таким:

$$P(B/A) = \frac{\alpha_1 \times \frac{n_{1B}}{N_1} + \alpha_2 \times \frac{n_{2B}}{N_2} + \alpha_3 \times \frac{n_{3B}}{N_3} + \dots}{\alpha_1 + \alpha_2 + \alpha_3 + \dots} = \frac{\sum_i \alpha_i \frac{n_{iB}}{N_i}}{\sum_i \alpha_i}$$

Это и есть вероятностный прогноз наступления события В при условии, что уже произошло событие А. При такой организации памяти становится понятным уже не только вероятностное прогнозирование и, в частности, классический условный рефлекс, но и еще одно явление.

Представим себе, что собака в определенных условиях после звонка всегда (в течение достаточно долгого времени) получала пищу. В результате после звонка у собаки начинается усиленное выделение слюны - результат того, что прием пищи прогнозируется с большой вероятностью. Однако в очередной раз после звонка собака не получила пищу, а последовало болевое раздражение лапы. То же повторилось и в следующий раз после звонка, а затем еще несколько раз. И вот в ответ на звонок собака отдергивает лапу, а усиления слюноотделения больше не наблюдается. Собака ведет себя так, как будто она забыла, что ее кормили после звонка, но помнит, что после звонка было больно лапе - собака переучилась. Однако опыт показывает, что собака ничего не забыла. Если оставить такую собаку на достаточно длительное время в покое - не кормить после звонка и не ударять током - а затем привести в лабораторию и включить звонок, у нее может возникнуть усиленное слюноотделение. Собака как бы вспомнила, что ее когда-то кормили после звонка, и как бы забыла более недавние события - болевое раздражение после звонка.

Похожие явления можно наблюдать и у человека. В квартире во время ремонта перенесли выключатель - он был слева от двери, а оказался справа. Первое время хозяин, заходя в квартиру, по привычке тянется в темноте налево. Но вскоре он переучивается и правильно ищет выключатель справа. И тут хозяин уезжает в отпуск. Вернувшись через месяц, он снова по ошибке тянется налево, как говорят, «по старой памяти». Та же картина - человек как бы забыл более недавний кратковременный опыт и как бы вспомнил ранее забытый более старый, но длительный опыт.

Такое «забывание» старого при изменении среды и «вспоминание» после перерыва объясняется тем, что со временем «весовой коэффициент давности» изменяется. Недавние события могут «перевешивать» даже длительный, но более старый опыт благодаря их высокому весовому коэффициенту давности. На рисунке 2 у нас весовой коэффициент менялся скачками, дискретно. Но естественнее представлять себе, что он меняется плавно, непрерывно (рис.3). В первое время недавний опыт (площадь S_2) сильнее старого опыта (площадь S_2 больше площади S_1) — и старый опыт как бы «забыт». Через некоторое время после обучения старый опыт уже перевешивает недавний кратковременный опыт (площадь S'_1 больше площади S'_2) При повторном обучении требуется уже меньше времени для восстановления вроде бы забытого навыка (сумма площадей S_3 и S''_2 больше площади S''_1 , соответствующей старому длительному опыту). Повторное обучение требует тем меньше времени, чем быстрее оно проводится после первичного и чем длительнее было первичное обучение.

В описанной выше модели вероятностное прогнозирование осуществляется по отношению только к одному событию, которое непосредственно предшествовало прогнозируемому событию. Ясно, однако, что такой прогноз не может быть достаточно надежным в среде, в которой имеются связи не только между событиями, стоящими рядом, но и между более отдаленными. Так, например, событие *D* может следовать с большей вероятностью за событием *C*, если перед *C* было *A*, и с малой вероятностью, если перед *C* было *B*. В описанной же выше модели прогноз вероятности наступления события *D* после *C* осуществляется без учета того, какое событие предшествовало *C* (в ящике *C* не содержится информации о том, что предшествовало событию *B*).

Наряду с ящиками, на передней стенке которых записано одно какое-либо событие (ящики первого порядка), заведем в нашей модели еще ящики второго порядка — такие, на передней стенке которых записаны два последовательных события - *AA*, *AB*, *AC*,... *BA*, *BB*, *BC*,..., *CA*, *CB*, *CC*,... Карточка *D*, поставленная, например, в ящик *AC*, означает, что событие *D* произошло после события *C*, которому предшествовало событие *A*. В приведенном выше примере карточка *D* будет часто встречаться в ящике *AC* и редко -- в ящике *BC* (рис.4).

