

УЧЕНЕ, РАЗВИТИЕ И ЕВОЛЮЦИЯ

Джефри Елман

Едно от интригуващите свойства на живите същества, и особено на интелигентните живи същества, е умението за адаптация, т.е. промяна в резултат на опита, на преживяното. Много неща се променят в резултат на преживяното, разбира се. Планини се превръщат в хълмове или пък, поради ерозия, скали се превръщат в пясък. Но това са примери за промяна, довела до намаляване на сложността на системата. Интересното за промяната при живите същества е, че адаптацията обикновено води до по-голяма степен на сложност. Струва ми се, че това е важна разлика, която може би даже е едно от определящите свойства на живото.

Съществуват, естествено, множество различни форми на адаптация, които се извършват в различни мащаби във времето. Във всеки отделен момент ние се учим от опита, от преживяното. В течение на целия си живот също така се *развиваме*, а през погледа на много по-дългото съществуване на даден животински вид *еволюцията* е главната движеща сила на адаптацията.

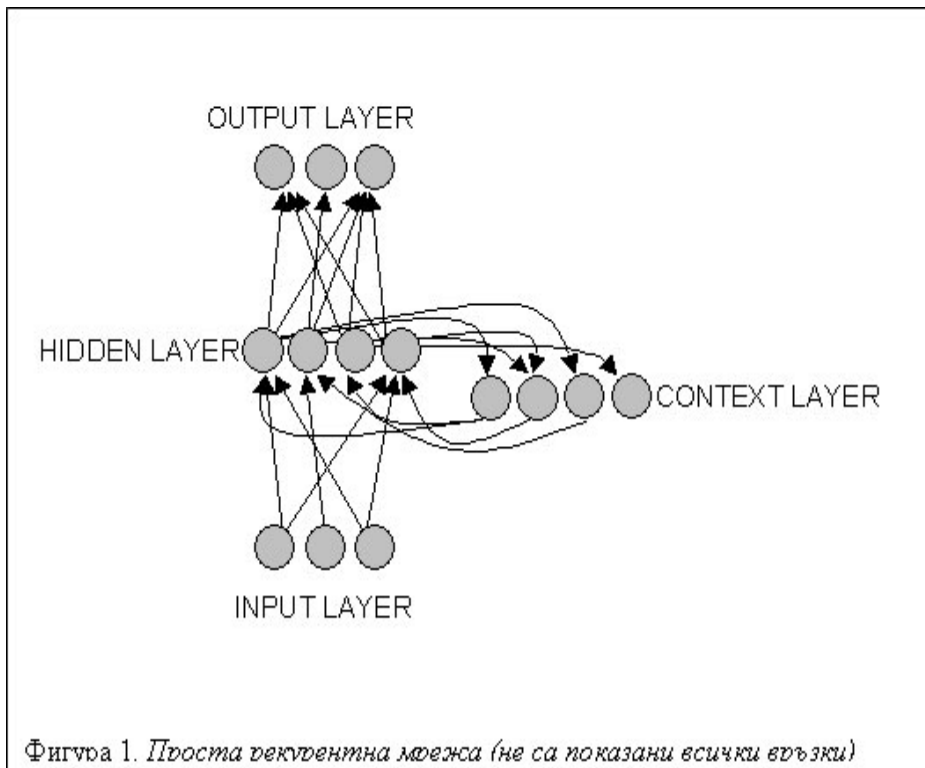
Отдавна се интересувам от адаптацията и промяната. Днес бих искал да говоря за някои резултати, до които достигнах чрез използването на компютърни симулации за изследване на ученето, развитието и еволюцията. Ще се постарая да се насоча към тези аспекти на работата, които са на едно по-високо ниво и, поради това, ще пропусна много от подробностите. Надеждата ми е, че по този начин ще можем да се съсредоточим върху резултатите.

Невронни мрежи

Ще започна с кратко обяснение за видовете компютърни модели, с които работя, тъй като те са относително нови и може би не всички тук са запознати с тях. Тези модели се основават на това, което наричаме "изкуствени невронни мрежи" (или "конекционистки модели") и представляват един нов подход в областта на изкуствения интелект. Този подход произтича от разбирането, изразено за първи път преди повече от 50 години в трудовете на психолога Доналд Хеб, че човешкият мозък притежава уникални изчислителни свойства. Те сериозно се различават от свойствата на дигиталния компютър, който обикновено се има предвид при метафоричното сравнение на компютър и човешки мозък. Преди около 15 години една група учени-когнитивисти започна да изследва изчислителните следствия на тези различия. Тези изследвания доведоха до разработването на нов подход в изкуствения интелект – *невронните мрежи*. Понякога тези мрежи са подробни и правдиви модели на реални мозъчни структури, но по-често (особено при моделирането на висши когнитивни функции) изкуствените невронни мрежи опростяват нещата и се опитват да обхванат по-висшите и по-абстрактните свойства на невронните изчисления.

