

Jocurile ca Problemă de Căutare

Jocurile reprezintă o arie de aplicație pentru algoritmi euristici.

Jocurile de două persoane sunt în general complicate datorită existenței unui oponent ostil și imprevizibil. De aceea ele sunt interesante din punctul de vedere al dezvoltării euristiciilor, dar aduc multe dificultăți în dezvoltarea și aplicarea algoritmilor de căutare.

Oponentul introduce *incertitudinea*, întrucât nu se știe niciodată ce va face acesta la pasul următor. În esență, toate programele referitoare la jocuri trebuie să trateze așa numita *problemă de contingență*. (Aici contingență are sensul de întâmplare).

Incertitudinea care intervine în cazul jocurilor nu este de aceeași natură cu cea introdusă, de pildă, prin aruncarea unui zar sau cu cea determinată de starea vremii. Oponentul va încerca, pe cât posibil, să facă mutarea cea mai puțin benignă, în timp ce zarul sau vremea sunt presupuse a nu lua în considerație scopurile agentului.

Complexitatea jocurilor introduce un tip de *incertitudine complet nou*. Astfel, incertitudinea se naște nu datorită faptului că există informație care lipsește, ci datorită faptului că jucătorul nu are timp să calculeze consecințele exacte ale oricărei mutări. Din acest punct de vedere, jocurile se aseamănă infinit mai mult cu lumea reală decât problemele de căutare standard.

Întrucât, în cadrul unui joc, există, de regulă, limite de timp, jocurile *penalizează ineficiența* extrem de sever. Astfel, dacă o implementare a căutării de tip A^* , care este cu 10% mai puțin eficientă, este considerată satisfăcătoare, un program pentru jocul de șah care este cu 10% mai puțin eficient în folosirea timpului disponibil va duce la pierderea partidei. Din această cauză, studiul nostru se va concentra asupra *tehnicilor de alegere a unei bune mutări atunci când timpul este limitat*. **Tehnica de “retezare”** ne va permite să ignorăm porțiuni ale arborelui de căutare care nu pot avea nici un rol în stabilirea alegerii finale, iar **funcțiile de evaluare euristice** ne vor permite să aproximăm utilitatea reală a unei stări fără a executa o căutare completă.

Ne vom referi la tehnici de joc corespunzătoare unor **jocuri de două persoane cu informație completă**, cum ar fi șahul.

În cazul jocurilor interesante, arborii rezultați sunt mult prea complecși pentru a se putea realiza o căutare exhaustivă, astfel încât sunt necesare abordări de o natură diferită. Una dintre metodele clasice se bazează pe „**principiul minimax**”, implementat în mod eficient sub forma **Algoritmului Alpha-Beta** (bazat pe așa-numita *tehnică de alpha-beta retezare*).

O definiție Formală a Jocurilor

Tipul de jocuri la care ne vom referi în continuare este acela al jocurilor de două persoane cu informație perfectă sau completă. În astfel de jocuri:

- există doi jucători care efectuează mutări în mod alternativ;
- ambii jucători dispun de informația completă asupra situației curente a jocului;
- jocul se încheie atunci când este atinsă o poziție calificată ca fiind “terminală” de către regulile jocului - spre exemplu, “mat” în jocul de șah.

Un asemenea joc poate fi reprezentat printr-un **arbore de joc** în care nodurile corespund situațiilor (stărilor), iar arcele corespund mutărilor. Situația inițială a jocului este reprezentată de nodul rădăcină, iar frunzele arborelui corespund pozițiilor terminale.

Vom lua în considerație cazul general al unui joc cu doi jucători, pe care îi vom numi **MAX** și respectiv **MIN**. MAX va face prima mutare, după care jucătorii vor efectua mutări pe rând, până când jocul ia sfârșit. La finalul jocului vor fi acordate puncte jucătorului câștigător (sau vor fi acordate anumite penalizări celui care a pierdut).