Ящики второго порядка, конечно, обеспечивают лучшее вероятностное прогнозирование, чем ящики первого порядка. Еще более надежное прогнозирование обеспечивают ящики третьего порядка, на которых записаны три последовательных события.

Ящики второго и более высоких порядков не только обеспечивают лучший прогноз, чем ящики первого порядка, но и делают возможной выработку «ситуационных» условных рефлексов. Представим себе такую ситуацию. Физиологи *A* и *B* работают на одних и тех же собаках, но в разное время. Физиолог *A* подкрепляет звонок пищей, а физиолог *B* - легким уколом лапы. В первом случае у собаки возникает пищевая реакция, во втором - оборонительная.

Если звонка нет, то на появление физиологов *A* или *B* собака не реагирует ни пищевой, ни оборонительной реакцией. На звонок же она после выработки условных рефлексов реагирует по-разному. Если перед звонком (*C*) она видела физиолога *A*, то, не дожидаясь подкрепления -- пищи (*D*), отвечает пищевой реакцией. Если же перед звонком (*C*) она видела физиолога *B*, то, не дожидаясь подкрепления - укола (*E*), реагирует отдергиванием лапы. В модели (рис. 4) это достигается благодаря тому, что карточки *D* (пища) попадают в ящик *AC*, а карточки *E* (укол) - в ящик *BC*. Поэтому при возникновении комплекса событий *A* и *C* прогнозируется высокая вероятность возникновения события *D*; если эта вероятность $P(D/AC)$ близка к 1, то наступает пищевая реакция. При возникновении же комплекса событий *B* и *C* прогнозируется высокая вероятность возникновения события *E* - если $P(E/BC)$ выше некоторого «порогового» уровня, наступает оборонительная реакция.

Наличие ящиков высоких (второго и выше) порядков позволяет моделировать не только классические условные рефлексы, описанные И. П. Павловым, но и условные рефлексы, названные Конорским условными рефлексами второго типа, об отличии которых от классических условных рефлексов уже было рассказано выше (в главе «Рефлекс или активное действие?»).

Модель с ящиками первого порядка достаточна для моделирования классических условных рефлексов, но не может обеспечить осуществление условных рефлексов второго типа.

Иначе обстоит дело, если использовать ящики высших порядков. Присутствие карточки *F* в ящике *EC* (рис.5) моделирует след в памяти того, что событие *F* произошло после *C*, которому предшествовало *E*. Из сказанного выше ясно, как ящик *EC* может быть использован для вероятностного прогнозирования предстоящих событий, в частности для предсказания вероятности наступления события *F*.

Но ящики второго (и более высокого) порядка могут быть использованы еще и другим образом. Под каждой буквой (в нашем примере *F*, *E*, *C*) можно понимать любое событие, воспринимаемое животным. Таким событием может быть и внешнее явление, и собственное

действие животного - сгибание лапы, лай и т. п. Пусть F означает получение мяса, E -- стук метронома, C -- лай. Послышался стук метронома (E). Что надо сделать, чтобы с наибольшей вероятностью получить мясо? В памяти-картотеке просматриваются ящики EA , EB , EC ... , где A , B , C ,...-- различные действия собаки. В каждом из этих ящиков подсчитывается вероятность наступления желаемого события (получение мяса). Таким образом, используя ящики второго порядка, можно не только осуществлять вероятностное прогнозирование внешних событий, но и строить планы собственных действий (C), приводящих с наибольшей вероятностью в заданных условиях (произошло E) к желаемому результату (F). Еще лучше эта задача решается с помощью ящиков более высоких порядков. Ящики n -го порядка позволяют строить планы действий, состоящие из $n - 1$ шагов, приводящих с наибольшей вероятностью к желаемому результату в заданных условиях. С помощью ящиков n -го порядка можно строить и более короткие планы действий -- из $n - k$ шагов, но зато при этом k первых элементов будут использованы для более точного прогноза. Чем более высокого порядка ящиками располагает модель, тем точнее она осуществляет вероятностное прогнозирование, тем более длинные планы действий можно с ее помощью строить. Но это дается ценой значительно большей громоздкости модели и более длительного «перебора» при выборе из памяти. Число ящиков в модели сильно увеличится при возрастании их порядка (n).

