

Gry, strategia minmax i AI

Dr hab. inż. Jerzy Balicki, prof. nadzw.

Plan

1. **Wczesne projekty AI oparte na algorytmach szukania**
2. **Szukanie min-max**
3. **Algorytmy szukania w grach**
4. **Szachy**
5. **Warcaby, tryktrak, go...**

Wczesne projekty

SAINT = Symbolic Automatic INTEGRator

J. Slagle, 1961, praca doktorska, MIT.

Całkowanie symboliczne, przekształcanie wyrażeń.

Program napisany w LISPIe, konieczne były heurystyki.

Redukcja problemów do podproblemów.

Rozwiązał 84 z 86 zdań z egzaminu na MIT.

Np: całka z $(\sec^2 t)/(1+\sec^2 t-3 \tan t) dt$

Po usprawnieniu rozwiązał też $\cos(x^{1/2})dx$, $x(1+x)^{1/2} dx$

SIN, Symbolic Integration, 1967, J. Moses, MIT.

SIN rozwiązywał najtrudniejsze całki.

MACSYMA, a potem **Mathematica**.

Inne projekty

STRIPS (R. Fikes, N. Nilsson, 1971, SRI International)

Planowanie ruchów robota w pokoju ze skrzynkami i pudełkami.

Model świata, plan ruchów.

Opis stanu: za pomocą rachunku predykatów.

Operatory: akcje robota.

Sprawdzanie wstępnych warunków stosowalności.

Wynikiem działania jest zmiana modelu świata.

ABSTRIPS (E. Sacerdoti, 1974)

Ulepszenia: hierarchiczne planowanie, czyli najpierw szkic działania, a potem konstruuje się plan szczegółowy.

Gry

Dobre pole do testów dla prostego podejścia do AI:
Szukanie najlepszego ruchu. Stany, operatory, strategie szukania.

Cechy problemu:

- niepewność związana z ruchem przeciwnika;
- duża przestrzeń szukania.

Mało pieniędzy na badania i rozwój programów do gier, z wyjątkiem szachów i od niedawna gier komputerowych.

Symulatory do gry w szachy zyskały popularność.

Wyzwaniem AI stało się osiągnięcie mistrzostwa w szachach.

Pierwszy program szachowy, 1958, Alex Bernstein - słaby.

Na Olimpiadzie Gier Komputerowych 2007 były turnieje 32 różnych gier, od **Abalone** do **Surakarta**.

W 2009 roku Polacy wygrali w **Counter-Strike**.

Kompromis pamięć/złożoność ocen.

Strategia minimaxu

Teoria gier: John von Neumann, Morgenstern 1944

Twr.

Dla każdej *dwuosobowej gry o sumie zerowej* istnieje wartość V i *mieszana strategia* dla każdego gracza, takie, że:

- (a) biorąc pod uwagę strategię gracza drugiego, najlepszą możliwą spłatą dla gracza pierwszego jest V ,
- (b) biorąc pod uwagę strategię gracza pierwszego, najlepszą możliwą spłatą dla gracza drugiego jest $-V$.

Strategia minimaxu

Oponenci w grze: **Min** i **Max** – **zaczynający**.

1. Utwórz drzewo dla gry do maksymalnej głębokości, na ile starczy czasu.
2. Oceń wartości funkcji heurystycznej poczynając od liści.
3. Cofnij się o jeden poziom i dokonaj ocen znajdujących się tam węzłów.
4. Po osiągnięciu korzenia wybierz decyzję maksymalizującą zyski.

Decyzja min-maks: dla skończonych drzew kompletna, dla racjonalnych oponentów najlepsza.

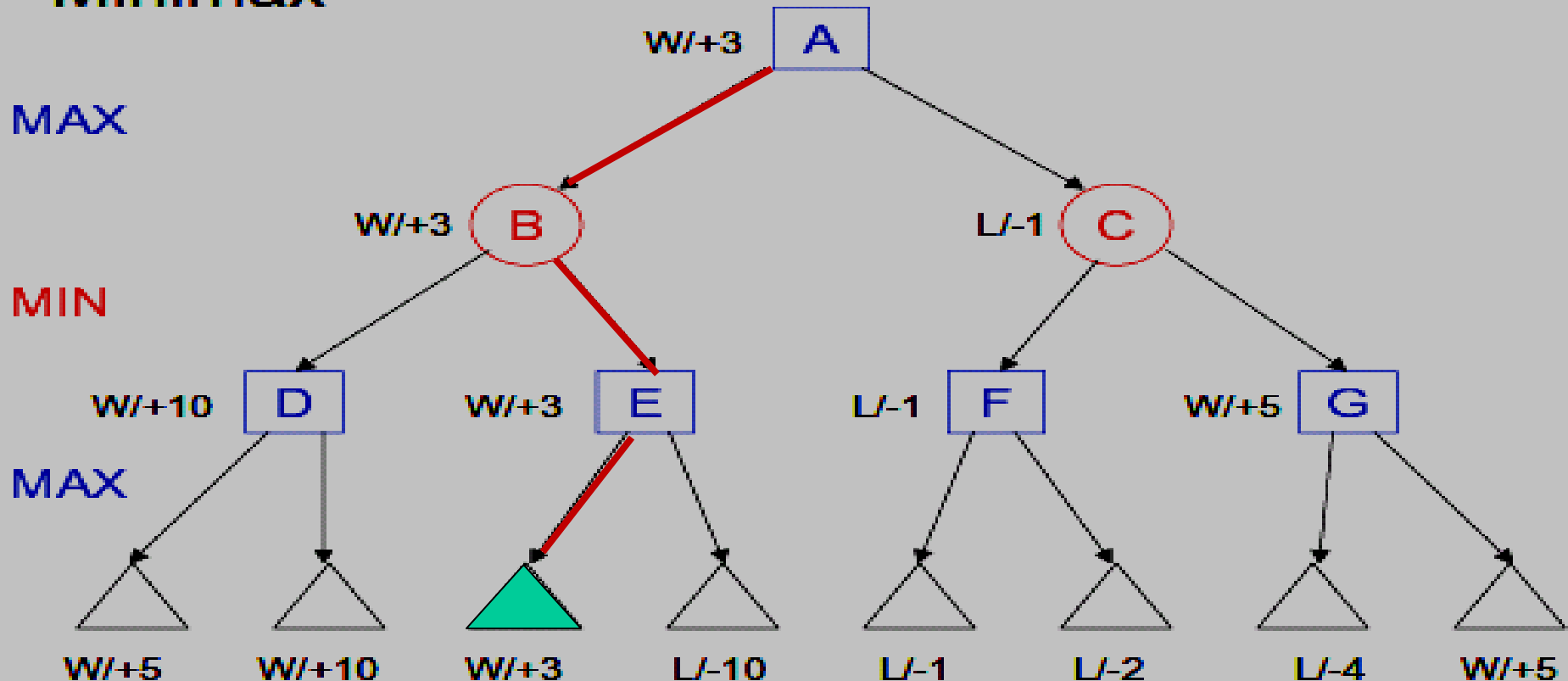
Złożoność $t \sim O(b^m)$, pamięć $O(bm)$ przy szukaniu w głąb.

Przykład mini-maxu

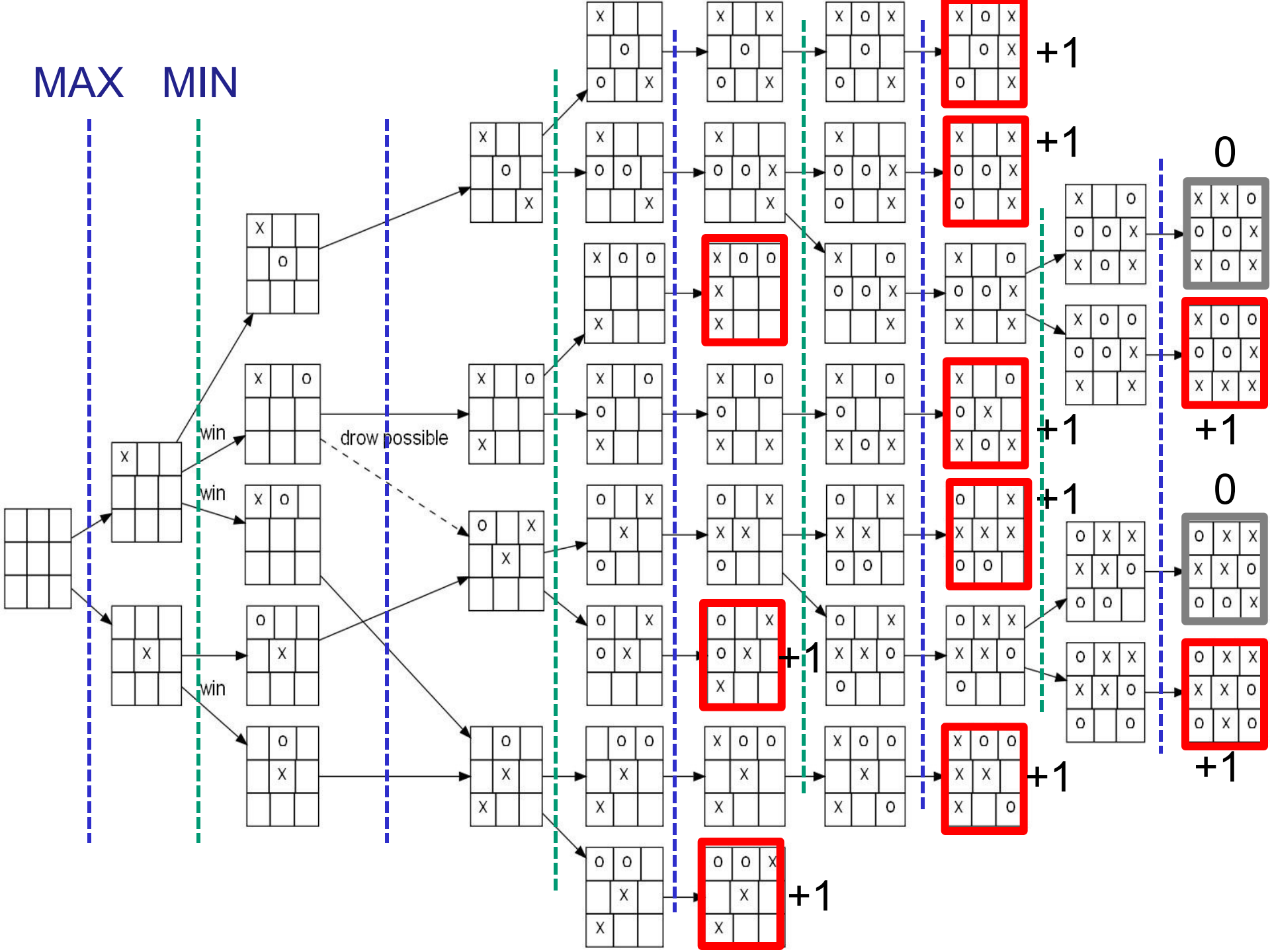
Którą drogę warto wybrać?

