

Od neurónových sietí ku Cogitoidu

Veľká časť výskumu v umelej inteligencii sa venuje modelovaniu mozgu a jeho činnosti. K najznámejším výsledkom v tejto oblasti nesporne patria *Neurónové siete*, ktoré dokážu veľmi dobre modelovať niektoré z jeho funkcií. Aj napriek veľkej snahe sa nevyhli niektorým nedostatkom a preto vznikli a stále vznikajú ďalšie modely, ktoré sa snažia simulovať kognitívne funkcie mozgu lepšie a presnejšie. Výrazným stupienkom v tomto smere je *Neuroid*, ktorého autorom je L. Valiant [11].

Neuroid

Na tomto modeli sa autor pokúša vysvetliť úžasnú schopnosť mozgu ukladať a spracúvať informácie za predpokladu, že neuróny a neurálne spojenia majú internú štruktúru, na ktorej záleží. Každý *Neuroid* je čiastočka s lineárnym prahom, rozšírená o stav a časovú závislosť tak, aby lepšie popisovala synchronizované rytmické správanie mozgovej kôry. Oproti štandardnej umelej neurónovej sieti je doplnený o dynamiku vnútorných parametrov siete. Základnou myšlienkou je teda spojenie perceptrónu s modelom konečno-stavového automatu, a ďalšími vlastnosťami umožňujúcimi *neuroidu* meniť svoj stav, váhy a prah. Výpočet siete pozostávajúcej z *neuroidov* je definovaný rovnako, ako výpočet štandardnej umelej neurónovej siete, len s tým rozdielom, že po každom kroku dochádza paralelne k zmenám váh a stavov všetkých elementov. Schopnosť *neuroidov* meniť svoje vnútorné charakteristiky im umožňuje „učenie“. Výpočtová sila *neuroidnej* siete” je rovnaká ako výpočtová sila umelých neurónových sietí, sú však omnoho lepším modelom na skúmanie učenia. Samy o sebe sa nedajú považovať za model mozgových alebo kognitívnych funkcií – za takýto model ich možno považovať až po nadefinovaní konkrétneho správania a topológie *neuroidálnej* siete.

Model povrchu pamäte

Významnou inšpiráciou pre model Cogitoidu je aj takzvaný *Memory Surface Model* (model povrchu pamäte), ktorého autorom je L. Goldschlager [2]. Model povrchu pamäte, narozdiel od predošlého, získal na flexibilitu abstrahovaním od vnútornej štruktúry, pričom sa sústreďuje len na "body" v mozgu (nazývajú sa stĺpce) a na tok informácií medzi nimi. Svojou funkcionalitou sú najbližšie impulzným perceptrónom. Komunikácia medzi stĺpcami prebieha vo forme sledov pulzov s premenlivou frekvenciou, pozdĺž váhovaných hrán grafu. Niektoré stĺpce majú spojenie na predspracované vnemy – reprezentujú vstup povrchu pamäte, ďalšie reprezentujú výstup. Reprezentácia znalostí je v modeli povrchu pamäte realizovaná pomocou konceptov zložených z obrazcov (*patterns*). Každý obrazec je tvorený množinou stĺpcov, spolu s ich (relatívnymi, normalizovanými) aktivitami, vyjadrujúcimi dôležitosť daného stĺpca v obrazci. Hovoríme, že obrazec je aktívny so silou s , ak sa jeho momentálna aktivita líši od predchádzajúcej s -násobne. Pokiaľ medzi aktivitami dvoch stĺpcov existuje korelácia, odrazí sa toto na váhach spojení medzi týmito dvoma stĺpcami. To isté platí aj pre obrazce – pokiaľ sú simultánne aktívne, stanú sa asociovanými. Tento mechanizmus umožňuje vznik zložitejších konceptov z jednoduchších. Napríklad, pokiaľ sú dva objekty často pozorované (fyzicky) spolu, stanú sa asociovanými. Toto platí rovnako aj pre udalosti, ktoré po

sebe často nasledujú (asociácia príčiny a dôsledku je špeciálnym prípadom tohto mechanizmu).