Самите мрежи се състоят от набор от съвсем прости изчислителни елементи, които обикновено се наричат "възли" и в известен смисъл наподобяват неврони. Един-единствен възел сам по себе си почти нищо не може да направи, но има определено ниво на активация по подобие на честотата на разреждане на неврона, която съответства на степента му на възбуда. Съществуват също така множество връзки между възлите, които им позволяват взаимно да се възбуждат или потискат, в зависимост от естеството на дадена връзка. Мрежата, като цяло, постоянно се намира в състояние на активност, което се променя във времето - както поради множеството взаимодействия между възлите, така и в резултат на всяко външно въздействие върху мрежата.

Бих искал да опиша пред вас един по-конкретен тип невронна мрежа, която много съм използвал при моите изследвания. Тя се нарича **проста рекурентна мрежа** (Elman, 1990) . На Фигура 1 виждате пример за проста рекурентна мрежа.



Фигура 1. Проста рекурентна мрежа (не са показани всички връзки)

Една група от възли образува слоя на входа – те получават въздействия от външния свят на входа на системата. Може да си ги представяте като “очите” и “ушите” на мрежата. Друга група образува слоя на изхода, който също си взаимодейства със света - тях може да си представяте като “устата” или двигателните органи на мрежата. Освен това има и два вътрешни слоя от възли. Единият (т.нар. “скрит” слой) има за задача да бъде посредник между входа и изхода; за него опростено може да се мисли като за “мозъка” на мрежата. Скритият слой е свързан с друг набор от единици, който образува “слоя на контекста” и представлява един вид памет. Тази памет е много важен компонент, за да може мрежата да запомни предишни събития и да се учи на схеми, които имат продължителност във времето.

И настина, има много интересувачи ни видове поведение, които могат да се разберат само ако се разгледа тяхната история във времето. В случая с човешките същества езикът е една от най-очевидните сред тях. Отделните думи сами по себе си твърде малко ни говорят. Това, което ни позволява да общуваме, е начинът, по който думите се свързват, образуват изречения, абзаци и т.н. Но има и много други области, при които времето е от най-съществено значение, включително зрението, двигателната дейност и т.н.

Най-накрая трябва да ви кажа, че ентузиазмът около този тип невронни мрежи произтича най-вече от това, че имаме на разположение няколко много прости и същевременно много мощни процедури за обучението им. Или казано по-добре, това са процедури, които позволяват на тези невронни мрежи да се самообучават. Вместо да даваме на мрежите експлицитни напътствия или да им казваме правила за дадена (предметна) област, те се учат на базата на примери. Получават различни видове входна информация и им се казва как да реагират спрямо нея (дори в някои случаи може изобщо да не реагират). Да вземем един прост пример: представяте си, че дадена мрежа е получила набора от съответствия между входна информация и целеви изход, представени на Фигура 2.

<u>ВХОД</u>	<u>ПРАВИЛЕН ИЗХОД</u>
2	4
9	81
10	100
7	49
1	1
...	

Фигура 2. Пример за схеми за обучение – всяка схема се състои от вход, който се подава на мрежата. Правилният изход, който трябва да даде мрежата, се намира във втората колона.

Тези двойки от вход и целеви изход се представят на мрежата в случаен ред. Тя научава, че когато види “2”, правилната реакция е “4”, когато види “9”, правилната реакция е “81” и т.н. Това, което не ѝ се задава, е естеството на отношението между тях или правилото за получаване на коректната изходна реакция от входната информация. Но след като види множество примери, ние подаваме на мрежата нова информация на входа, например “12”, и следим какво ще каже тя. Ако се е научила да прави правилно обобщение, реакцията на изхода ще е “144”.