Oceniamy liście i cofamy się do góry przenosząc najwyższe lub najniższe oceny na węzły z ruchami dla MAX i MIN.

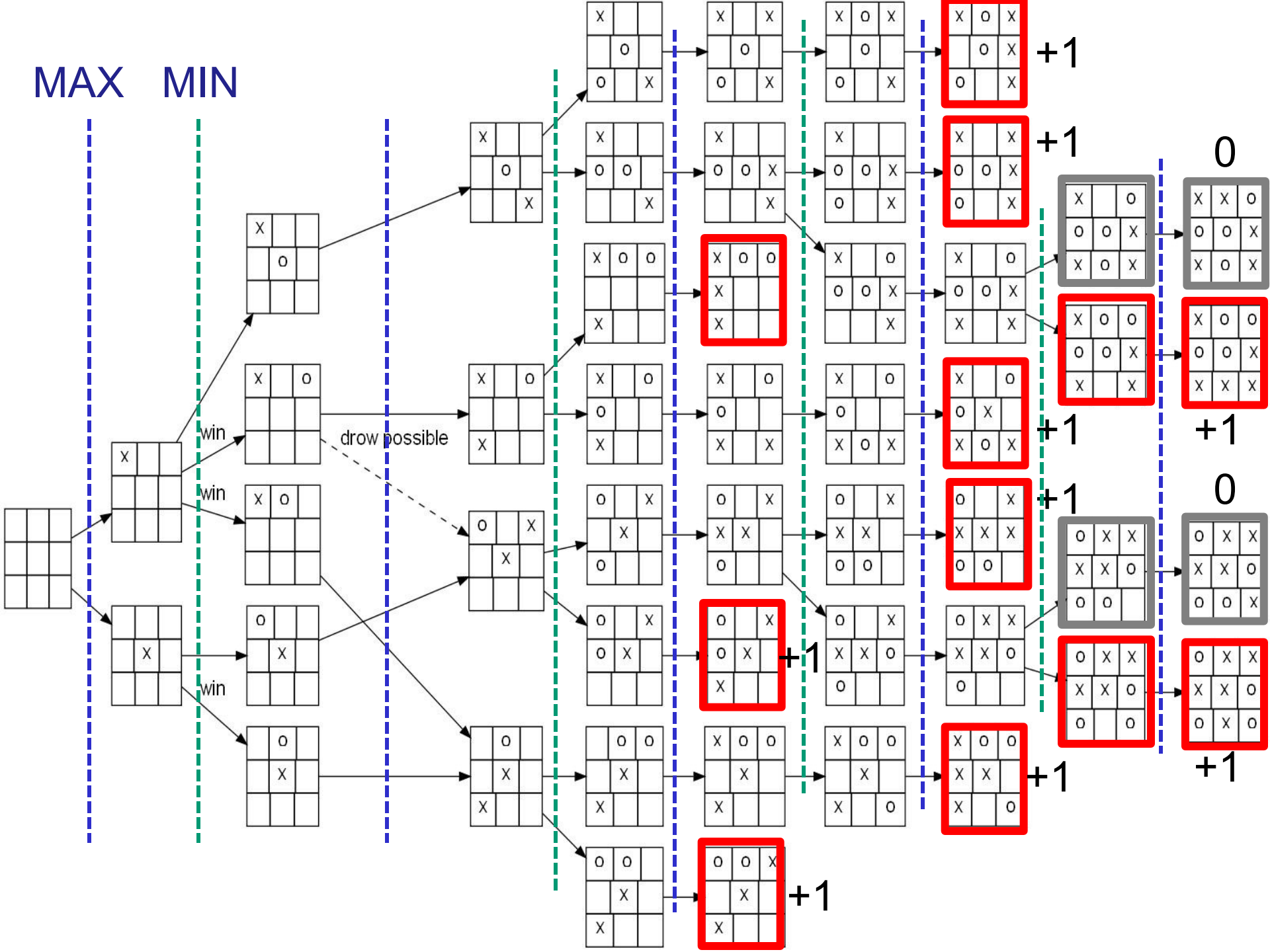
Minimax



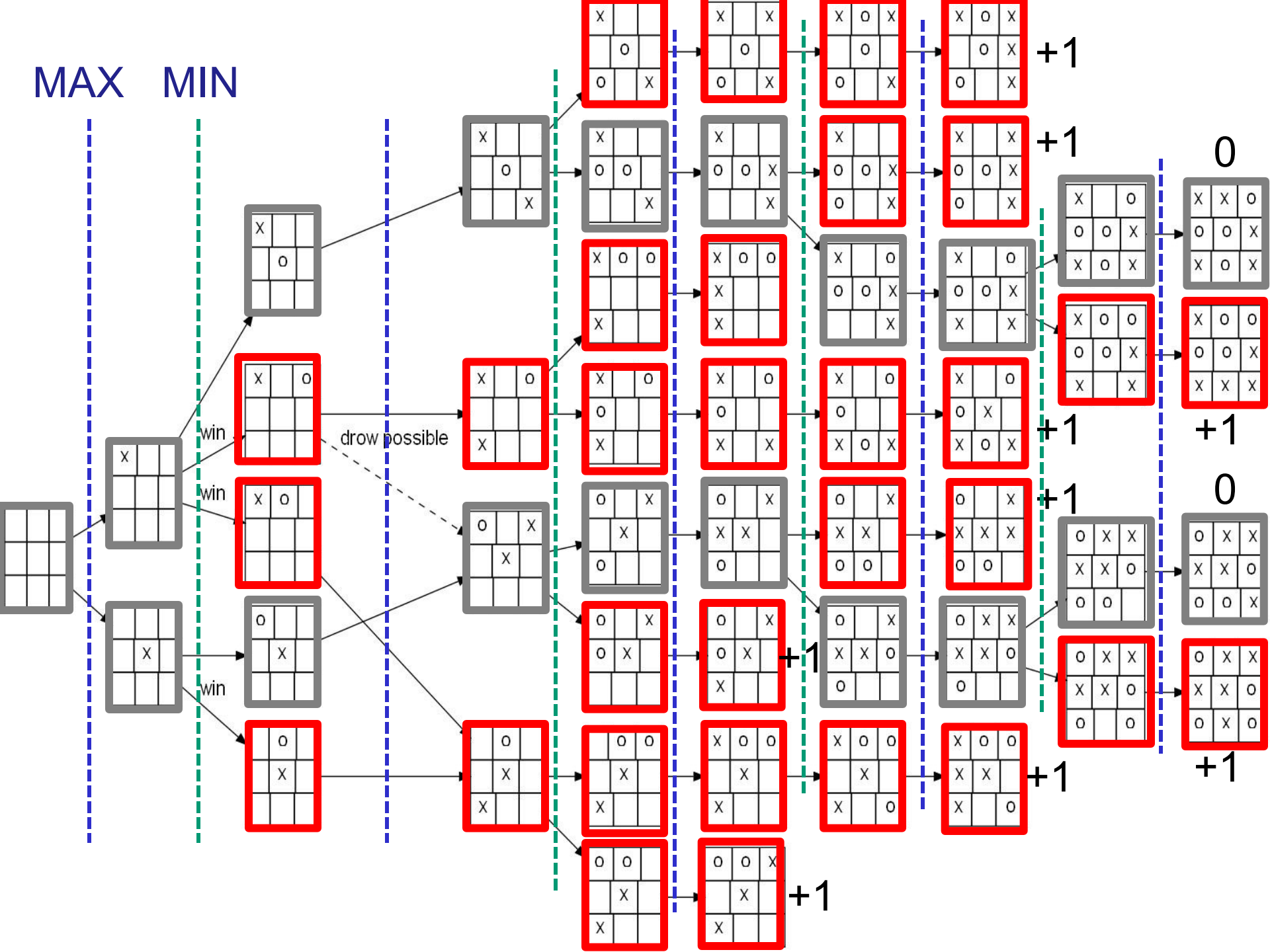
MAX MIN



MAX MIN



MAX MIN



Szachy - historia



1958, pierwszy program szachowy, Alex Bernstein.