Un joc poate fi definit, în mod formal, ca fiind un anumit tip de problemă de căutare având următoarele componente:

- **starea inițială**, care include poziția de pe tabla de joc și o indicație referitoare la cine face prima mutare;
- **o mulțime de operatori**, care definesc mișcările permise (“legale”) unui jucător;
- **un test terminal**, care determină momentul în care jocul ia sfârșit;
- **o funcție de utilitate** (numită și **funcție de plată**), care acordă o valoare numerică rezultatului unui joc; în cazul jocului de șah, spre exemplu, rezultatul poate fi *câștig*, *pierdere* sau *remiză*, situații care pot fi reprezentate prin valorile 1, -1 sau 0.

Dacă un joc ar reprezenta o problemă standard de căutare, atunci acțiunea jucătorului MAX ar consta din căutarea unei secvențe de mutări care conduc la o stare terminală reprezentând o stare câștigătoare (conform funcției de utilitate) și din efectuarea primei mutări aparținând acestei secvențe. Acțiunea lui MAX interacționează însă cu cea a jucătorului MIN. Prin urmare, MAX trebuie să găsească o strategie care va conduce la o stare terminală câștigătoare, indiferent de acțiunea lui MIN. Această strategie include mutarea corectă a lui MAX corespunzătoare fiecărei mutări posibile a lui MIN. În cele ce urmează, vom începe prin a arăta cum poate fi găsită strategia **optimă** (sau **rațională**), deși în realitate nu vom dispune de timpul necesar pentru a o calcula.

Algoritmul Minimax

JUCATORI: MIN și MAX

MAX reprezintă jucătorul care încearcă să câștige sau să își **maximizeze** avantajul avut. **MIN** este oponentul care încearcă să **minimizeze** scorul lui MAX. Se presupune că MIN folosește aceeași informație și încearcă întotdeauna să se mute la acea stare care este cea mai nefavorabilă lui MAX.

Algoritmul Minimax este conceput pentru a determina *strategia optimă corespunzătoare lui MAX* și, în acest fel, pentru a decide care este cea mai bună primă mutare.

Algoritmul Minimax:

1. Generează întregul arbore de joc, până la stările terminale.
2. Aplică funcția de utilitate fiecărei stări terminale pentru a obține valoarea corespunzătoare stării.
3. Deplasează-te înapoi în arbore, de la nodurile-frunze spre nodul-rădăcină, determinând, corespunzător fiecărui nivel al arborelui, valorile care reprezintă utilitatea nodurilor aflate la acel nivel. Propagarea acestor valori la niveluri anterioare se face prin intermediul nodurilor- părinte succesive, conform următoarei reguli:
 - a. dacă starea-părinte este un nod de tip MAX, atribuie-i maximul dintre valorile avute de fiii săi;
 - b. dacă starea-părinte este un nod de tip MIN, atribuie-i minimul dintre valorile avute de fiii săi.
4. Ajuns în nodul-rădăcină, alege pentru MAX acea mutare care conduce la valoarea maximă.

Observație: Decizia luată la pasul 4 al algoritmului se numește *decizia minimax*, întrucât ea *maximizează utilitatea*, în ipoteza că *oponentul joacă perfect cu scopul de a o minimiza*.

Un arbore de căutare cu valori minimax determinate conform Algoritmului Minimax este cel din Fig. 3.1:

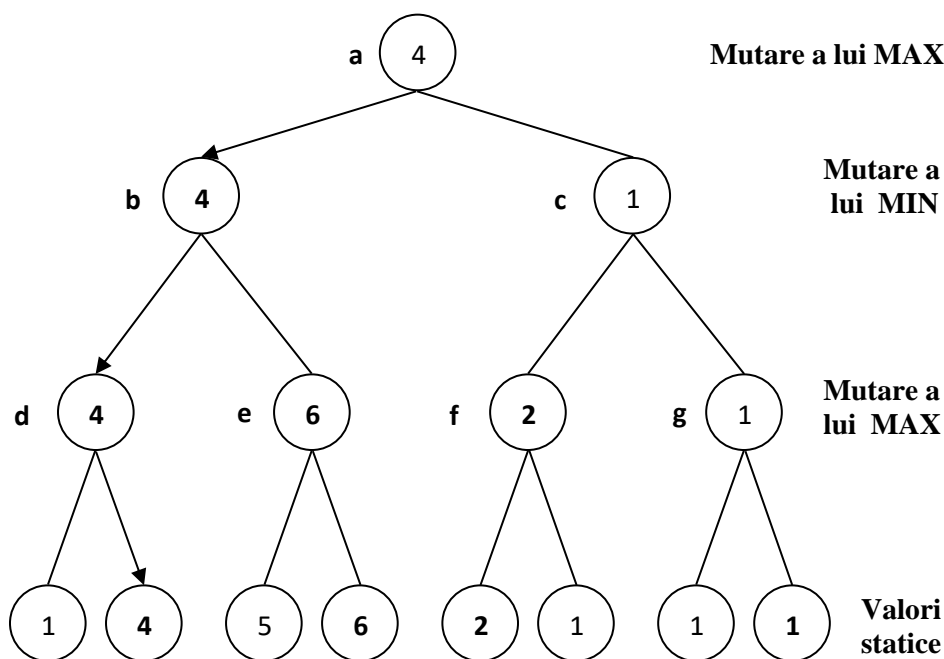


Fig. 3.1

Valorile pozițiilor de la ultimul nivel sunt determinate de către funcția de utilitate și se numesc *valori statice*. Valorile minimax ale nodurilor interne sunt calculate în mod **dinamic**, în manieră **bottom-up**, nivel cu nivel, până când este atins nodul-rădăcină. Valoarea rezultată, corespunzătoare acestuia, este 4 și, prin urmare, cea mai bună mutare a lui MAX din poziția a este *a-b*. Cel mai bun răspuns al lui MIN este *b-d*. Această secvență a jocului poartă denumirea de variație principală. Ea definește jocul optim de tip minimax pentru ambele părți. Se observă

că valoarea pozițiilor de-a lungul variației principale nu variază. Prin urmare, mutările corecte sunt cele care conservă valoarea jocului.

Dacă adâncimea maximă a arborelui este m și dacă există b “mutări legale” la fiecare punct, atunci complexitatea de timp a Algoritmului Minimax este $O(b^m)$. Algoritmul reprezintă o căutare de tip depth-first (deși aici este sugerată o implementare bazată pe recursivitate și nu una care folosește o coadă de noduri), astfel încât cerințele sale de spațiu sunt numai liniare în m și b .

În cazul jocurilor reale, cerințele de timp ale algoritmului sunt total nepractice, dar acest algoritm stă la baza atât a unor metode mai realiste, cât și a analizei matematice a jocurilor.

Întrucât, pentru majoritatea jocurilor interesante, arborele de joc nu poate fi alcătuit în mod exhaustiv, au fost concepute diverse metode care se bazează pe căutarea efectuată numai într-o anumită porțiune a arborelui de joc. Printre acestea se numără și tehnica Minimax, care, în majoritatea cazurilor, va căuta în arborele de joc numai până la o anumită adâncime, de obicei constând în numai câteva mutări. Ideea este de a evalua aceste poziții terminale ale căutării, fără a mai căuta dincolo de ele, cu scopul de a face economie de timp. Aceste estimări se propagă apoi în sus de-a lungul arborelui, conform principiului Minimax. Mutarea care conduce de la poziția inițială, nodul-rădăcină, la cel mai promițător succesori al său (conform acestor evaluări) este apoi efectuată în cadrul jocului.

Algoritmul general Minimax a fost amendat în două moduri: funcția de utilitate a fost înlocuită cu o **funcție de evaluare**, iar testul terminal a fost înlocuit de către un așa-numit **test de tăiere**.