Интересното и вълнуващото при тази процедура, е, че мрежата трябва сама да научи правилото. Това, естествено, е валидно и за много от най-важните неща, които научаваме, особено като деца. Така например, ние научаваме езика най-вече чрез наблюдение на възраст от 4 до 5 години. Следващото, по-късно във времето, експлицитно обучение ни помага да направим фината настройка на нашите познания, но изглежда не играе съществена роля при най-основното изучаване на езика.

Темата днес е “Учене, развитие и еволюция”, затова нека ви дам няколко примера за това как изучаваме тези три движещи сили на адаптацията, като използваме простите рекурентни мрежи. Ще започна с един пример, свързан с обучението.

Учене на думи

Първият пример е свързан с ученето на думи. Приемаме за дадено, че думите са измежду първите неща, които трябва да научи малкото дете за езика си, тоест най-основния речников състав. Но тази задача се усложнява от два много важни и затрудняващи детето факта. Първо, то чува поток от звуци на входа, който не е разчленен по никакъв начин, подсказващ къде започват и завършват думите (къде са техните граници). Потокът от звуци е непрекъснат във времето, а думите (за разлика от напечатаните) не са отделени една от друга от “празни пространства” или паузи. Този факт представлява изненада за много хора, но на практика именно той е едно от нещата, които се оказват такъв сериозен проблем при автоматичното разбиране на реч от машини. Така че първото предизвикателство е да знаеш къде са думите в този непрекъснат поток от звуци.

Вторият проблем е, че макар в много езици да съществува онова, което безспорно можем да наречем “думи” (в разпечатката на тази лекция думите са оградени с интервали), редица други езици не се състоят от единици, които така охотно ще наречем думи.

Например, в турския език единиците са редици от смислоносещи морфеми, свързани последователно в дълги поредици. В някои езици тези поредици могат да представляват цели групи от думи или фрази. *Затова второто предизвикателство е, че определението за това какво е дума много се различава в отделните езици.* Горкото дете не само не знае къде са единиците в този поток от звуци, които трябва да научи, но не знае и какви са те - думи, морфеми или фрази.

Да видим сега как една обикновена невронна мрежа може да се справи с тези предизвикателства. Ще опростим донякъде задачата, но по начин, който не прави проблема по-лесен за решаване, а само улеснява моето обяснение. Започваме с един писмен текст от детска приказка. Той започва така: "Имало едно време едно момче и едно момиче, които живеели в един замък край морето. Те си играели с един дракон ..." и т.н. Нека направим две неща. Най-напред, да отстраним всички празни места между думите, така че сега текстът да изглежда като една непрекъсната поредица от букви: "Ималоедновременноедномомчеедномомичекоитоживееливединазамъккраймореето..." И така, той представлява непрекъсната звукова информация на входа, която чува едно бебе. Освен това, нека заменим всички букви с числови кодове (случайни поредици от нули и единици), защото всъщност това се възприема от възлите на входа (подобно на конфигурацията от активация, която може да получи даден неврон).

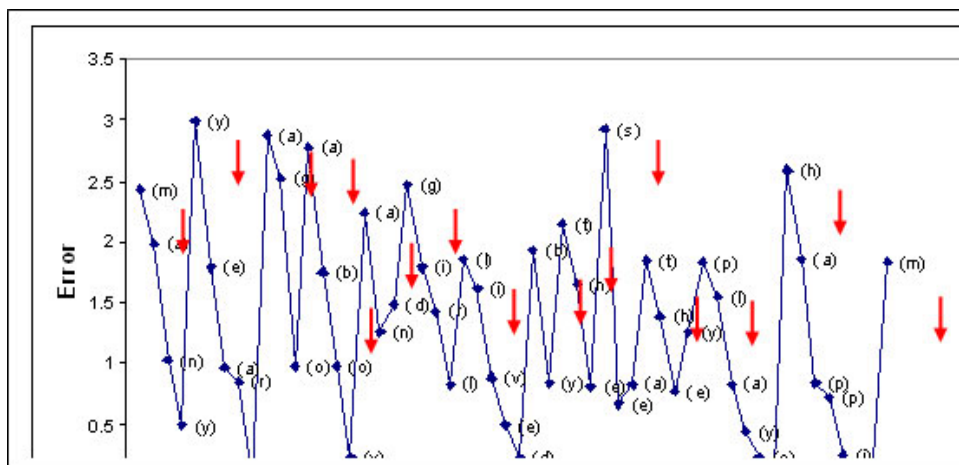
След това ще предоставим всяка буква (представена от числов код) една по една на мрежата. След като я прочете, мрежата ще трябва да отговори с определен изход. На Фигура 3 се вижда какъв изход се свързва с даден вход (тук, за ваше улеснение, те са представени в буквен вид).