Szkocki międzynarodowy mistrz szachowy funduje nagrodę dla programu, który ogra go chociaż raz na cztery partie.

W 1985 roku przegrał wszystkie cztery partie.

1985, program **HiTech** wśród najlepszych 800 graczy, oceniał ok. 10 mln pozycji, w 1988 roku wygrał z arcymistrzem.

Intel+IBM: szachy to dobra reklama.

1994 **Chess Genius** na PC Pentium, kilka razy zwyciężył Gary Kasparova; czas grania ograniczono do 25 minut na zawodnika.

1996 – Deep Blue przegrał z Kasparowem 2:3

1997 – Deep Blue wygrał 3.5:2.5 !!!

2002 – program **Deep Fritz** na PC remisuje z Władimirem Kramnikiem
Przez 2 miesiące Kramnik trenował z programem Deep Fritz.

Deep Thought i Deep Blue

Deep Thought, od 1985 roku, 4 studentów z USA.

Program **Deep Blue** + hardware do gry w szachy:

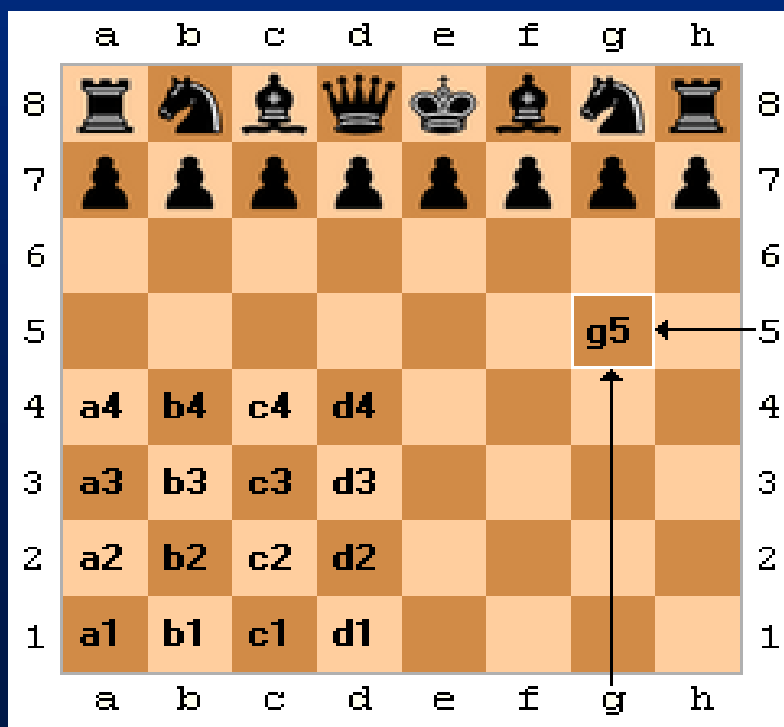
32 procesory IBM RS6000/SP2 + 256 ASIC.

Oceniał **200-1000 milionów pozycji/s.**

Duża biblioteka otwarć i końcówek.

Deep Thought – **szukanie alfa-beta**, ok. 10 ruchów w skomplikowanych sytuacjach.

Deep Blue - ok. 14 ruchów, 3000 punktów.



Ostatni mecz?

Kasparov vs. X3D **Deep Fritz junior**.
Nowy Jork, listopad 2003.

Mecz zakończył się remisem;
atrakcją była gra w wirtualnej
rzeczywistości.



Komputer był około *100 razy wolniejszy* niż Deep Blue
i do tego zajęty generowaniem obrazu w 3 wymiarach,
co spowodowało, że mecz zakończył się **remisem**.

Następny mecz odbędzie się z programem
działającym w telefonie?

Kandydaci

Magnus Carlsen



Viswanathan Anand



Ranking szachowy (ELO, FIDE)

1. Metoda obliczania relatywnej siły gry szachistów.
2. Nazwa ELO pochodzi od nazwiska Arpada Elo – amerykańskiego naukowca węgierskiego pochodzenia.
3. Porównanie przed turniejem rankingu zawodnika i jego przeciwników powinno określić **oczekiwany rezultat (zmienna losowa o rozkładzie normalnym)** tego zawodnika. Jeśli w turnieju uzyskał wynik lepszy od oczekiwanego, jego ranking powinien wzrosnąć, jeśli gorszy – zmaleć.
4. Mistrz międzynarodowy 2400-2500
5. Arcymistrz od 2500
6. Kasparow 2851, Carlsen 2813, Anand 2803, Krasenkow 2702, Polgar 2735 (K), Soćko 2502 (K)

Szachy

Statyczna ocena sytuacji na planszy: **liczba figur, wartość figur, położenie figur, możliwości ruchów.**

Funkcja oceny: suma ważona $W_i F_i$,
dobierane wsp. W_i

Mistrz świata > 2800 punktów.

Przewidywania na:

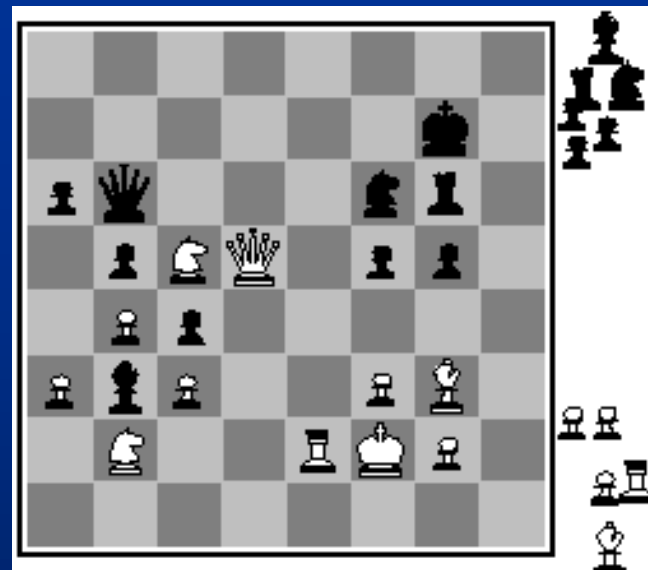
5 ruchów (10 poziomów) to ok. 1500 punktów.

Od 5-10 poziomów mamy 200 punktów/poziom.

Dla 10 ruchów ok. 2500 punktów.

Ok. 35 ruchów/poziom, heurystyki redukują to do 6/poziom;
dla 1000 ocen/s, 150 s/ruch, $b=35$, ok. 3-4 ruchy - nowicjusz.

Jakość gry zależy od szybkości obliczeń.



Funkcje oceny

Wartość figur szachowych może być **absolutna** lub **względna**.

Wartości absolutne: pionek =1, skoczek=3, goniec=3, wieża =5, hetman=9 ...

Pozycja figur: dla każdej konfiguracji dokonuje się innej oceny, np. pionek w pobliżu końcowego pola jest dużo ważniejszy niż pionek zablokowany.

Ocena następnego ruchu, zagrożeń dla poszczególnych figur.

Złożone oceny konfiguracji wielu figur.

Kombinacja liniowa różnych elementów oceny f_i z wagami W_i

$$E(f) = \sum_i W_i f_i$$

Współczynniki W_i dobiera się tak, by maksymalizować zyski. Nieliniowe funkcje oceny mogą dać lepsze rezultaty.

Precyzyjne wartości f_i nie mają znaczenia - liczy się tylko względny porządek, jest to więc „porządkowa funkcja kosztu”.

Przewaga materialna a przewaga pozycyjna

- W konkretnej sytuacji na szachownicy wartość względna jest ważniejsza od wartości absolutnej.
- W określonych sytuacjach figury mogą mieć różne wartości.
- W pozycjach o charakterze zamkniętym względna wartość skoczka przewyższa wartość gońca.
- W pozycjach otwartych wartość gońca jest większa niż skoczka (a tym bardziej pary gońców nad parą skoczków).
- Dochodzący do pola przemiany pion może mieć wartość o wiele wyższą niż przypisana mu wartość 1.
- Dwie wieże mają przewagę nad hetmanem, mimo że gra tym ostatnim jest z reguły łatwiejsza.
- Wartości punktowe poszczególnych figur mają charakter umowny i zależą od konkretnej pozycji na szachownicy