Cea mai directă abordare a problemei deținerii controlului asupra cantității de căutare care se efectuează este aceea de a fixa o limită a adâncimii, astfel încât testul de tăiere să aibă succes pentru toate nodurile aflate la sau sub adâncimea d . Limita de adâncime va fi aleasă astfel încât cantitatea de timp folosită să nu depășească ceea ce permit regulile jocului. O abordare mai robustă a acestei probleme este aceea care aplică „iterative deepening”. În acest caz, atunci când timpul expiră, programul întoarce mutarea selectată de către cea mai adâncă căutare completă.

Funcții de Evaluare

O **funcție de evaluare** întoarce o estimatie, realizată dintr-o poziție dată, a utilității așteptate a jocului. Ea are la bază evaluarea șanselor de câștigare a jocului de către fiecare dintre părți, pe baza calculării caracteristicilor unei poziții. Performanța unui program referitor la jocuri este extrem de dependentă de calitatea funcției de evaluare utilizate.

Funcția de evaluare trebuie să îndeplinească anumite condiții evidente:

- ✓ trebuie să concorde cu funcția de utilitate în ceea ce privește stările terminale;
- ✓ calculele efectuate nu trebuie să dureze prea mult;
- ✓ trebuie să reflecte în mod corect șansele efective de câștig.

O valoare a funcției de evaluare acoperă mai multe poziții diferite, grupate laolaltă într-o categorie de poziții etichetată cu o anumită valoare. Spre exemplu, în jocul de șah, fiecare *pion* poate avea valoarea 1, un *nebul* poate avea valoarea 3 șamd.. În poziția de deschidere evaluarea este 0 și toate pozițiile până la prima captură vor avea aceeași evaluare. Dacă MAX reușește să captureze un nebul fără a pierde o piesă, atunci poziția rezultată va fi evaluată la valoarea 3. Toate pozițiile de acest fel ale lui MAX vor fi grupate într-o categorie etichetată cu “3”. Funcția de evaluare trebuie să reflecte șansa ca o poziție aleasă la întâmplare dintr-o asemenea categorie să conducă la câștig (sau la pierdere sau la remiză) pe baza experienței anterioare (VEZI tehnici de învățare).

Funcția de evaluare cel mai frecvent utilizată presupune că valoarea unei piese poate fi stabilită independent de celelalte piese existente pe tablă. Un asemenea tip de funcție de evaluare se numește **funcție liniară ponderată**, întrucât are o expresie de forma:

$$w_1f_1 + w_2f_2 + \dots + w_nf_n,$$

unde valorile w_i , $i = \overline{1, n}$ reprezintă ponderile, iar f_i , $i = \overline{1, n}$ sunt caracteristicile unei anumite poziții. În cazul jocului de șah, spre exemplu w_i , $i = \overline{1, n}$ ar putea fi valorile pieselor (1 pentru pion, 3 pentru nebul etc.), iar f_i , $i = \overline{1, n}$ ar reprezenta numărul pieselor de un anumit tip aflate pe tabla de șah.

În construirea formulei liniare trebuie mai întâi alese caracteristicile, operație urmată de ajustarea ponderilor până în momentul în care programul joacă suficient de bine. Această a doua operație poate fi automatizată punând programul să joace multe partide cu el însuși, dar alegerea unor caracteristici adecvate nu a fost încă realizată în mod automat.

O Implementare Eficientă a Principiului Minimax: Algoritmul Alpha-Beta

Tehnica pe care o vom examina, în cele ce urmează, este numită în literatura de specialitate **alpha-beta pruning** (“**alpha-beta retezare**”). Atunci când este aplicată unui arbore de tip minimax standard, ea va întoarce aceeași mutare pe care ar furniza-o și Algoritmul Minimax, dar într-un timp mai scurt, întrucât realizează o retezare a unor ramuri ale arborelui care nu pot influența decizia finală.

