# Bazele teoretice ale calculului evolutiv

Ideea de a aplica principiile darwiniste ale evoluţiei în rezolvarea automată a problemelor (*Problem Solving* - PS) datează din anii 1940, înaintea apariţiei calculatoarelor electronice. În 1948 Turing propunea o tehnică de rezolvare a problemelor numită *căutare genetică* sau *evolutivă*. În anii 1960 Fogel, Qwens şi Walsh introduceau conceptul de *programare evolutivă* (sau evoluţionistă), în timp ce Holland dezvolta *algoritmii genetici*. În aceeaşi perioadă, Rechenberg şi Schwefel introduceau *strategiile evolutive* (evoluţioniste) ca modalităţi alternative de rezolvare automată a problemelor. În anii 1990, Koza dezvoltă o nouă tehnică de căutare în spaţiul soluţiilor, *programarea genetică*. În terminologia actuală, întregul spectru de metode de rezolvare automată de inspiraţie darwinistă este desemnat prin termenul de *calcul evolutiv* (evoluţionist) şi include subdomeniile: programare evolutivă, strategii evolutive, algoritmi genetici şi programare genetică.

Calculul evolutiv este un domeniu al informaticii inspirat din procesul evoluţiei naturale; ideea care stă la baza calculului evolutiv este conexiunea evoluţie naturală – tehnica de rezolvare a problemelor de tip experiment-eroare (sau generare-testare). Cu alte cuvinte, într-un mediu dat, indivizii constituiţi într-o populaţie intră în competiţie pentru a supravieţui şi a se reproduce. Abilitatea indivizilor de a-şi atinge aceste scopuri în mediul în care trăiesc este strict corelată cu şansele lor de supravieţuire şi multiplicare şi determină evoluţia în timp a populaţiei. În contextul modalităţii de rezolvare a problemelor de tip generare-testare stochastice, populaţia este modelată ca o colecţie de elemente candidat la soluţie. Calitatea candidaţilor la soluţie, definită în termenii gradului în care fiecare element rezolvă problema, determină şansa lor de a fi menţinuţi şi utilizaţi pentru construirea unor noi candidaţi.

**Suportul de natură biologică al calculului evolutiv**

Teoria evoluţionistă a lui Darwin oferă o explicaţie a diversităţii biologice şi a mecanismului care stă la baza acesteia. În centrul interpretării macroscopice a evoluţiei este plasată selecţia naturală. Considerând un mediu care poate susţine un număr limitat de indivizi şi instinctul primar al fiecărui individ de a se reproduce, procesul de selecţie este esenţial şi inevitabil în controlul dimensiunii populaţiei. Selecţia naturală favorizează indivizii cei mai competitivi în însuşirea resurselor, adică acei indivizi care sunt cel mai bine adaptaţi condiţiilor de mediu. Fenomenul este cunoscut drept supravieţuirea celor mai bine adaptaţi/potriviţi (*survival of the fittest*).

Teoria evoluţiei are ca principii fundamentale selecţia bazată pe competiţie şi variaţiile de *fenotip* în rândul membrilor populaţiei. Fenotipul este ansamblul de însuşiri şi caractere care se manifestă în mod vizibil la un individ şi care este determinat pe bază ereditară şi de condiţiile de mediu (DEX). Caracteristicile fenotipului unui individ determină gradul lui de adaptabilitate la condiţiile de mediu (*fitness*). Fiecare individ reprezintă o combinaţie unică de caracteristici ale fenotipului şi este evaluat de condiţiile de mediu. Dacă evaluarea este favorabilă, atunci fenotipul individului este propagat spre urmaşi (progenituri), altfel caracteristicile fenotipului dispar şi individul moare fără a se putea reproduce. Viziunea lui Darwin despre evoluţie este aceea că, în procesul trecerii de la o generaţie la alta prin reproducere, apar mutaţii (variaţii) mici, aleatoare, în caracteristicile fenotipului. Ca rezultat al acestor variaţii, apar şi sunt evaluate noi combinaţii de caracteristici ale fenotipului. Cele mai bune dintre ele supravieţuiesc şi se reproduc şi în acest mod evoluţia conduce la progres.