Przewaga materialna a przewaga pozycyjna

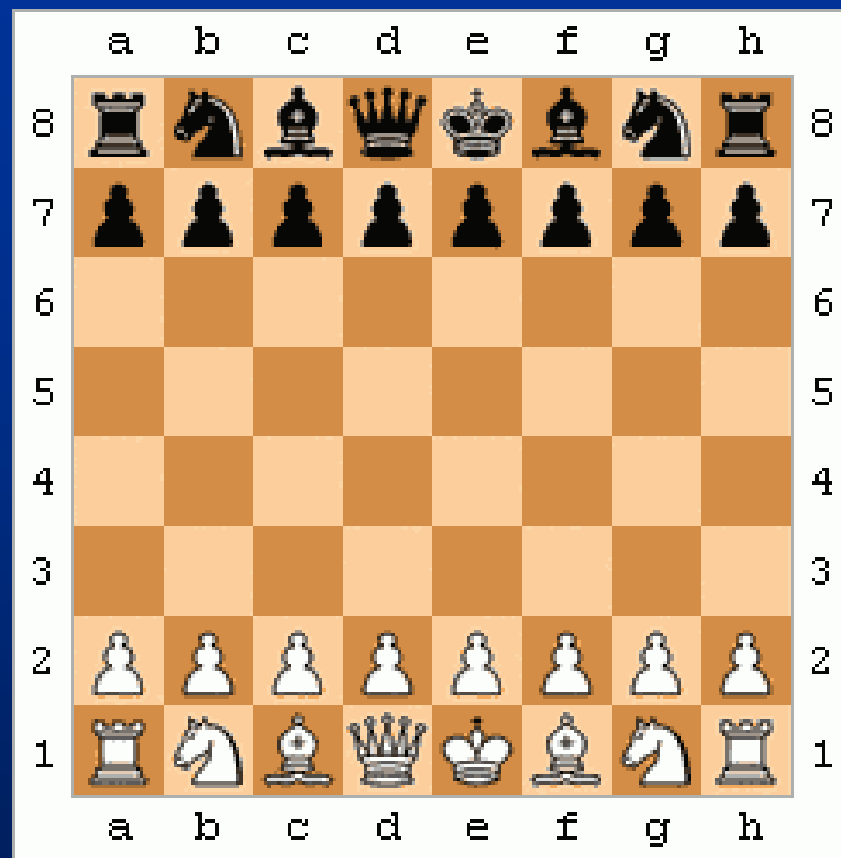
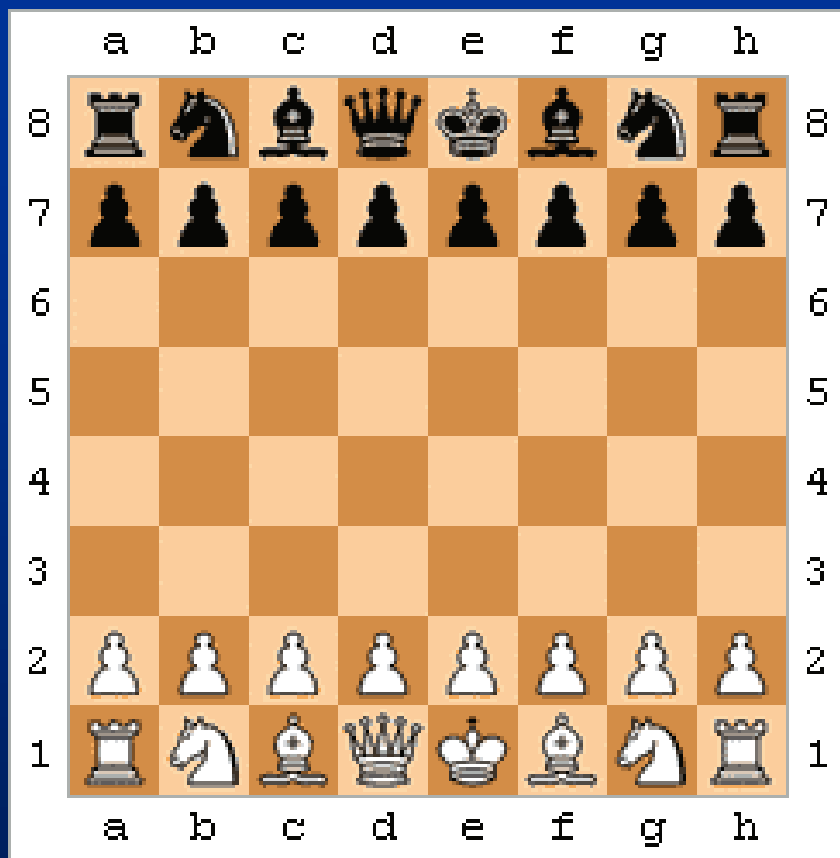
Przy przewadze materialnej powinno dążyć się do uproszczenia pozycji (poprzez wymianę figur) sprowadzając grę do końcówki.

Wartość punktowa pionka rośnie wraz ze zbliżaniem się do pola przemiany.

Wartość skoczka zależy od charakteru pozycji:
w zamkniętych pozycjach skoczek (3,4) ma przewagę nad gońcem (3), w otwartych pozycjach na odwrót (skoczek – 1,7, goniec – 3); wszystko zależy od konkretnego ustawienia figur na szachownicy.

Wartość króla rośnie wraz ze zmniejszaniem się liczby figur na szachownicy: jest on wtedy znacznie mniej narażony na bezpośredni atak matowy przeciwnika.

Przebieg partii

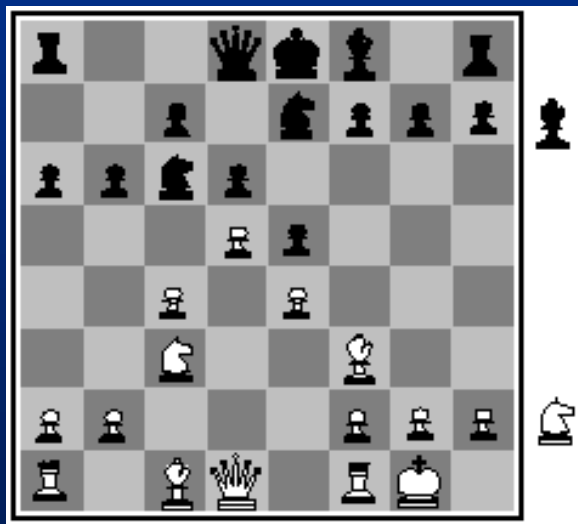


Szacowana liczba sytuacji na szachownicy 10^{123}

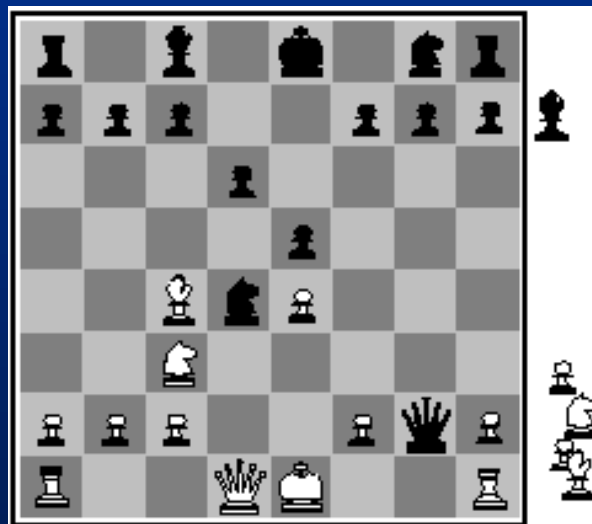
Przykłady ocen

Szachiści oceniają jakościowo, wystarczają względnego porównania.

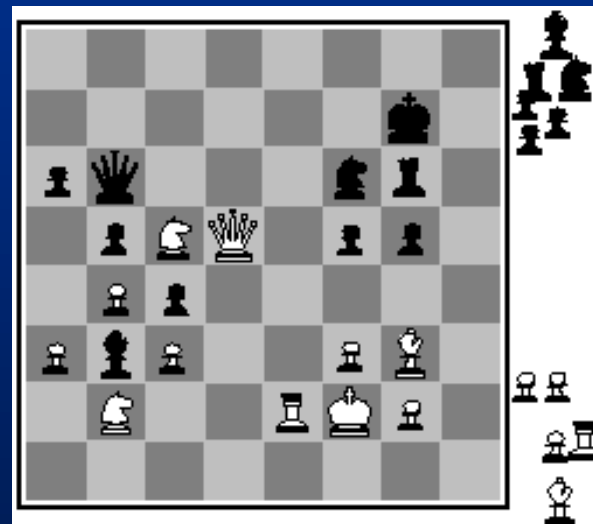
Na podstawie doświadczenia tworzy się złożone funkcje oceny.