Фигура 3. Начало на примерите за обучение. След прочитането на всяка буква

Виждате ли повтарящата се схема? От мрежата се очаква да отгатне коя ще е следващата буква на входа. С други думи, задачата ѝ е да *предскаже* следващата буква. На пръв поглед това може да ви изглежда непосилно. Как може мрежата да предсказва бъдещето? Всъщност, тя не може наистина да предсказва следващата буква със сигурност. Но трябва да се научи, че за някои поредици има по-голяма вероятност да се появят, както и че други са направо невъзможни. Например, след подаване на първоначалната буква "ц" мрежата следва да усвои, че за следващата буква е по-вероятно да е някаква гласна и е съвсем невероятно да е още веднъж буквата "ц".

Ако тя действително се научи да го прави, тогава може да се очаква да се случи следното. След като бъде обучавана на този текст известно време, можем да погледнем доколко добре се справя с предсказването. Това ще оценим, като измерим колко грешки прави мрежата в предсказанията си след всяка представена буква. Ако направим графика на грешките, ще видим схемата на Фигура 4. Буквата в скоби е тази, която мрежата се опитва да предскаже. Веднага се забелязва нещо очевидно, а именно, че грешката не е еднаква за цялата поредица. Всъщност, налице е повтаряща се схема на покачвания и спадове. Ако се вгледаме по-внимателно, ще забележим нещо любопитно – грешката се покачва в началото на думата, но с времето, тъй като мрежата чете все повече от буквите в думата, грешките намаляват. Това поведение е разбираемо. Колкото повече букви от дадена дума бъдат прочетени, толкова по-добри би трябвало да стават предсказанията за следващите, защото допълнителните букви ни дават допълнителна информация, която ограничава възможностите за последващия резултат.



Да предположим, че малките пеленачета извършват нещо подобно на онова, което прави мрежата: слушат пасивно речта наоколо и просто се опитват да предугадят какво следва. Ако правят така, то те скоро ще научат, че някои поредици от звуци вървят “заедно” и са по-предсказуеми. Също така би следвало да забележат, че има някои места в поредиците, които са твърде непредсказуеми. На Фигура 4 те са отбелязани със стрелки. Тези места всъщност бележат началото на това, което ще наречем думи. Детето, естествено, е съвсем като мрежата, в смисъл, че не знае първоначално нито къде започват или свършват думите, нито дори че има такива единици. Излиза обаче, че само като слушат и се опитват да предугадат, детето и мрежата имат на разположение информация, която им помага да разрешат този труден проблем. Преди няколко години Джени Сафрърн, Дик Аслин и Елиса Нюпорт докладваха резултатите от експеримент със 7-месечни бебета. Тук няма да описвам експеримента, само ще кажа, че децата от това изследване изглежда са използвали система за учене, която много наподобява системата, използвана от мрежата.

Това е пример за учене. Нека сега дам пример за развитие, което разбирам като протичащи промени, особено в ранна възраст, които са не толкова резултат от опита и преживяното, колкото от вътрешни изменения. Този пример също ще е свързан с езика и ще използва мрежа и задача, много подобни на току-що описаните.

Развитие: колко е важно да започнеш с малко

При тази симулация исках да видя дали мрежата може да бъде обучена на сложна граматична структура. Тя трябваше да се научи на изречения от типа “Момчетата, които Мери мрази, гонят кучетата”. Забележете, че макар в английския (и българския) това да е едно-единствено изречение, всъщност става дума за две неща - момчетата гонят кучетата и Мери мрази момчетата. Второто действие е поставено в така нареченото “подчинително наклонение”. То е нещо като мини-изречение, което ни дава допълнителна информация за момчетата, а именно, че Мери ги мрази.

Става ясно, че такива сложни структури са свързани с онова, което информатиците наричат “рекурсия”. Тя е нещо много важно при езиците за програмиране, както и при човешките езици. Рекурсията изисква даден елемент да е съставен от по-малки, които също са фрази със съществителни. (Един сходен и донякъде по-прост пример е да си представим какво става, когато получим отражение на образ в огледало и поставим друго огледало срещу първото... получава се привидно безкрайна “рекурсия” на образи.) Някои лингвисти и психолози като Ноам Чомски и Стивън Пинкър твърдят, че рекурсията е явление уникално за човешкия език и при това - че не може да се научи чрез примери. Вместо това, смятат те, рекурсията трябва да е вродена. Според тези учени, тя трябва да е част от общочовешките дарби.