Nás však bude zaujímať model, ktorý navrhol J. Wiedermann - *Cogitoid*. Autor sa snaží poukázať na nedostatky týchto modelov a navrhol vlastný model v ktorom sa snaží tieto nedostatky odstrániť. Žiadny z predchádzajúcich modelov nedokáže uspokojivo zvládať zložitejšie kognitívne úlohy. V Goldschlagerovom modeli absentuje prítomnosť detailov, ktoré sú potrebné pri formálnom usudzovaní. Naopak neuroid poskytuje dostatočné detaily, ale modelovanie zložitejších kognitívnych úloh na úrovni samotného neuroidu je takisto nevyhovujúce. J. Wiedermann ponúka model, ktorý je dostatočne formálny, aby zvládol matematické usudzovanie, pričom vynecháva detaily na neuronálnej úrovni a teda poskytuje flexibilitu.

Cogitoid

Cogitoid je teda formálny model učiaceho sa systému, samotným autorom (J. Wiedermann) prezentovaný ako model ľudskej mysle [4,5,6]. Je to výpočtový model určený pre umelú inteligenciu, ktorý vychádza tak zo symbolického, ako aj z konekcionistického prístupu. Filozoficky možno tento model zaradiť ako behavioristický. Cogitoid reprezentuje všetky vstupy a výstupy ako koncepty ktoré sú mu buď predkladané, alebo on prezentuje na svojom výstupe. Koncept pritom tvorí množina elementárnych vlastností nejakého javu, prípadne objektu. Cogitoid sa snaží spájať príbuzné koncepty za pomoci asociácií. Samotný výpočet Cogitoidu teda prebieha tak, že sa postupne aktivujú koncepty, ktoré môžeme reprezentovať ako jeho myšlienky, alebo spomienky, prislúchajúce vstupom, prípadne sú spojené s týmito už aktivovanými vstupmi za pomoci asociácií.

Od klasickej umelej inteligencie preberá myšlienku konceptov, ktoré zodpovedajú logickým formulám (konjunkcia predikátov pre jednotlivé elementárne vlastnosti). Odvodzovanie nových konceptov však neprebieha logickým odvodzovaním, ale je založené na asociáciách, čo sa dá prirovnať k procesom ktoré prebiehajú v neurónových sieťach.

Ako to celé funguje?

Koncept je základným kameňom Cogitoidu. Takýto koncept však môže byť určený niekoľkými elementárnymi prvkami, ktoré sa nazývajú črty alebo javy. Koncept je teda tvorený množinou niektorých črt. Keď si predstavíme, že rozoznávame n – črt a očísľujeme ich 1.. n , každý binárny vektor takejto dĺžky zodpovedá práve jednému konceptu. Spolu je takýchto vektorov 2^n a jednotka na i -tom mieste daného vektoru hovorí o prítomnosti resp. neprítomnosti črty číslo i v koncepte zodpovedajúcom tomuto vektoru. (Keď konceptom priradíme na základe ich binárnej reprezentácie číslo v desiatkovej sústave, môžeme s nimi pracovať len ako s číslami - od 1 do 2^n). Množiny, ktoré tieto koncepty vytvárajú, nazývame univerzum. Táto množina je definovaná pre každý Cogitoid a vymedzuje koncepty, s ktorými môže pracovať.

Jednotlivé koncepty sú navzájom prepojené, pričom jeden koncept môže byť abstrakciou (zovšeobecnením), alebo konkretizáciou iného konceptu. Ak si zoberieme nejaký všeobecný koncept a pridáme k nemu nejakú ďalšiu vlastnosť (črtu), dostaneme koncept, ktorý bude konkrétnejší ako ten z ktorého sme vychádzali. Opačne, ak z konceptu odstránime nejakú vlastnosť, dostaneme niečo všeobecnejšie, čo pokrýva väčšiu množinu objektov. Ďalšou reláciou medzi konceptmi je relácia podobnosti. Dva koncepty sú si podobné, ak majú nejaké spoločné vlastnosti. Je dobré si uvedomiť, že jeden koncept nemusí byť konkretizáciou ani abstrakciou druhého konceptu, ale pritom môže platiť, že sú si podobné. Vzťahy abstrakcia, konkretizácia a podobnosť sú určené významom konceptov a teda sa počas celého výpočtu Cogitoidu nemenia.