Ruch czarnych
Przewaga białych



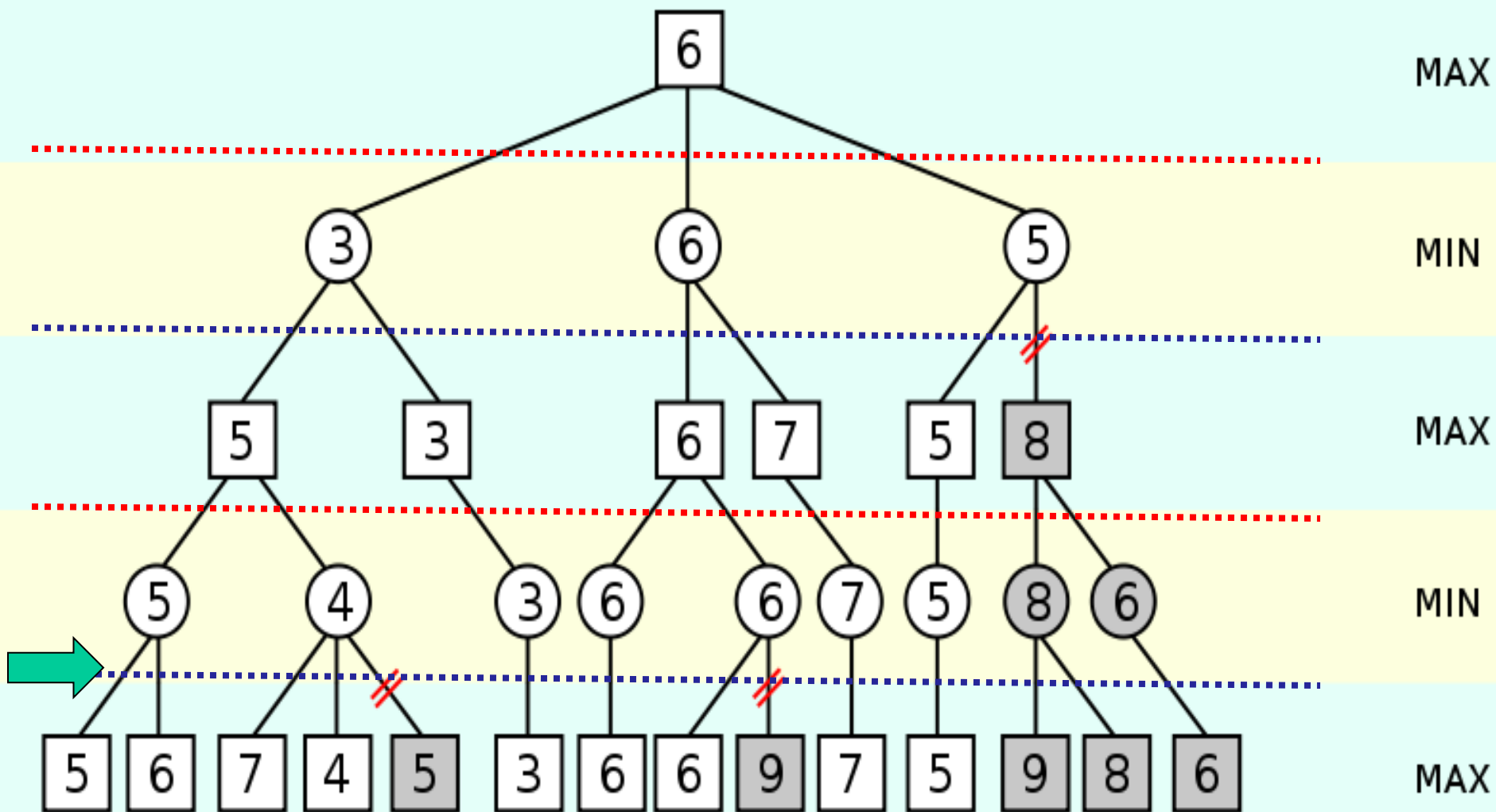
Ruch białych
Czarne wygrywają



Ruch czarnych
Białe bliskie przegranej

Obcinanie alfa-beta

W algorytmie min-max obcina się gałęzie, które nie mają wpływu na końcowy wybór.



Programy do gry w szachy

- <http://chessdb.sf.net> – baza danych partii szachowych
- <http://www.tim-mann.org/chess.html> - **Winboard** jest wykorzystywany jako interfejs dla silników szachowych, umożliwia też grę na serwerach szachowych.
- <ftp://ftp.cis.uab.edu/pub/hyatt/> , <http://ruffian.hkust.se/>
<http://members.home.nl/matador/prodeo.htm> - **Crafty** i **Ruffian** to bardzo silne , konkurujące z komercyjnymi, silniki szachowe. Umożliwią rozegranie oraz analizę partii na poziomie arcymistrzowskim.
- **ProDeo** to bezpłatna kontynuacja enginu **Rebel 12**, szacowana siła gry to: **2700 elo**.

Rybka

- <http://rybkachess.com/> - **Rybka** to najsilniejszy engin na świecie;
- bezpłatna wersja **Rybka beta** nie ma sobie równych wśród silników komercyjnych;
- komercyjna wersja 2.1 gra z niedostępną dla nikogo siłą **3000 elo**.

Rybka

1. GUIs:

- Aquarium, · Arena, · Chess Assistant
- Chessbase, · Shredder, · Chess Partner

2. **Vasik Rajlich (ur. 1971 w Cleveland) – szachista, mistrz międzynarodowy. Ożenił się z Iwetą Radziejwicz, polską arcymistrzynią szachową.**



Warcaby

- 1952, Samuel, pierwszy program uczący się gry w warcaby.
- 1992, **Chinook** (J. Schaeffer, UoA) wygrywa US Open.
Program używa **szukania $\alpha-\beta$**
- Mistrzostwa człowiek-maszyna, Londyn 1992
Dr. Marion Tinsley, wygrał z Chinookiem 4-2, 33 remisy.

Użyto 8-proc. stacji Silicon Graphics 4D/480, 256 MB RAM, baza danych wszystkich końcówek z 1-7 figurami + prawie połowa wszystkich partii z 8 figurami.

- 1994, remis 1-1 i 18 remisów.
- **1995**, wygrana Chinooka 1-0 i 31 remisów.
Użyto komputera SGI z 512 MB RAM.

Obcinanie alfa-beta w warcaby

Dokładne ocenienie ruchu bez osiągnięcia końcowego stanu jest niemożliwe. Dlatego efektywność posunięcia musimy oszacować. Oszacowania tego dokonuje się za pomocą funkcji oceniającej aktualne ustawienie.

Oszacowaniem jest różnica między liczbą białych a czarnych figur (im więcej białych figur tym ruch lepszy dla grającego białymi).

Należy wprowadzić wartościowanie (pionek damce nie jest równy), dlatego oszacowano, iż wartość damki jest trzy razy większa niż pionka. Czas pojedynczego wykonania funkcji jest krótki.

Ważne jest , aby znaleźć kompromis między dokładnością oceniania a liczbą analizowanych ruchów.

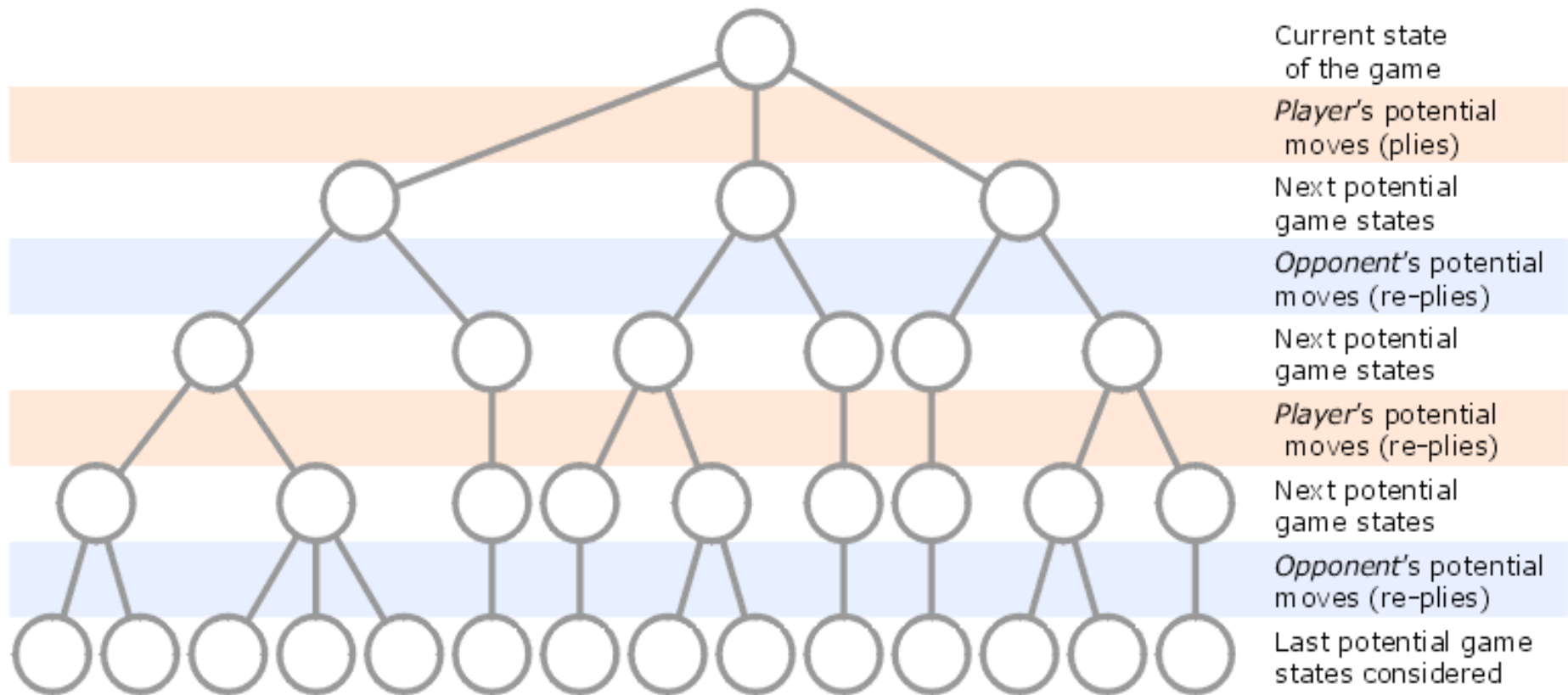
Obcinanie alfa-beta w warcaby

Nie trzeba przeszukiwać wszystkich gałęzi drzewa (oczywiście do pewnej ustalonej głębokości).

Jeśli Max ma wykonać ruch, po którym Min będzie mógł wykonać jak najgorsze ruchy, to nie trzeba przeszukiwać wszystkich gałęzi.