Лично аз съм твърдо убеден, че биологичното начало играе много важна роля за това кои сме ние и какво можем да направим. Но освен това вярвам, че тези зададени биологични ограничения вероятно не приемат формата на генетични инструкции, които са уникални и присъщи единствено за неща като езика - например гени, които да са специализирани за него. (А има такива учени, които твърдят, че при човека има гени, предназначени специално за езика.) Вместо това считам, че е по-вероятно той да е продукт на голям брой малки изменения в гените, които могат да се открият и при много от най-близките ни роднини в животинския свят, и че взети заедно, тези изменения правят възможно явлението език. Нека ви дам един пример за това как малки изменения във времето могат да позволят да се стигне до обучение на рекурсия.

Целта на моята симулация беше да се види дали една проста рекурентна мрежа може да се обучи на сложни изречения с рекурсия. Първоначално се получи пълен провал. Мрежата трябваше да обработи голям брой изречения, някои от които виждате на Фигура 5. Част от тях са прости, но редица бяха сложни и с по няколко подчинени изречения. За мое удивление, мрежата изобщо не можа да се научи как да се справи. Опитах най-различни алтернативни подходи за период от около 6 месеца, но не можах да открия мрежа, която да постигне успех. Този непрестанен провал ми действаше доста обезсърчително и изглежда говореше в полза на казаното от Чомски и Пинкър - че рекурсията изисква вроден механизъм, който липсва на мрежата и не може да се научи чрез примери.

Момчетата, които харесва Мери, гонят кучетата.
Котката спи.
Момичето, което яде сандвича, е симпатично.
Кучетата ядат сирене.
Кучето, което харесва котките, които гонят мишките, вижда момчетата.
Момичетата съществуват.
.....

Фигура 5. Някои изречения, използвани при обучението на мрежата на рекурсия.

Тогава ми хрумна нещо. Мрежата, с която работех, приличаше на зрял, възрастен човек в смисъл, че макар и да започваше обучението без предварителни познания, цялата механика на модела вече бе напълно развита и в зряло състояние. Работната памет – под формата на рекурентни връзки – например беше напълно разработена от самото начало. Това бе точно обратното на положението на учещите език деца. Децата сами по себе си претърпяват драматични промени (мозъците им се променят) точно по времето, когато се опитват да научат езика. А това се различава доста от моите мрежи, които бяха статични и не търпяха развитие по време на обучението.

През ранното детство се променят много неща, но особено драматично нараства работната памет. Отначало тя е доста ограничена и постепенно се увеличава през първите години на човешкия живот. При моите мрежи работната памет е представена чрез рекурентни връзки. Затова реших да си направя един експеримент. Започнах да обучавам мрежи, които имат недоразвита работна памет. Рекурентните връзки бяха с доста шум (така, както връзките между невроните са с определено ниво на шум през първата част на живота). Така че през първата фаза на обучение на мрежите се подаваха всички тези много трудни изречения, но те бяха с ограничена работна памет, която им позволяваше да обработят само по няколко думи наведнъж, защото паметта за по-старите думи се влошаваше. Мрежите почти нищо не научиха от дългите изречения, но се научиха на простите, които бяха достатъчно кратки и можеха да влязат в ограниченията от време на “бешката” работна памет. След известен период постепенно започнах да намалявам шума и да увеличавам паметта. Най-накрая тя достигна нива както при възрастните. Към този момент вече мрежата имаше изумителната способност да обработва всички изречения, включително и най-сложните, с рекурсия.

Как е станало това? Отговорът е относително прост. Сложните изречения са били прекалено голямо предизвикателство в началото. Когато започна с ограничена работна памет, на мрежата ѝ бе спестено усилието да се справи с тях. Работната памет е послужила като един вид филтър, който ѝ е позволявал да види само най-простите изречения. Когато мрежата научила тези изречения, тя усвоила някои важни основни характеристики на езика, включително опростената граматика и граматичните категории. Тези познания са били от изключително значение за обучението с по-сложните структури. И така, когато работната памет се усъвършенствала така, че мрежата да може вече да “види” този тип изречения, тя вече е имала нещо като “скеле”, което ѝ е позволило да научи по-трудните факти относно рекурсията.