Modelul primar poate fi deci rezumat după cum urmează. O populaţie este formată dintr-un număr de indivizi, priviţi ca unităţi de selecţie; succesul fiecărui individ în încercarea de a se reproduce depinde de cât de bine este adaptat condiţiilor de mediu comparativ cu restul indivizilor. Pe parcursul reproducerii celor mai buni (bine adaptaţi la mediu) indivizi, apar mutaţii ocazionale, care generează noi indivizi, ce vor fi ulterior evaluaţi. În consecinţă, pe măsură ce timpul trece, se produc schimbări în structura populaţiei, cu alte cuvinte populaţia reprezintă o unitate a evoluţiei. Întregul proces poate fi asimilat din punct de vedere intuitiv cu modelul unui peisaj adaptiv (dinamic) (respectiv o suprafaţă adaptivă) în spaţiul 3D (0-x-y-z). Fiecare punct p(x,y,z) al suprafeţei este asimilat unui individ, unde în planul x-y este figurată combinaţia de caracteristici ale fenotipului individului, iar altitudinea lui p (valoarea lui pe axa 0z) corespunde nivelului de adaptabilitate (fitness) a individului reprezentat de p. În acest context, evoluţia este procesul de „avans” al populaţiei spre zone aflate la o altitudine mai mare, acest avans fiind realizat pe baza mutaţiilor şi selecţiei naturale. Este obţinută astfel legătura cu conceptul de *probleme multimodale*, adică probleme în care există o mulţime de puncte de optim local (superioare tuturor soluţiilor din vecinătatea lor), cel mai bun element al mulţimii fiind optimul global. O problemă în care există un singur optim local este numită *unimodală*.

Legătura dintre procesul evoluţiei şi un proces de optimizare este pe cât de directă, pe atât de înşelătoare, pentru că evoluţia nu presupune întotdeauna creşterea globală a calităţii populaţiei (în termenii funcţiei de fitness). Deoarece, la fiecare epocă, populaţia este finită şi operatorii de selecţie şi mutaţie includ componente aleatoare, poate fi constatată o inactivitate genetică, manifestată fie prin dispariţia din populaţie a unor indivizi foarte adaptaţi condiţiilor de mediu, fie prin variaţia foarte mică sau chiar lipsa de variaţie a unor caracteristici ale fenotipului indivizilor din populaţie. Unul din efectele posibile este acela al concentrării indivizilor într-o zonă de adaptabilitate scăzută. Efectul rezultat prin combinarea procesului de selecţie cu inactivitatea genetică poate conduce în egală măsură la creşterea nivelului de adaptabilitate globală a populaţiei, respectiv la descreşterea acestuia. În plus, nu există nici o garanţie că, dacă evoluţia duce la creşterea calităţii populaţiei, nivelul de optim local nu a fost deja atins înaintea unui fenomen de inactivitate genetică. În scopul evitării „ciclării” evoluţiei populaţiei într-o regiune de optim local, au fost dezvoltate o serie de teorii, una dintre cele mai cunoscute fiind teoria lui Wright conform căreia poate fi determinat optimul global în cazul unei suprafeţe fixe (teorie referită drept *shifting balance*).

**Tipuri de probleme ce pot fi rezolvate pe baza calculului evolutiv**

În literatura de specialitate sunt evidenţiate două clase de metode PS de inspiraţie biologică: calculul neuronal (*neurocomputing*), prin intermediul căruia problemele sunt rezolvate imitând modul de funcţionare (raţionament) a (al) creierului uman şi procesele de tip evolutiv, problemele fiind rezolvate prin imitarea creării creierului uman.

