Politechnika Wrocławska

Wyznaczanie strategii w grach

Dariusz Banasiak

Katedra Informatyki Technicznej W4/K9

Politechnika Wrocławska

Definicja gry

Teoria gier i konstruowane na jej podstawie programy stanowią jeden z głównych wyznaczników postępu sztucznej inteligencji. Głównie ze względów komercyjnych spore nakłady przeznaczane są na badania dotyczące szachów (np. Intel i IBM zaangażowały w konstrukcję komputerów szachowych oraz w badania nad rozwojem algorytmów dla gier duże środki).

Pierwszy program szachowy został napisany w roku 1958 (Alex Bernstein).

W roku 1997 Kasparow podczas pojedynku z Deep Blue na sześć partii dwie przegrał, trzy zremisował i tylko jedną wygrał.

Definicja gry

Gra jest to rozgrywka (dowolna sytuacja konfliktowa) prowadzona przez graczy, zgodnie z ustalonymi zasadami, w celu osiągnięcia określonego celu.

Gra składa się z zestawu reguł określających możliwości wyboru postępowania jednostek (graczy).

Rozpoczyna się ona od pewnego stanu początkowego i kończy się stanem, który według określonego kryterium może być uznany za wygrany, przegrany lub remisowy (z punktu widzenia jednego z graczy). Z każdym ze stanów końcowych można związać pewną funkcję wypłaty.

Definicja gry

Charakterystyczną cechą gry jest istnienie sytuacji konfliktowej (konflikt interesów):

każdy gracz stara się maksymalizować swój własny zysk i jednocześnie zminimalizować zysk przeciwnika.

Podział gier ze względu na liczbę graczy:

- bezosobowe (np. gra w życie)
- jednoosobowe (np. puzzle, pasjans)
- dwuosobowe (np. warcaby, szachy, kółko i krzyżyk)
- wieloosobowe (np. brydż, poker, chińczyk)

Podział gier ze względu na liczbę ruchów:

- jednochodowe (np. papier- kamień-nożyce)
- wielochodowe (np. szachy, warcaby, brydż)

Podział gier ze względu na postać funkcji wypłaty:

- gry o sumie zerowej (wygrana jednego gracza oznacza przegraną drugiego gracza)
- gry o sumie niezerowej

Definicja gry

Podział gier ze względu na posiadaną informację:

- gry o pełnej informacji (gracz wybierając kolejny ruch posiada pełną informację o aktualnej sytuacji oraz możliwościach przeciwnika np. warcaby, szachy)
- gry o niepełnej informacji (np. poker, brydż, domino)

Podział gier ze względu na występowanie elementu losowości:

- całkowicie losowe (np. ruletka)
- częściowo losowe (np. brydż, domino, chińczyk)
- deterministyczne (warcaby, szachy, go)

Definicja gry

Każdą z gier można scharakteryzować za pomocą przyjętych kryteriów:

- szachy to gra dwuosobowa, wielochodowa, o sumie zerowej, o pełnej informacji i deterministyczna
- ruletka to gra jednoosobowa, jednochodowa, o sumie zerowej, o niepełnej informacji oraz całkowicie losowa

Celem wykładu jest przedstawienie metod opisu oraz poszukiwania strategii dla gier dwuosobowych, wielo-chodowych, o sumie zerowej, o pełnej informacji i deterministycznych.

Proces poszukiwania strategii gry można opisać drzewem.