За мен този резултат представлява интерес по две причини. Най-напред, той ни подсказва едно ново обяснение за това защо децата учат езици толкова по-лесно от възрастните. Лингвисти като Чомски предполагат, че те имат в мозъка си т.нар. механизъм за усвояване на език, но че той изчезва по някакъв начин в периода на късното детство и вече не е на разположение на възрастните. Според този възглед, критичният период за научаване на даден език зависи от невронно оборудване, от механизма за усвояване на език, който съществува само в детството. Но нашите резултати подсказват алтернативно обяснение. Може научаването на език да е най-лесно в този период, защото то изисква определен ред в ученето от по-простите към по-сложните факти, което е възможно благодарение на ранните ограничения в процеса на човешкото развитие. С други думи, ранното научаване на език е свързано с *ограниченията* на мозъчните механизми в детските ни години, а не със специални устройства, които притежаваме само тогава.

Този резултат подсказва и отговора на една друга загадка. Обикновено за животните не е препоръчително дълго време да са незрели организми. Младите индивиди не могат да се защитават от хищници и други опасности, нито да се изхранват. Затова може да се предположи, че еволюцията ще работи в полза на животински видове, при които отделните представители се развиват бързо, а детството им е относително кратко. Ако е така, то изглежда странно, че тъкмо при човешката раса, която в много отношения е с висока степен на адаптация към средата си, се наблюдава период на начално развитие, много по-дълъг (в сравнение с общата продължителност на живота), отколкото при който и да е друг животински вид. Детството и юношеството при нас са доста удължени във времето, а това може на пръв поглед да ни изглежда като признак на лоша адаптация.

Резултатите от симулацията обаче сочат друго възможно обяснение (би следвало да спомена, че можем да мислим и за други отговори на тази загадка, включително проблеми, свързани с раждането на деца с големи мозъци). Тази симулация ни подсказва, че незрялостта в развитието може да има определени преимущества, свързани с изучаването на сложни понятийни области като езика. Някои проблеми може да изискват поэтапно усвояване, а ограниченията в развитието - да представляват механизъм, подпомагащ преминаването през тези етапи. Противно на нашите очаквания, ако бебетата имаха работна памет като тази на възрастните, то те може би щяха да имат повече трудности с научаването на езика. Тази хипотеза наричам още “колко е важно да започнеш с малко.”

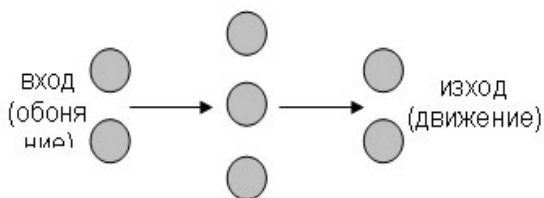
Еволюция на желанията и целите

Сега искам накратко да премина към последната симулация, която е свързана с еволюцията. Защо да се интересуваме от еволюцията? Ами, макар да мисля, че човешките способности за учене са много големи и да съм силно впечатлен от възможностите на невронните мрежи в тази посока, съществува една много ясна разлика между човешките бебета и невронните

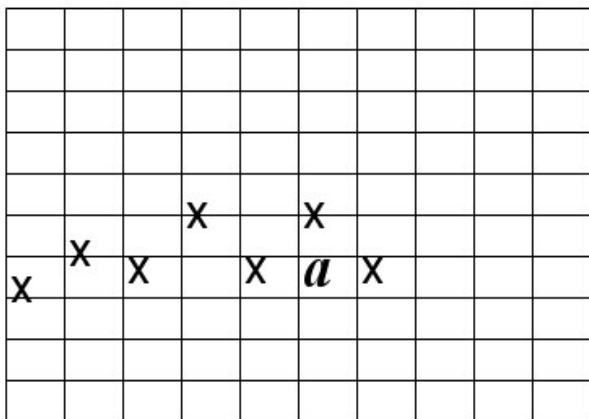
мрежи. Последните нямат собствена инициатива. Те усвояват много от средата си, но са пасивни по отношение на избора на това какво да се научи или дори кои среди са предпочитани. От друга страна, бебетата си имат собствени намерения. Имат си желания и потребности и те играят основна роля при формирането на онова, на което обръщат внимание и което, в крайна сметка, научават.