Нетрудно сообразить, что при том же числе m возможных событий (т. е. таких событий, которые наступили хоть раз) модель может содержать m ящиков первого порядка, ящиков второго порядка -- m^2 , ящиков третьего порядка -- m^3 , ..., ящиков n -го порядка -- m^n . В силу этого в простой среде пользоваться ящиками высоких порядков нецелесообразно: процедура прогнозирования становится громоздкой и длительной, а выигрыш в точности может быть малым или даже совсем отсутствовать, если, например, среда представляет собой бернуллиевскую последовательность -- то есть такую случайную последовательность, где вероятность очередного её члена не зависит от того, каковы были предыдущие члены (простым примером может быть вероятность выпадения той или иной стороны брошенной монеты). Итак, стремление улучшить вероятностное прогнозирование приводит к тому, что модель памяти становится чрезвычайно громоздкой. А нельзя ли несколько разгрузить память, не ухудшая результатов вероятностного прогнозирования? Нельзя ли запоминать не все в равной степени, а кое-что и совсем не запоминать? Вслед за этим возникает и другой вопрос: какие события (сигналы) должны более основательно запоминаться для вероятностного прогнозирования?

Чтобы ответить на этот вопрос, подумаем, от чего зависит глубина запоминания того или иного события (сигнала)? Она зависит от многих факторов, но здесь мы хотим обратить внимание лишь на один из них - на «степень удивительности», степень неожиданности, несоответствия происшедшего события тому, что прогнозировалось.

Поясним примером. Вечером осеннего дня мне трудно точно вспомнить, не встречал ли я сегодня на улице человека в сером пальто. А вот если со мной в автобусе ехал человек в индийской чалме, я это буду хорошо помнить вечером. Встреча с человеком в чалме в московском автобусе маловероятна (и неожиданна для меня), а с человеком в сером пальто -- весьма вероятна в осенний день.

В главе «Вероятностное прогнозирование в педагогике» описаны результаты экспериментов, которые выявили четкую зависимость степени запоминания слов от того, с какой вероятностью прогнозировалось появление именно этого слова. Чем более неожиданным было слово, тем лучше оно запоминалось.

Пусть теперь и наша модель запоминает не все, что поступит в нее, с одинаковой глубиной. В нашем предыдущем рассуждении результат работы памяти - вероятностный прогноз - служил только для преднастройки других рабочих органов. Теперь же этот прогноз будет использован и для организации самой памяти.

Если произошло событие A , то с помощью карточек ящика A подсчитывается P_A -- вероятность того, что снова произойдет A , точно так же P_B -- вероятность того, что наступит событие B , ..., P_K - вероятность того, что наступит событие K . Это и есть вероятностный

прогноз. Предположим, что после этого наступило событие J , вероятность наступления которого, согласно прогнозу, равнялась P_j . Теперь по установленному порядку следует поставить карточку J в ящик A . Но при этом карточке придается различный «вес» в зависимости от степени неожиданности события J , она ставится с коэффициентом $1 - P_j$. Назовем эту величину коэффициентом неожиданности. Если в прошлом опыте после A всегда следовало J , то $P_j = 1$ (т. е. хранящаяся в памяти «модель среды» точно соответствует среде и сигнал J ожидался с полной уверенностью). Тогда $1 - P_j = 0$, т. е. повторно наступившее событие J не запоминается, не загружает память. В жестко детерминированной среде модель перестает запоминать, как только научается безошибочно прогнозировать ход событий. Но если среда изменится и наступит неожиданное событие, оно врежется в память с максимальным коэффициентом неожиданности $1 - 0 = 1$. Каждое событие запоминается тем сильнее, чем более неожиданным было его возникновение в данный момент. К вопросу о том, что высоко ожидаемое событие не требует каждый раз фиксации в памяти (и о том, как это сказывается в патологии), мы еще вернемся в главе «Вероятностное прогнозирование в патологии».

Живому организму или другой сложной системе, функционирующей в вероятностно организованной среде, вероятностное прогнозирование полезно тем, что позволяет осуществлять преднастройку - подготовку к действиям в тех предстоящих ситуациях, наступление которых прогнозируется с наибольшей вероятностью. Исследования, проведенные в нашей лаборатории (М. А. Цискаридзе, В. А. Иванниковым и др.), показали, что именно благодаря преднастройке при наличии нескольких сигналов, на каждый из которых нужно было реагировать нажатием соответствующей кнопки, реакция совершается тем быстрее, чем с большей вероятностью прогнозировалось возникновение именно этого сигнала. Подробнее об этих экспериментальных исследованиях рассказано ниже – в главе «Экспериментальные исследования вероятностного прогнозирования». Многочисленные эксперименты показали, что при случайной последовательности сигналов, вероятность которых одинакова, время реакции растет с увеличением их числа. Однако странным казался тот факт, что удлинение времени реакции происходит лишь при увеличении числа таких сигналов (примерно, до десяти). Дальнейшее увеличение (в некоторых экспериментах число различных сигналов превышало тысячу) не ведет к еще большему удлинению времени реакции.