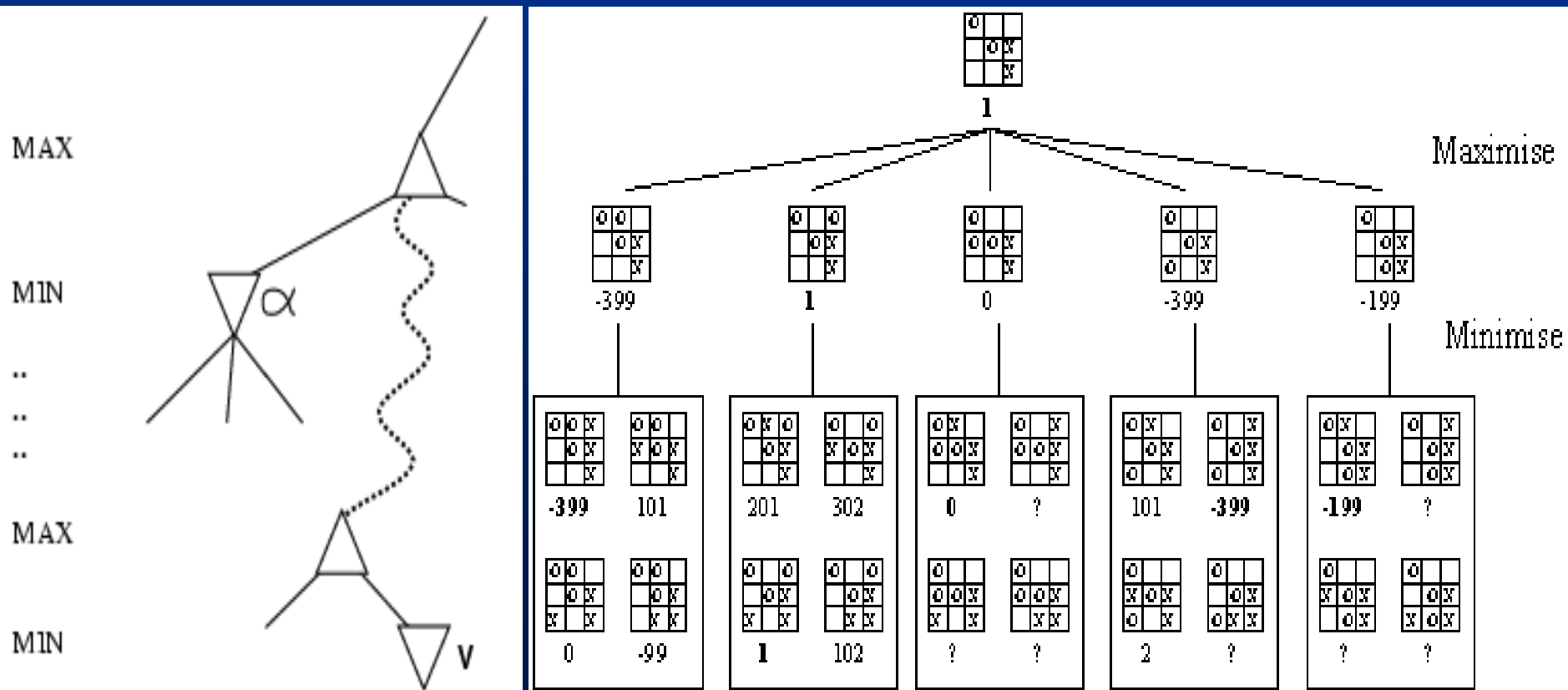
Powinniśmy zadbać o to, aby pierwszy analizowany ruch był jak najlepszy.

Minimax with alpha-beta pruning on a two-person game tree of 4 plies



Obcinanie alfa-beta

- α - najlepsza wartość dla MAXa dla kolejnych kroków;
- Porzuć V jeśli są lepsze wartości α w innej części grafu (dla MAX);
- β - najlepsza wartość dla MINa dla kolejnych kroków;
- oszczędność w najlepszym razie $t \sim O(b^{m/2})$



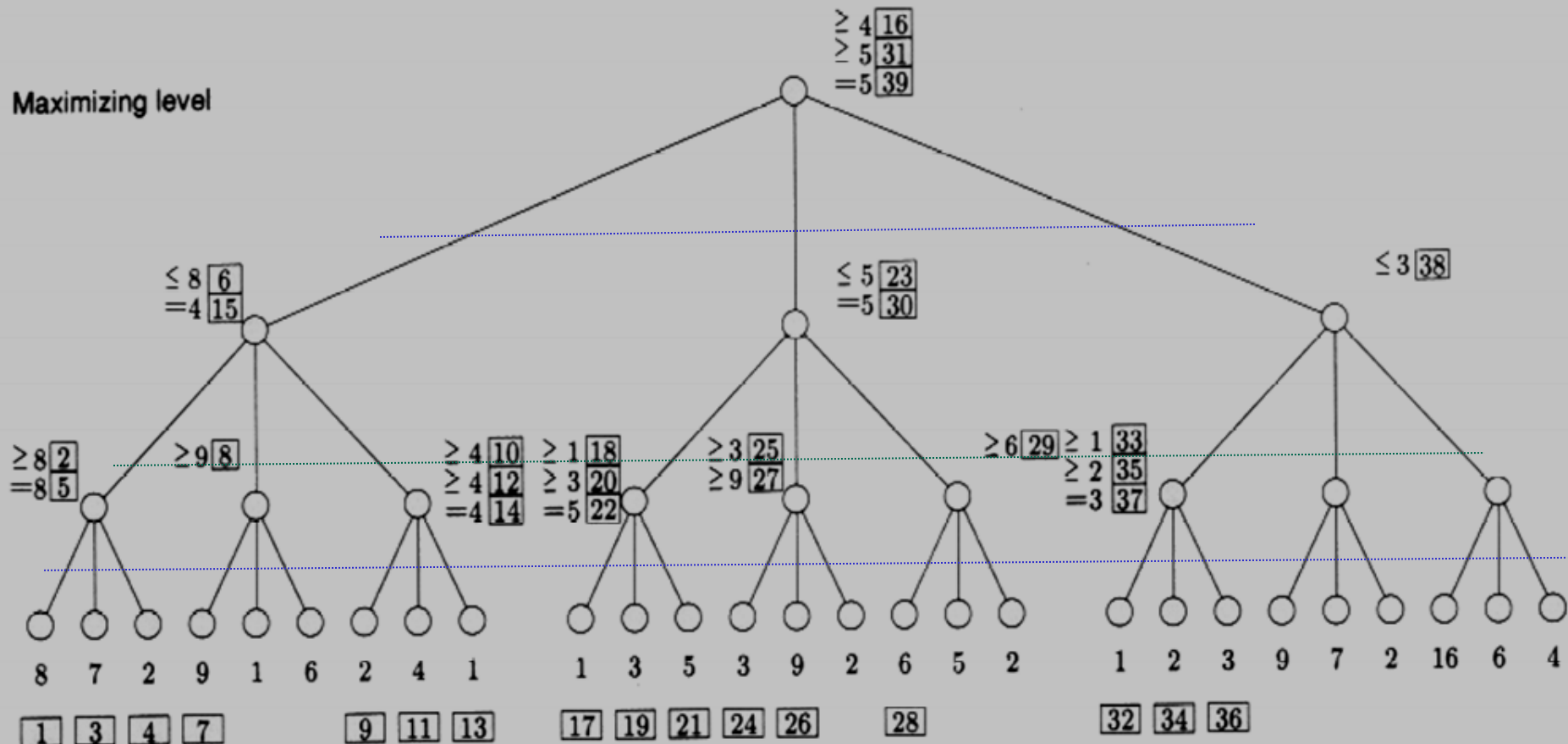
Przykład zastosowania α - β

Kolejność odwiedzanych węzłów zaznaczono w kwadratach, a wartości funkcji heurystycznej obok.

Węzeł nr 1 ma $f=8$; węzeł nr 2 ma $f \geq 8$; po sprawdzeniu $=8$;

Węzeł nr 2 ma $f \leq 8$, po sprawdzeniu węzła nr 9 reszta węzłów jest pomijana.

W gałęzi (2,4,1) max $f=4$, a więc na tyle może liczyć MIN.