Principiul general al acestei tehnici constă în a considera un nod oarecare n al arborelui, astfel încât jucătorul poate alege să facă o mutare la acel nod. Dacă același jucător dispune de o alegere mai avantajoasă, m , fie la nivelul nodului părinte al lui n , fie în orice punct de decizie aflat mai sus în arbore, atunci n nu va fi niciodată atins în timpul jocului. Prin urmare, de îndată ce, în urma examinării unora dintre descendenții nodului n , ajungem să deținem suficientă informație relativ la acesta, îl putem înlătura.

Ideea tehnicii de alpha-beta retezare: a găsi o mutare “suficient de bună”, nu neapărat cea mai bună, dar suficient de bună pentru a se lua decizia corectă. Această idee poate fi formalizată prin introducerea a *două limite*, *alpha* și *beta*, reprezentând **limitări ale valorii de tip minimax corespunzătoare unui nod intern**. Semnificația acestor limite este următoarea:

- *alpha* este *valoarea minimă* pe care este deja garantat că o va obține MAX;
- *beta* este *valoarea maximă* pe care MAX poate spera să o atingă.

Din punctul de vedere al jucătorului MIN, *beta* este valoarea cea mai nefavorabilă pentru MIN pe care acesta o va atinge. **Valoarea efectivă** care va fi găsită se află între alpha și beta.

Valoarea *alpha*, asociată nodurilor de tip MAX, nu poate niciodată să descrească, iar valoarea *beta*, asociată nodurilor de tip MIN, nu poate niciodată să crească. **Alpha este scorul cel mai prost pe care îl poate obține MAX, presupunând că MIN joacă perfect.**

Dacă, spre exemplu, valoarea alpha a unui nod intern de tip MAX este 6, atunci MAX nu mai trebuie să ia în considerație nici o valoare internă mai mică sau egală cu 6 care este asociată oricărui nod de tip MIN situat sub el. În mod similar, dacă MIN are valoarea beta 6, el nu mai trebuie să ia în considerație nici un nod de tip MAX situat sub el care are valoarea 6 sau o valoare mai mare decât acest număr.

Cele două reguli pentru încheierea căutării, bazată pe valori *alpha* și *beta*, sunt:

1. Căutarea poate fi oprită dedesubtul oricărui nod de tip MIN care are o valoare *beta* mai mică sau egală cu valoarea *alpha* a oricăruia dintre strămoșii săi de tip MAX.
2. Căutarea poate fi oprită dedesubtul oricărui nod de tip MAX care are o valoare *alpha* mai mare sau egală cu valoarea *beta* a oricăruia dintre strămoșii săi de tip MIN.

Dacă, referitor la o poziție, se arată că valoarea corespunzătoare ei se află în afara intervalului *alpha-beta*, atunci această informație este suficientă pentru a ști că poziția respectivă nu se află de-a lungul *variației principale*, chiar dacă nu este cunoscută valoarea *exactă* corespunzătoare ei. **Cunoașterea valorii exacte a unei poziții este necesară numai atunci când această valoare se află între alpha și beta.**

Din punct de vedere formal, putem defini o **valoare de tip minimax** a unui nod intern, P , $V(P, \alpha, \beta)$, ca fiind “suficient de bună” dacă satisface următoarele cerințe:

$$V(P, \alpha, \beta) < \alpha, \quad \text{dacă } V(P) < \alpha$$

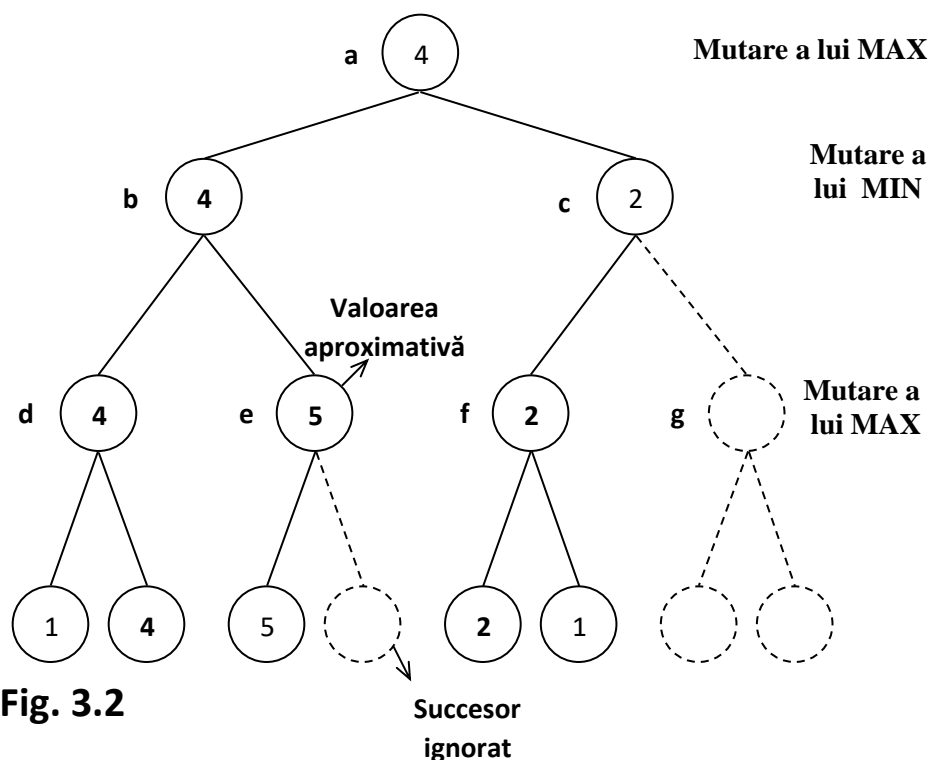
$$V(P, \alpha, \beta) = V(P), \quad \text{dacă } \alpha \leq V(P) \leq \beta$$

$$V(P, \alpha, \beta) > \beta, \quad \text{dacă } V(P) > \beta,$$

unde prin $V(P)$ am notat valoarea de tip minimax corespunzătoare unui nod intern.

Valoarea exactă a unui nod-rădăcină P poate fi întotdeauna calculată prin setarea limitelor după cum urmează: $V(P, -\infty, +\infty) = V(P)$.

Fig. 3.2 ilustrează acțiunea Algoritmului Alpha-Beta în cazul arborelui din Fig. 3.1. Așa cum se vede în figură, unele dintre valorile de tip minimax ale nodurilor interne sunt aproximative. Totuși, aceste aproximări sunt suficiente pentru a se determina în mod exact valoarea rădăcinii. Se observă că Algoritmul Alpha-Beta reduce complexitatea căutării de la 8 evaluări statice la numai 5 evaluări de acest tip:

**Fig. 3.2**

Corespunzător arborelui din Fig. 3.2 procesul de căutare decurge după cum urmează:

1. Începe din poziția *a*.
2. Mutare la *b*.
3. Mutare la *d*.
4. Alege valoarea maximă a succesorilor lui *d*, ceea ce conduce la $V(d) = 4$.
5. Întoarce-te în nodul *b* și execută o mutare de aici la *e*.
6. Ia în considerație primul succesor al lui *e* a cărui valoare este 5. În acest moment, MAX, a cărui mutare urmează, are garantată, aflându-se în poziția *e*, cel puțin valoarea 5, indiferent care ar fi celelalte alternative plecând din *e*. Această informație este suficientă pentru ca MIN să realizeze că, la nodul *b*, alternativa *e* este inferioară alternativei *d*. Această concluzie poate fi trasă fără a cunoaște valoarea exactă a lui *e*. Pe această bază, cel de-al doilea succesor al lui *e* poate fi neglijat, iar nodului *e* i se poate atribui valoarea *aproximativă* 5.