Niektoré koncepty reprezentujú pocity, alebo emócie pričom môžu byť buď pozitívne (spokojnosť), alebo negatívne (nespokojnosť). Takéto koncepty sa nazývajú afekty. Každý koncept má priradenú hodnotu sily, ktorá hovorí o tom, ako často sa Cogitoid s týmto konceptom počas výpočtu stretol. Koncepty ďalej môžeme rozdeľovať podľa toho, či sú v daný okamih aktívne, alebo neaktívne. Aktívne sú tie, s ktorými Cogitoid v daný časový okamih práve pracuje a teda sa nachádzajú v „operačnej pamäti“.

Asociácia je vždy relácia medzi dvomi konceptmi, pričom je orientovaná. Asociácie rozdeľujeme na excitačné a inhibičné. Excitačné sú také, ktoré posilňujú aktivitu druhého konceptu, pokiaľ je prvý koncept aktívny. Naopak inhibičné túto aktivitu potláčajú. Každá asociácia má svoju silu, ktorou ovplyvňuje koncept, na ktorý „ukazuje“. Sila asociácií sa dynamicky počas výpočtu Cogitoidu mení v závislosti od toho, čo dostáva na vstup a jeho vnútorných výpočtov. Asociácie sa posilňujú dvomi spôsobmi:

1. na princípe kauzality (časovej postupnosti), ktorá funguje tak, že keď je v jednom kroku výpočtu aktivovaný koncept A a v nasledujúcom kroku koncept B, tak excitačná asociácia z A do B sa posilní. Ak je toto posilnenie dostatočne veľké, tak pri nasledujúcej aktivácii konceptu A si Cogitoid „spomenie“ na to, že tesne po A bol aktivovaný koncept B a teda tento koncept B tiež aktivuje.
2. na princípe podobnosti konceptov, ktorá nastáva automaticky hneď keď je nejaký koncept aktivovaný, Cogitoid „nájde“ podobné koncepty a posilní asociácie medzi nimi symetricky (obojsmerne). Toto má za následok, že ak bude aktivovaný jeden z podobných konceptov, tak Cogitoid si „spomenie“ na všetky podobné koncepty.

Asociácie teda tvoria akúsi dlhodobú pamäť Cogitoidu.

Pozitívne afekty majú za následok posilnenie excitačných asociácií. Negatívne afekty vedú k posilneniu inhibičných asociácií.

Výpočet Cogitoidu prebieha v neustále sa opakujúcich krokoch. Na začiatku každého kroku Cogitoid aktivuje koncepty, ktoré má na vstupe. Následne použije asociácie vedúce z konceptov aktívnych v predošlom kroku výpočtu a aktivuje koncepty, ktoré sú týmito asociáciami najviac excitované. Ďalej nastane posilnenie asociácií a aktívnych konceptov podľa zadaných kritérií. Následne prebieha

oslabenie neaktívnych konceptov a ich asociácií, čo má za následok postupné „zabúdanie“. Samozrejme toto „zabúdanie“ musí prebiehať pomalšie ako posilňovanie používaných konceptov a asociácií, inak by klesli sily všetkých neaktívnych asociácií na nulu a Cogitoid by už neaktivoval nič, čo nie je aktivované priamo zo vstupu. V konečnej fáze sú na výstup odoslané práve aktívne koncepty a deaktivované všetky aktívne koncepty, ktoré neboli aktivované v tomto výpočtovom kroku. Týmto je zavŕšený cyklus a môže začať nový výpočet. Podrobnejšie rozpísané to vyzerá nasledovne:

- 1. fáza - vstup/výstup

Vstup: Zo vstupu načítame v čase t množinu konceptov I_t . Každý koncept z tejto množiny aktivujeme (aktivácia vonkajším podnetom). Ďalej aktivujeme konkretizáciu (suprémum) i_t tejto množiny (aktivácia súčasným výskytom). Jej aktivácia spôsobí aktiváciu každej abstrakcie tohto konceptu, čím vytvorí ideál¹ L_{it} . Výstup: Výstupom je množina O_t , ktorú sme vytvorili v predošlom kroku.