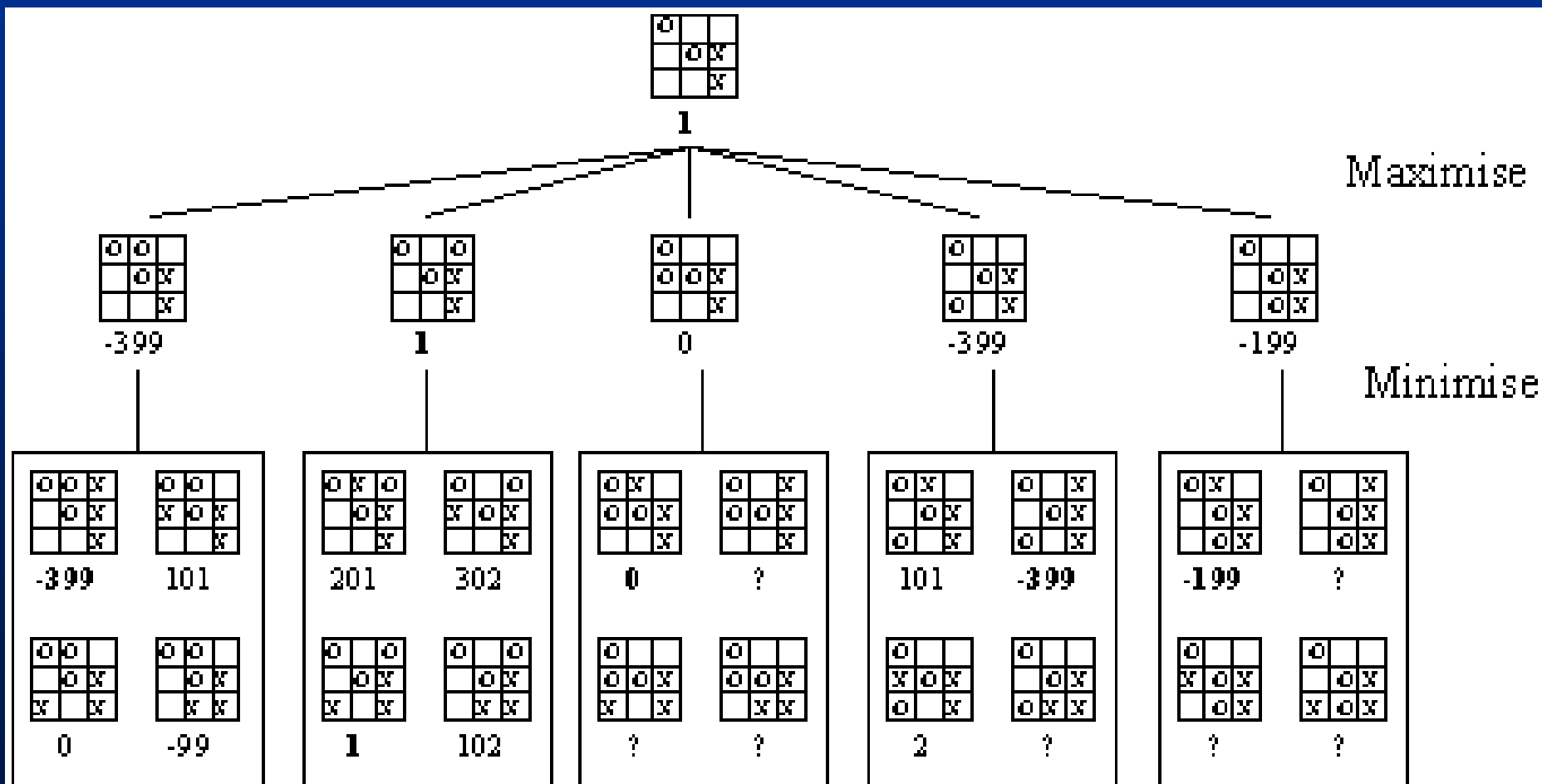


Obcinanie alfa-beta

Ignoruj mało obiecujące gałęzie.

α – kres dolny do oceny **Max**,

β – kres górny do oceny **Min**.



Tryktrak

- **Trik-trak (backgammon), popularny w Japonii. Zawiera element niepewności (rzut kostką).**
- **1980, program BKG wygrał z mistrzem świata.**
- **1992, program Tesauro, techniki jak dla warcabów + uczenie się (sieć neuronowa) lepszych ocen, ranking wśród 3 najlepszych graczy.**
- **1995, Logistello zwycięża mistrza świata Takeshi Murakami 6:0 !!!**

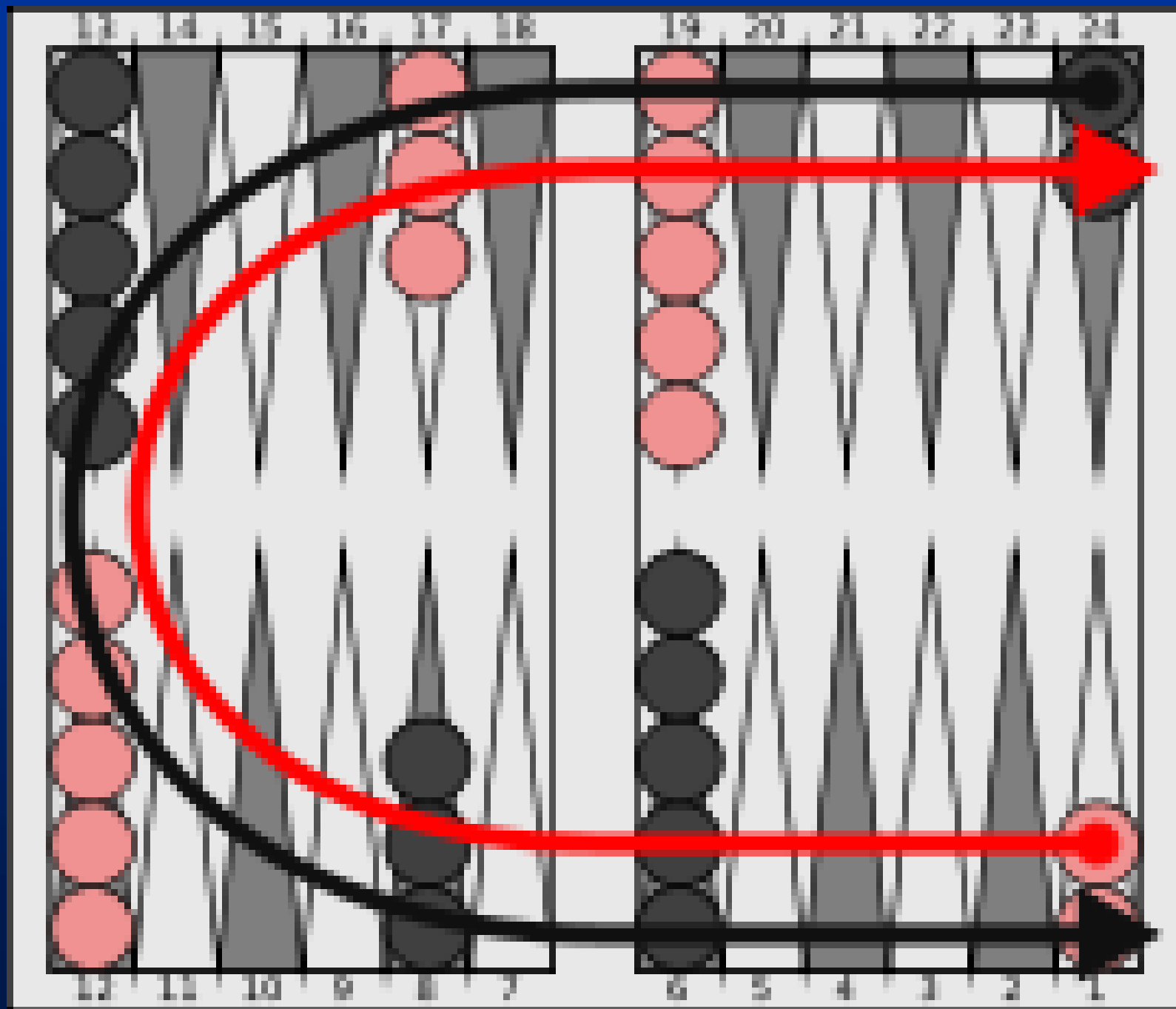
Program grał parę tygodni sam z sobą poprawiając swoje funkcje oceny heurystycznej.

**Mistrzowskie rezultaty w wielu grach osiągnięto w latach 1990.
Moc obliczeniowa to warunek konieczny, ale nie wystarczający ...**