Это явление может быть понято, если предположить, что преднастройка осуществляется лишь по отношению к действиям, адекватным такой ситуации, возникновение которой прогнозируется с вероятностью, превышающей некоторую минимальную величину P_0 . Величина P_0 играет роль абсолютного порога вероятностного прогноза. К ситуациям, появление которых прогнозируется с вероятностью меньшей, чем P_0 , преднастройка не осуществляется. Если $P_0 \approx 1/10$ (а из экспериментальных данных вытекает, что это примерно так), то при наличии равновероятных сигналов вероятностный прогноз достигает пороговой величины P_0 лишь при числе сигналов до десяти, при большем числе равновероятных сигналов вероятностный прогноз появления каждого из них меньше порогового. Если же сигналы возникают с неодинаковой вероятностью, то среди них могут оказаться и сигналы с $P > 1/10$, но число их (независимо от общего числа сигналов) всегда будет меньше десяти. А это значит, что как бы велико ни было число возможных сигналов, при наличии порога P_0 организм упрощает ситуацию так, как будто число их не больше чем $1/P_0$, (т. е. в приводимом примере не больше десяти). Наличие порога P_0 позволяет при большом общем числе возможных событий принимать во внимание лишь небольшую часть -- не более $1/P_0$ (т.е. в приводимом примере меньше десяти) наиболее вероятных из них (рис.6).

Различие в преднастройке к действиям на сигналы наблюдается лишь тогда, когда различие прогнозируемой вероятности появления этих сигналов достаточно велико, больше некоторого ΔP - дифференциального вероятностного порога. В частности, для выработки условного рефлекса не необходима выработка прогноза с $P=1$ (такая ситуация создавалась И.П.Павловым в эксперименте); необходимо, чтобы

прогнозируемая вероятность подкрепления была достаточно близкой к единице ($P \geq 1 - \Delta P$).

Таким образом, в описанной модели хранение сведений о прошлом и способ пользования ими делает возможным вероятностное прогнозирование предстоящих событий и построение плана действий, которые в заданных условиях с наибольшей вероятностью приведут к желаемому результату.

Память и вероятностное прогнозирование являются лишь частью большой системы (будь то живой организм или техническое устройство). Здесь в каждый момент любому событию A , B , C ..., которое может произойти, приписывается определенная вероятность $P(A)$, $P(B)$, $P(C)$ и т. д. Учитывая этот вероятностный прогноз, а также значимость прогнозируемого события, исполнительные системы осуществляют преднастройку -- подготовку к действиям тех органов, которые должны будут работать в прогнозируемой ситуации для достижения заданной цели (или для удовлетворения потребностей организма). Степень преднастройки каждого исполнительного органа тем больше, чем вероятнее и чем значимее наступление того события, при котором потребуется работа этого органа. В главе «Экспериментальное исследование вероятностного прогнозирования» мы вернемся подробнее к этому вопросу.

С позиций описанной структуры памяти стоит, может быть, рассмотреть и такое явление как импринтинг, описанное Конрадом Лоренцом. Маленькие, недавно вылупившиеся утята всегда следуют за своей мамой. Память маленьких утят еще почти не заполнена. И тот «ящик» (в нашей модели), который занят крупным движущимся объектом (еще нет различий деталей этого объекта) и собственным движением - приближением к этому объекту, заполнен «карточками» удовлетворения (согревание, насыщение, безопасность). Других «карточек» в этом ящике пока нет или почти нет. И эти «карточки», несмотря на их «малочисленность», являются почти единственными. И если в эксперименте утка-мама подменяется другим крупным движущимся объектом, маленькие утята следуют за ним - до тех пор, пока их память не заполнится более обширным опытом.

Мы описали здесь модель памяти в виде каталожного шкафа. Но эта система может быть реализована и использована в виде программы для компьютера. Она будет получать информацию и в соответствии с этим формировать (а если надо, то и изменять) свой вероятностный прогноз. Она будет самообучающейся.