- 2. fáza - aktivácia vnútorným podnetom

Spomedzi všetkých pasívnych konceptov z množiny P_t vyberieme ten koncept, ktorý je najviac *excitovaný*. Excitácia konceptu sa vypočíta ako súčet síl všetkých jeho excitačných asociácií s aktívnymi konceptmi, zmenšená o súčet inhibičných asociácií s aktívnymi konceptmi. Vyberieme teda najviac excitovaný koncept a (prípadne, ak je takých viac, tak ich konkretizáciu) a súčasným výskytom aktivujeme všetky jeho abstrakcie - L_a .

- 3. fáza - výpočet kvality

Nech A_t je množina všetkých aktívnych konceptov (L_{it}, O_t, L_a). Kvalitu konceptov z tejto množiny vypočítame podľa nasledovného pravidla: Z aktivovaných operačných konceptov propagujeme ich kvalitu na väčšie prvky (podľa usporiadania v úplnom zväze). Ak nejaký koncept získa kvalitu od kladného aj záporného operačného konceptu, potom je jeho kvalita rovná ± 1 . Koncepty, ktoré týmto spôsobom nezískali žiadnu kvalitu, ju zdedia od najmenších konceptov, ktoré ich obklopujú. Ak v danom čase nie je aktívny žiadny operačný koncept, všetky kvality budú rovné $+1$.

- 4. fáza - ukladanie do dlhodobej pamäte

Posilňovanie konceptov: Silu všetkých aktívnych konceptov z A_t zvýšime o pevne danú kladnú konštantu c_1 .

Posilňovanie asociácií: Silu všetkých asociácií medzi aktívnymi konceptmi zvýšime o pevne danú kladnú konštantu c_2 nasledovne: Prvý koncept z dvojice vyberieme z množiny O_t , ďalší vyberieme z množiny novo-aktivovaných konceptov $L_{it} + L_a - O_t$. Ak je kvalita prvého konceptu záporná, posilníme inhibičnú asociáciu, ak je kladná, posilníme excitačnú. Ak je kvalita rovná ± 1 , posilníme obidve asociácie. Navyše, ak prvý koncept bol v predošlom kroku aktivovaný vonkajším podnetom, posilňujeme asociácie o hodnotu c_3 .

- 5. fáza - postupné zabúdanie

Sila všetkých konceptov, ktoré nie sú aktívne aj asociácií medzi nimi, sa zníži o pevne danú konštantu c_0 .

- 6. fáza - deaktivácia

Deaktivujeme všetky koncepty z $O_t - L_{it} - L_a$. Množina aktívnych konceptov zostane