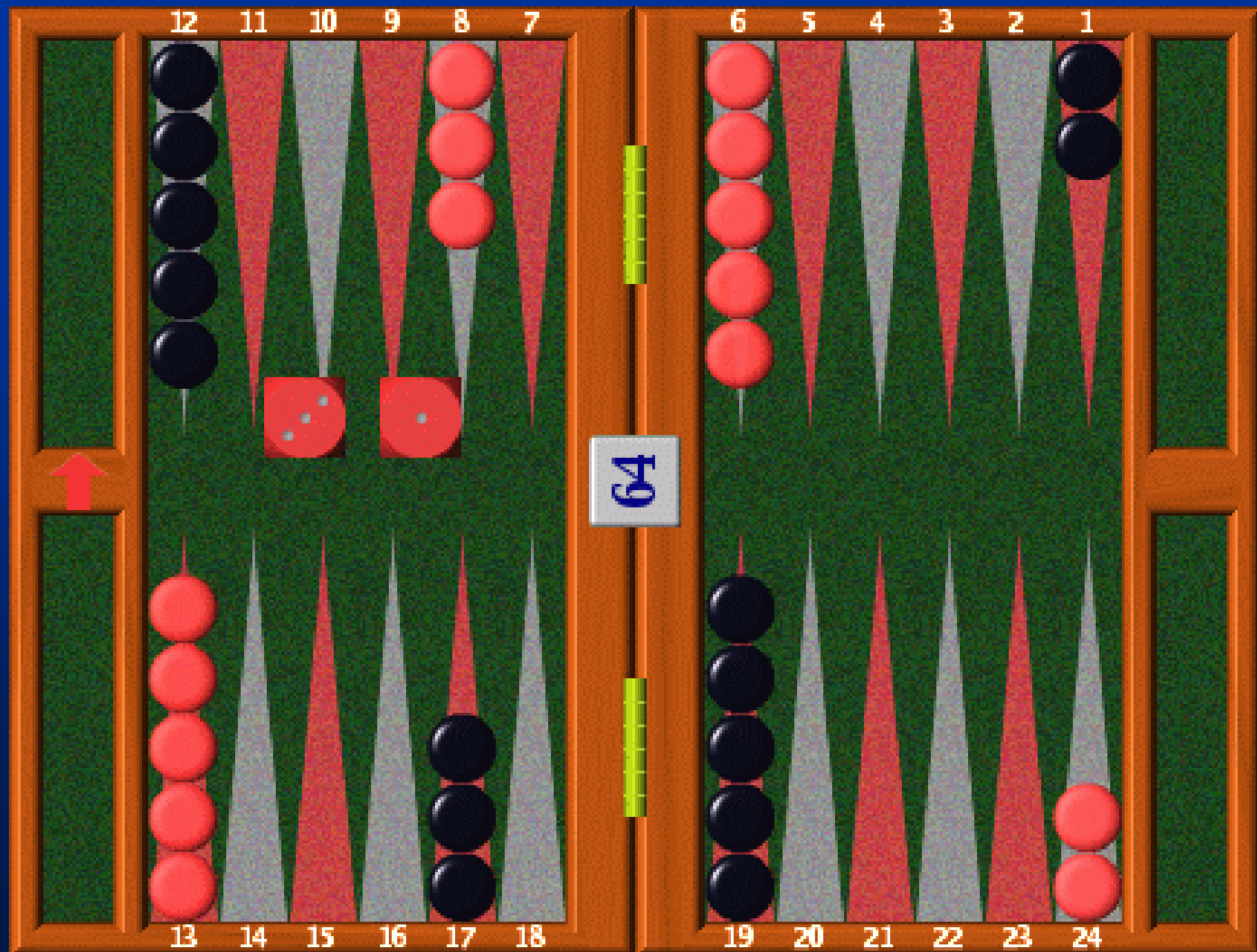
Tryktrak

- (ang. *backgammon*; fr. *tric-trac*) – gra planszowa dla dwóch graczy.
- Każdy z graczy dysponuje 15 pionami w różnych kolorach, które przesuwa po planszy, składającej się z 24 pól (trójkątów), nazywanych liniami, zgodnie z liczbą oczek wyrzuconych na dwóch kostkach.
- Celem gry jest zdjęcie wszystkich swoich pionów z planszy. Wygrywa gracz, który uczyni to pierwszy.









- Gracz posługujący się pionkami czarnymi porusza się po polach oznaczonych na rysunku od liczby 1 do 24.
- Gracz posiadający pionki czerwone porusza się w przeciwnym kierunku - od 24 do 1.
- Po rzucie może poruszyć swe dwa pionki o liczby wskazane na kostkach (w tym jednym pionem dwa razy).
- Wyjątkiem jest "dubel" (ta sama ilość oczek na obu kostkach) - w tym przypadku gracz może poruszyć cztery, a nie dwa pionki o liczbę oczek wskazanych na obu oczkach.

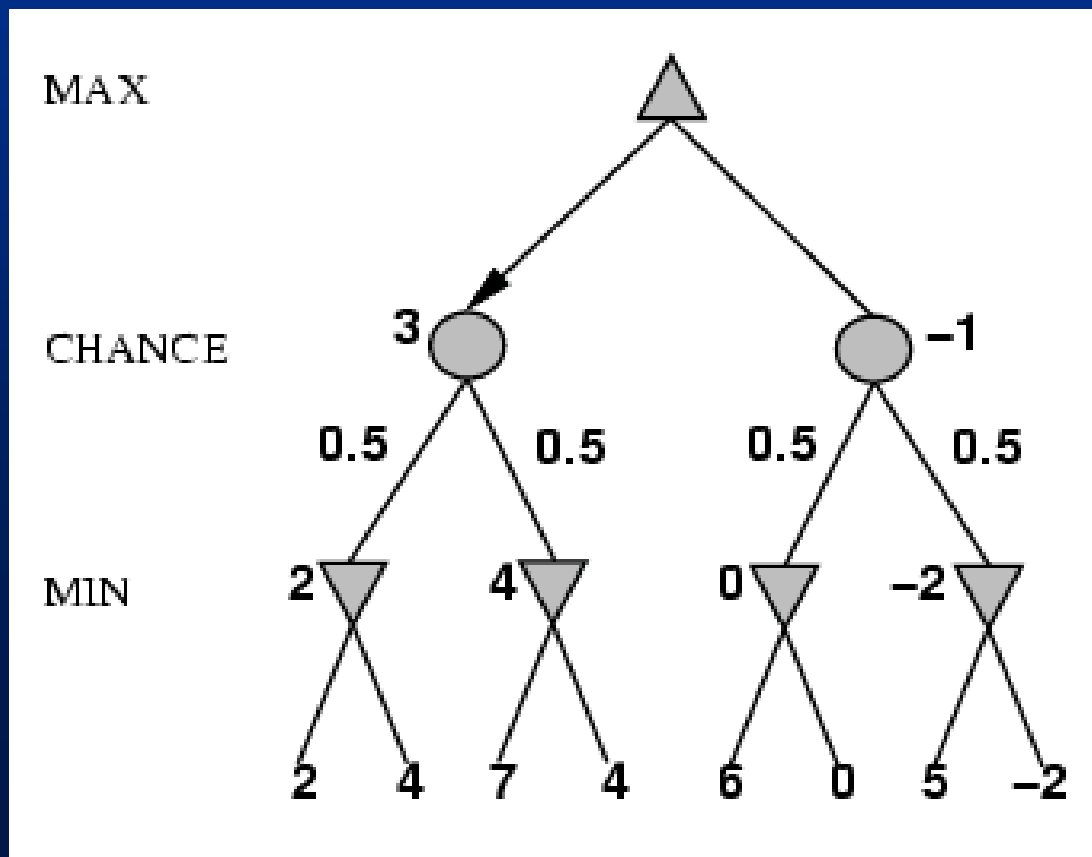
- Z pola nr 1 można w jednym ruchu wyprowadzić tylko jeden pion.
- W pierwszej fazie gracz musi doprowadzić wszystkie swe pionki do "domu" tzn. pól oznaczonych numerami 20-24.
- Dopiero wtedy może pozbywać się pionków z planszy. Odbywa się to poprzez wyrzucenie kostką ilości oczek potrzebnej pionkowi do opuszczenia planszy.
- W czasie gry można zbijać. Zbicie polega na umieszczeniu swego pionka na polu na którym już znajduje się pionek przeciwnika (dokładnie jeden).
- Po biciu pionek bijący nie może się poruszać w tym ruchu, pionek zbity jest odkładany z planszy i przeciwnik nie może kontynuować gry jeśli jego pionek najpierw nie wróci na planszę.
- By pionek wrócił na planszę trzeba rzucić kostką i rozpocząć grę tym pionkiem jak gdyby znajdował się on na polu zerowym.

- W grze podstawową umiejętnością jest takie poruszanie pionkami, aby nigdy nie pozostawały samotne (nie można ich wtedy zbić).
- Gracz stara się zablokować jak najwięcej pól w swym "domu", a następnie zbić pion przeciwnika - wtedy graczowi trudno będzie odzyskać pion (w najlepszej sytuacji można zablokować cały dom - wtedy przeciwnik musi czekać aż gracz zacznie opuszczać planszę).

Gry niedeterministyczne

Strategia min-max połączona z oceną probabilistyczną szans na generowanie kolejnego ruchu.

Obcinanie α - β można stosować, ale jest znacznie mniej efektywne, wzrasta liczba możliwych rozgałęzień.



Go: większe wyzwanie

Liczba ruchów w Go to średnio 260 (szachy tylko 35).

Liczba możliwych partii: 10^{260} (szachy 10^{123}).

Liczba różnych pozycji na planszy: 10^{172} (szachy 10^{46}).

Techniki szukania są mało przydatne: obecne programy są kiepskie na standardowej planszy 19x19, ale na poziomie mistrzowskim na planszach 9x9.

Pierwszy program w 1968 roku, a wiodące obecnie programy to: Go4++ napisał M. Reiss, MogoTW (Ojima Yoji) i Zen (Yamato).

Konieczne do dobrego grania w Go jest:

- rozpoznawanie struktur (typ lokalnych konfiguracji),
- uczenie maszynowe,
- strategie i planowanie,
- metody reprezentacji wiedzy.

Nagroda 1 mln \$ dla programu, który pokona mistrza z Taiwanu!

Go: większe wyzwanie

- Szachownica ma tylko 64 pola, zaś deska do Go - 361 przecięć, na które można stawiać piony;
- Tylko pierwsze dwa ruchy można wykonać na 129 960 sposobów, a pierwsze 3 ruchy na 46 i pół miliona sposobów;
- Celem gry jest zdobycie terytorium.
- Dwaj gracze ustawiają na przemian piony na planszy, wpierw w rogach i na brzegach, a później łączą je w łańcuchy, które zamykają teren.
- Raz postawionego piona nie wolno już przesuwować.
- Całkowicie otoczone kamienie przeciwnika zostają zbite i zdjęte z planszy.
- Wszystkie kamienie mają tę samą wartość, jeden samotny kamień jest bezwartościowy.

Go: większe wyzwanie

- Kamienie w swym ustawieniu przypominają zdyscyplinowane wojsko, które rozbija wrogie szyki.
- W trakcie ataku wdzierają się na terytorium przeciwnika, aby pochwycić jego kamienie.
- W szachach przegrywa zazwyczaj ten z graczy, który jako pierwszy popełni błąd. W Go można, pomimo wcześniej popełnionych błędów, zebrać siły na innym froncie, aby ostatecznie pokonać przeciwnika.
- Gracz może w celu wygrania zupełnie świadomie poświęcić grupy swych pionów.
- Pod koniec partii, przed podliczeniem wyniku, każdy z graczy ustawia kamienie zbite przez siebie na terenie przeciwnika.
- Zwycięzcą staje się ten, kto dysponuje większym terenem.

Go: przykładowa partia