Căutarea de tip alpha-beta retează nodurile figurate în mod discontinuu. Ca rezultat, câteva dintre valorile intermediare nu sunt exacte (nodurile *c*, *e*), dar aproximările făcute sunt suficiente pentru a determina atât valoarea corespunzătoare rădăcinii, cât și variația principală, în mod exact.

Un alt exemplu de aplicare a Algoritmului Alpha-Beta este cel din Fig. 3.4. Astfel, corespunzător spațiului de stări din Fig. 3.3, alpha-beta retezarea produce arborele de căutare din Fig. 3.4, în care stările fără numere nu sunt evaluate.

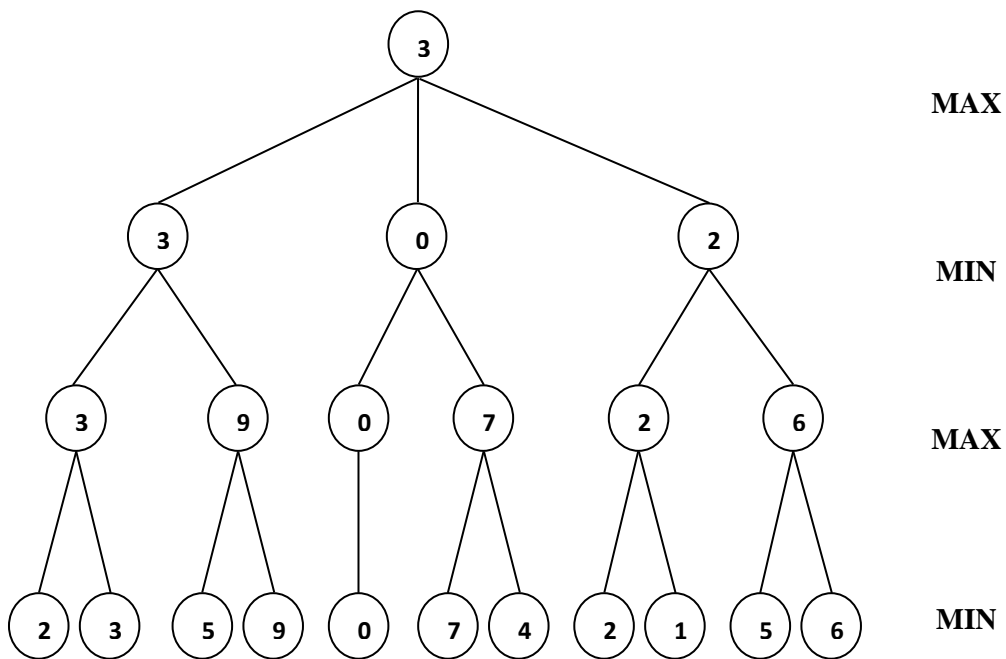


Fig. 3.3

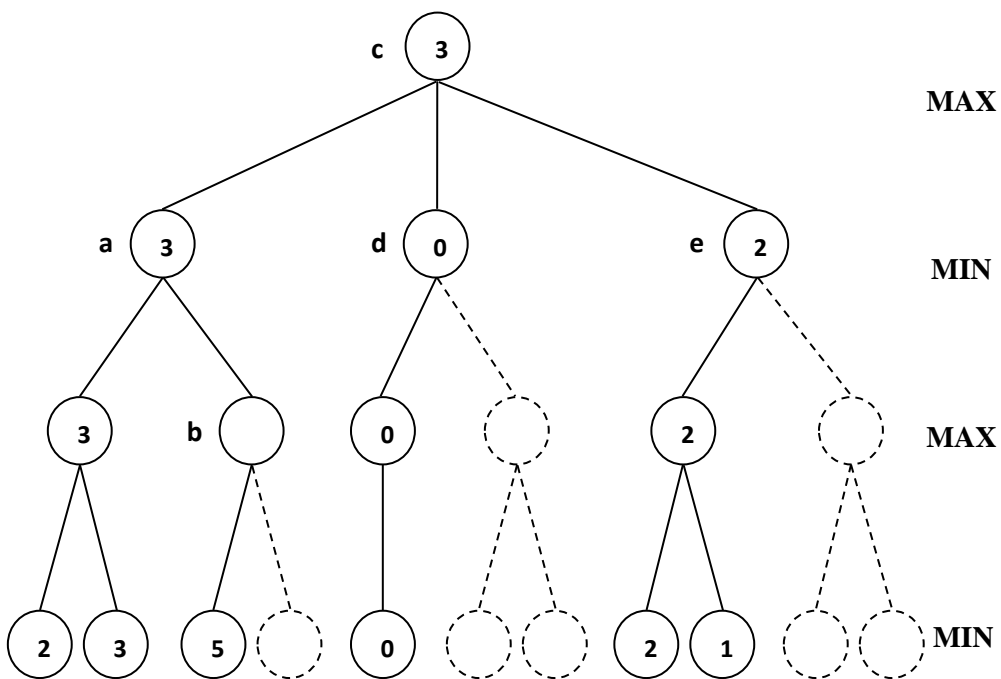


Fig. 3.4

În Fig. 3.4: a are $\beta = 3$ (valoarea lui a nu va depăși 3);

b este β - retezat, deoarece $5 > 3$;

c are $\alpha = 3$ (valoarea lui c nu va fi mai mică decât 3);

d este α - retezat, deoarece $0 < 3$;

e este α - retezat, deoarece $2 < 3$;

c este 3.

Considerații privind Eficiența

Eficiența Algoritmului Alpha-Beta depinde *de ordinea în care sunt examinați succesorii*. Este preferabil să fie examinați mai întâi succesorii despre care se crede că ar putea fi cei mai buni. În mod evident, acest lucru nu poate fi realizat în întregime. Dacă el ar fi posibil, funcția care ordonează succesorii ar putea fi utilizată pentru a se juca un joc perfect. În ipoteza în care această ordonare ar putea fi realizată, s-a arătat că Algoritmul Alpha-Beta