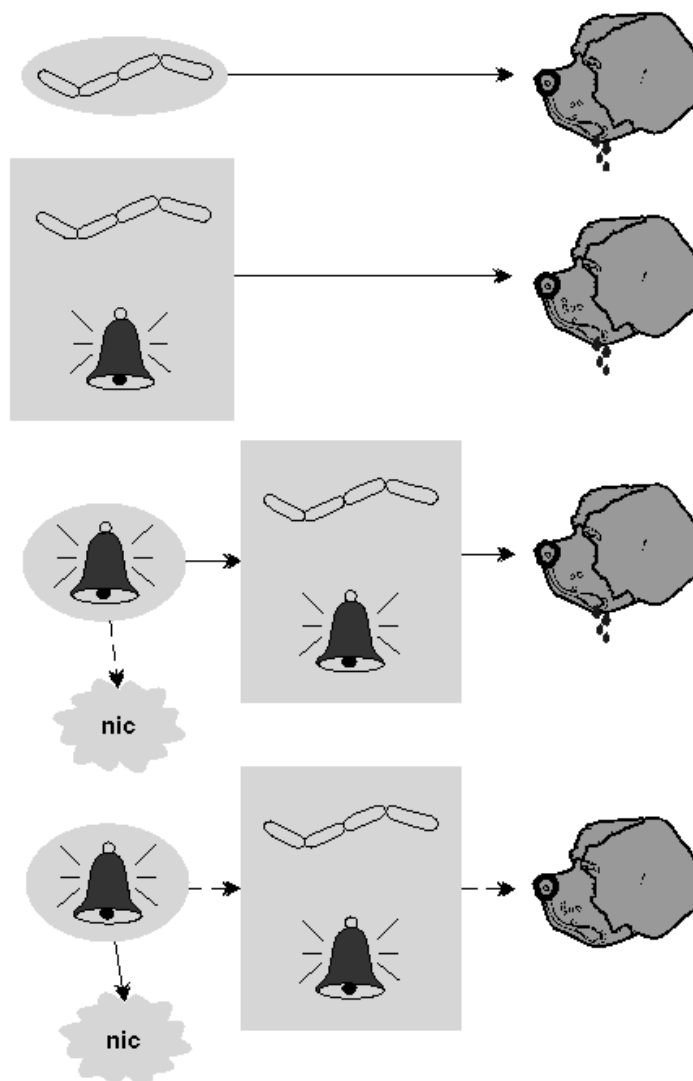
¹ Neprázdnu podmnožinu \mathbf{I} okruhu \mathbf{A} nazývame ideálom okruhu \mathbf{A} , ak

1. pre $\forall x, y \in \mathbf{I}$ je $x - y \in \mathbf{I}$,
2. pre $\forall x \in \mathbf{I}, a \in \mathbf{A}$ je $ax \in \mathbf{I}, xa \in \mathbf{I}$

iba $L_{it} + L_a$ a bude tvoriť množinu aktívnych konceptov v ďalšom kroku výpočtu (O_{t+1}).

K názornejšej predstave uvidíme konkrétny príklad simulácie Pavlovovského reflexu, ktorý bol rozpracovaný v diplomovej práci Michaely Ačovej, na ktorej výsledky sa snažím nadviazať. Model Cogitoidu odskúšala na príklade klasického Pavlovovského pokusu so psom (Obr. 1). Podnet (koncept S) je v tomto prípade jedlo, r je reakcia, čiže produkcia slín so zámerom konzumácie jedla. Koncept a je napríklad zasvietenie žiarovky, alebo zazvonenie zvončeka, ktoré sprevádza podanie jedla. Po určitom čase pes reaguje tvorbou slín aj bez podania jedla, iba pri zazvonení (zasvietení).

Predpokladáme, že koncept S je tak silný, že jeho samostatné prezentovanie ako vstup pre Cogitoid, vyvolá na výstupe koncept r . Ak teda budeme na vstup prezentovať dvojicu a, S tak na výstupe sa určite objaví r , pričom ale Cogitoid výrazne posilní asociáciu a, S . Tým sa stane, že ak na vstupe dáme následne len koncept a , tak ten aktivuje asociáciu a, S , čím sa na výstup dostane opäť koncept r . Keby sme pokračovali s opätovným prezentovaním konceptu a , bez konceptu S , dosiahneme ochabenie asociácie a, S a Cogitoid prestane aktivovať na výstupe koncept r . Toto zodpovedá štandardnému zvieraciemu správaniu, čím sa potvrdilo, že Cogitoid je schopný simulovať jednoduché kognitívne procesy ako : formovanie abstrakcii, asociatívny výber, podmienené učenie, alebo správanie na základe podobnosti.



Obr. 1 – Schéma pokusu na podmienený reflex