Но желанията и потребностите не могат да бъдат усвоени. Те са в истинския смисъл на думата част от това, с което ни е надарила природата. Продукт са, но не на ученето на индивидите, а на промени, осъществили се в течение на съществуването на даден животински вид. Или казано по-просто: смея да твърдя, че целите и потребностите са *продукт на еволюцията*. Тъй като една цялостна картина на интелигентното поведение би трябвало да ги вземе предвид, то ние се нуждаем от известно разбиране по въпроса как може да са се зародили в хода на еволюцията.

За да изследваме тези проблеми, заедно със Стефано Нолфи и Доменико Паризи проведохме един прост експеримент с колония от примитивни невронни мрежи, създадени така, че да приличат на примитивни амебоподобни същества. На Фигура 6 виждате една от тях. Тя имаше органи за възприятие, които ѝ позволяваха да "помирише" храната в средата си, както и два двигателни органа, които може да си представим като един вид "крака" и с които тя може да се придвижва. Мрежата обаче първоначално бе свързана по случаен начин, тоест връзките в този организъм не бяха установени с цел да постигнат някакво полезно поведение. Когато дадено същество бъде поставено в свят с храна (както е показано на Фигура 7), то може просто да си седи бездейно по време на цялото си съществуване. Други същества пък се въртят в кръг, а трети вървят по права линия, докато не стигнат до стена. Ако обаче по случайност съществото премине през място с храна, то я изяжда.



Фигура 6. Невронна мрежа, която е еволюирала да търси храна. Мрежата има поведение на много просто амебоподобно същество.



В края на живота си повечето същества измряха, без да са оставили поколение. Някои по-големи късметлии обаче случайно бяха намерили достатъчно храна, за да се осигурят енергия за самовъзпроизводство. Техните деца обикновено бяха съвсем същите като родителите. Понякога обаче се получаваша малки грешки при процеса на възпроизводство (можем да си ги представим като един вид мутации), така че детето по някакъв начин се оказваше малко по-различно от родителя. Тогава потомството бе поставено в света и оставено да живее известен период от време. Когато неговите представители измираха, те също оставяха деца, но само ако по случайност се бяха натъкнали на храна.

Позволихме на този процес да продължи в рамките на 50 поколения. Когато стигнахме до последното, погледнахме какво е поведението на тези същества. За наша изненада, то бе доста по-различно от това на предците им. Вместо да седят пасивно или да се разхождат по случаен начин из средата, индивидите от последното поколение систематично изследваха своя свят, търсеха всяко парче храна и го изяждаха, докато накрая не оставаше нищо. Ако разгледаме тяхната дейност, оставаме с впечатление, че те са целенасочени и имат ясни желания - искат храна и знаят как да я получат.

Тези желания, естествено, са просто резултат от натрупването на малки случайни изменения, привнесени от мутациите. Тъй като се

възпроизвеждаха само съществата, намерили храна, деца можеха да имат само тези, които в резултат на случайни промени имаха поведение, увеличаващо способността им за намиране на храна. Техните деца също бяха подложени на еволюционен натиск и така процесът премина през период на натрупване от 50 поколения. Накрая поведението бе действително целенасочено, макар вероятно не по начин, който ние бихме си представили. Тук не става въпрос за съзнателно и интроспективно познание, а само за поведение, основаващо се на рефлексни - пораждащо поведение, което позволява на съществата да оцелеят и да имат поколение. Но може би това не е толкова различно от нашите собствени желания и потребности, както и да ги мислим?

* * *

С тези примери преминах през много голяма част от областта на изследванията си и се извинявам, че се наложи така да бързам и да изпускам много подробности (но ако бях включил всички детайли, несъмнено щях да ви се извинявам за това, че ви задържам с часове). Надявам се да съм ви дал представа за различните въпроси, които са ми били интересни през изминалите години, и се надявам да съм ви убедил, че този тип модели, които описах тук, може да доведе до някои нови хипотези за това как да се търси отговор на тези въпроси при хората. Не мисля, че моделите работят точно като хората и че само защото решават проблемите по даден начин, то това непременно е и начинът, валиден за хората. Ценността им е, че ни дават по-различен подход към проблема и подсказват хипотези, за които може би не сме се досетили сами и които впоследствие можем да изследваме чрез пряко експериментиране с хора. Моделите не правят нищо повече, но надявам се ще се съгласите с мен, че това само по себе си е много ценно.