În general, un sistem funcţional de rezolvare automată a problemelor conţine trei componente: datele de intrare, datele de ieşire şi modulul intern care conectează primele două componente. Modalitatea de funcţionare a modelului permite explicarea modului de funcţionare a sistemului, în sensul că, răspunsul sistemului poate fi calculat pentru orice date de intrare specificate. În funcţie de componentele din sistem cunoscute, pot fi diferenţiate următoarele tipuri de probleme:

* *Probleme de optimizare*: sunt cunoscute modelul şi datele de ieşire dorite (respectiv o descriere a acestora), iar problema este de a determina datele de intrare care corespund rezultatelor dorite. Un exemplu de astfel de problemă este cea a comis voiajorului (în care trebuie determinat cea mai scurtă sau ieftină rută care să lege un număr dat de oraşe): modelul este cunoscut şi corespunde formulei de calcul al lungimii unei rute date, în care lungimea (sau costul) calculată (calculat) este data de ieşire. Proprietatea pe care rezultatul trebuie să o îndeplinească este un criteriu de optim (lungime minimă), iar problema este de a determina acea dată de intrare, corespunzătoare unei rute, care să conducă la rezultatul dorit. O problemă de optimizare rezolvată cu succes prin calcul evolutiv este generarea orarului în cadrul unei universităţi. În cursul unei zile sunt programate în general mii de activităţi, restricţiile pe care trebuie să la îndeplinească o programare corectă a acestora fiind multiple iar soluţiile fezabile ale problemei în număr foarte mic relativ la mulţimea tuturor programărilor posibile.
* *Probleme de modelare sau de identificare a sistemului*: sunt cunoscute datele de intrare şi rezultatele corespunzătoare lor, iar modelul este necunoscut. Modelul trebuie determinat astfel încât, pentru fiecare intrare dată, să calculeze rezultatul corect. Un exemplu de astfel de problemă: clasificarea supervizată în cazul modelului cu două clase. Problema este de a determina un clasificator care să separe corect elementele celor două clase. Datele de intrare corespund elementelor celor două clase, pentru fiecare dată de intrare rezultatul fiind eticheta clasei de provenienţă. Identificarea modelului revine la determinarea unei funcţii de decizie, care, de exemplu, să calculeze valori pozitive pentru exemplele care provin din prima clasă, respectiv valori negative pentru celelalte. Cu alte cuvinte, scopul este de a determina o formulă (în acest caz expresia analitică a unei funcţii de decizie) care să „lege” datele de intrare (cunoscute) de rezultate (cunoscute). Problemele de acest tip apar în medii în care sunt disponibile foarte multe date (observaţii, înregistrări etc.), de exemplu în situaţii în care există un set de dimensiuni considerabile de observaţii/înregistrări asupra/relative la un fenomen/eveniment. Identificarea modelului care să explice conexiunile dintre datele de intrare şi rezultate trebuie realizată şi astfel încât acesta să asigure o capacitate de generalizare rezonabilă (pentru noi date de intrare, sistemul trebuie să furnizeze în general răspunsuri corecte). Astfel de probleme sunt cele din domeniile instruirii automate (*machine learning*) şi *data mining*.
* *Probleme de simulare*: se cunosc modelul şi o serie de date de intrare şi cerinţa este de a determina datele de ieşire corecte, corespunzătoare intrărilor date. Un exemplu de probleme de simulare care pot fi rezolvate utilizând calculul evolutiv sunt cele în care este căutat răspunsul la întrebări de tipul „ce se întâmplă dacă” (*what-if questions*), în condiţiile în care problema subiectului investigat evoluează (în termenii operaţiilor de variaţie şi selecţie). Economia evolutivă este un domeniu de cercetare relativ nou, care, în principiu, are la bază ideea că jocul şi jucătorii din arena socio-economică sunt asemenea evoluţiei vieţii (a „jocului” şi „jucătorilor” ei).