nu trebuie să examineze, pentru a alege cea mai bună mutare, decât $O(b^{d/2})$ noduri, în loc de $O(b^d)$, ca în cazul Algoritmului Minimax. Aceasta arată că factorul de ramificare efectiv este \sqrt{b} în loc de b – în cazul jocului de șah 6, în loc de 35. Cu alte cuvinte, Algoritmul Alpha-Beta poate “privi înainte” la o adâncime dublă față de Algoritmul Minimax, pentru a găsi același cost.

În cazul unei ordonări neprevăzute a stărilor în spațiul de căutare, Algoritmul Alpha-Beta poate dubla adâncimea spațiului de căutare (Nilsson 1980). Dacă există o anumită ordonare nefavorabilă a nodurilor, acest algoritm nu va căuta mai mult decât Algoritmul Minimax. Prin urmare, în cazul cel mai nefavorabil, Algoritmul Alpha-Beta nu va oferi nici un avantaj în comparație cu căutarea exhaustivă de tip minimax. În cazul unei ordonări favorabile însă, dacă notăm prin N numărul pozițiilor de căutare terminale evaluate în mod static de către Algoritmul Minimax, s-a arătat că, în cazul cel mai bun, adică atunci când mutarea cea mai puternică este prima luată în considerație, Algoritmul Alpha-Beta nu va evalua în mod static decât \sqrt{N} poziții. În practică, o funcție de ordonare relativ simplă (cum ar fi încercarea mai întâi a capturilor, apoi a amenințărilor, apoi a mutărilor înainte, apoi a celor înapoi) ne poate apropia suficient de mult de rezultatul obținut în cazul cel mai favorabil.