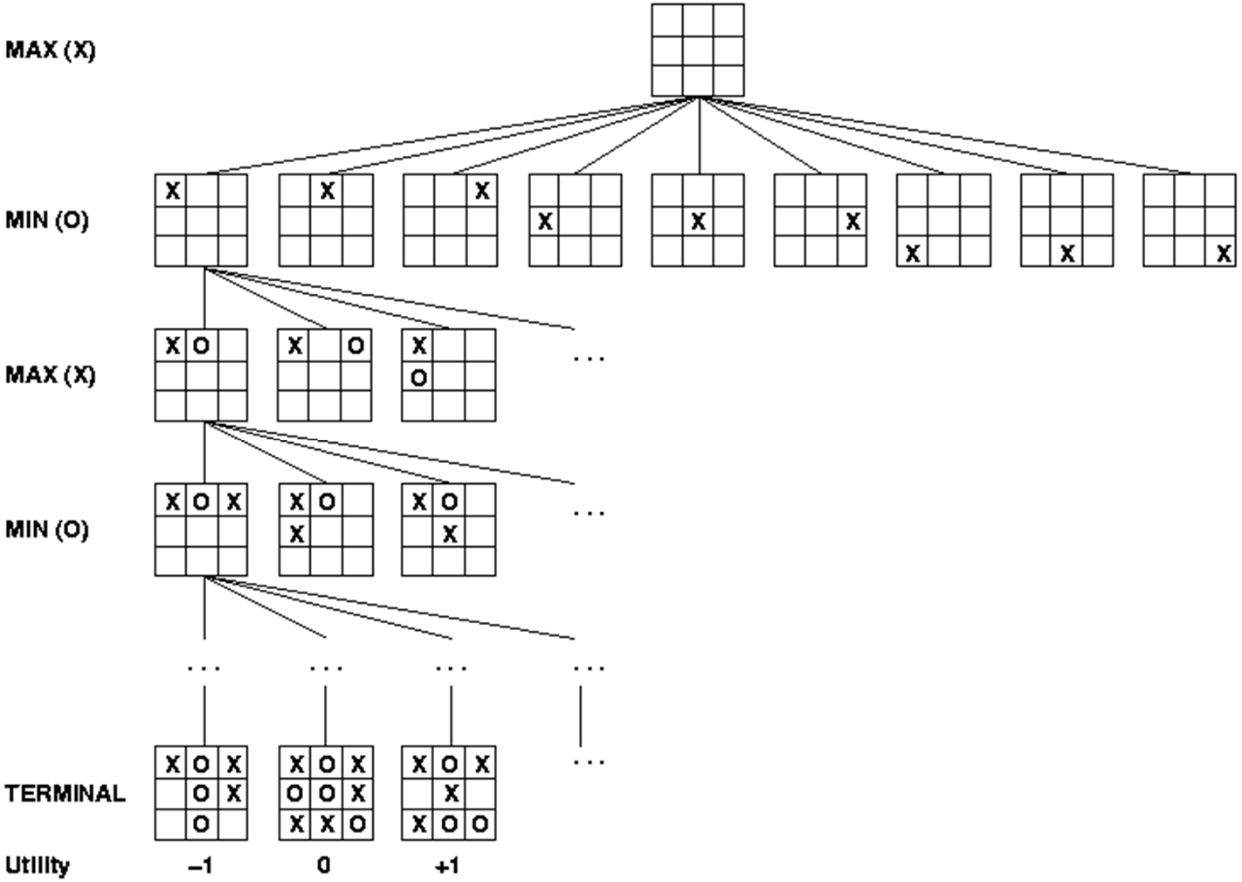
Wykonujący pierwszy ruch (ze stanu początkowego) nazywany jest **graczem**, a drugi **przeciwnikiem**. Określenie funkcji wypłaty dla stanów końcowych dokonywane jest z punktu widzenia gracza.

Korzeń jest początkowym stanem gry, jego potomkowie reprezentują stany osiągalne przez gracza w pierwszym ruchu. Z kolei ich potomkowie określają stany możliwe do uzyskania przez przeciwnika itd. Węzły końcowe (liście drzewa) reprezentują stany wygrane (wartości dodatnie), przegrane (wartości ujemne) lub remisowe (zero) dla gracza. Każda droga z węzła początkowego do liścia opisuje jedną rozgrywkę.

Wyznaczenie strategii gry (planu rozgrywki) polega na określeniu kolejnych posunięć gracza dla różnych posunięć przeciwnika.

Wyznaczenie strategii gry

Fragment drzewa dla gry „kółko i krzyżyk”:



Metody określenia strategii wygrywającej dla gracza:

- na podstawie macierzy gry
- na podstawie drzewa gry

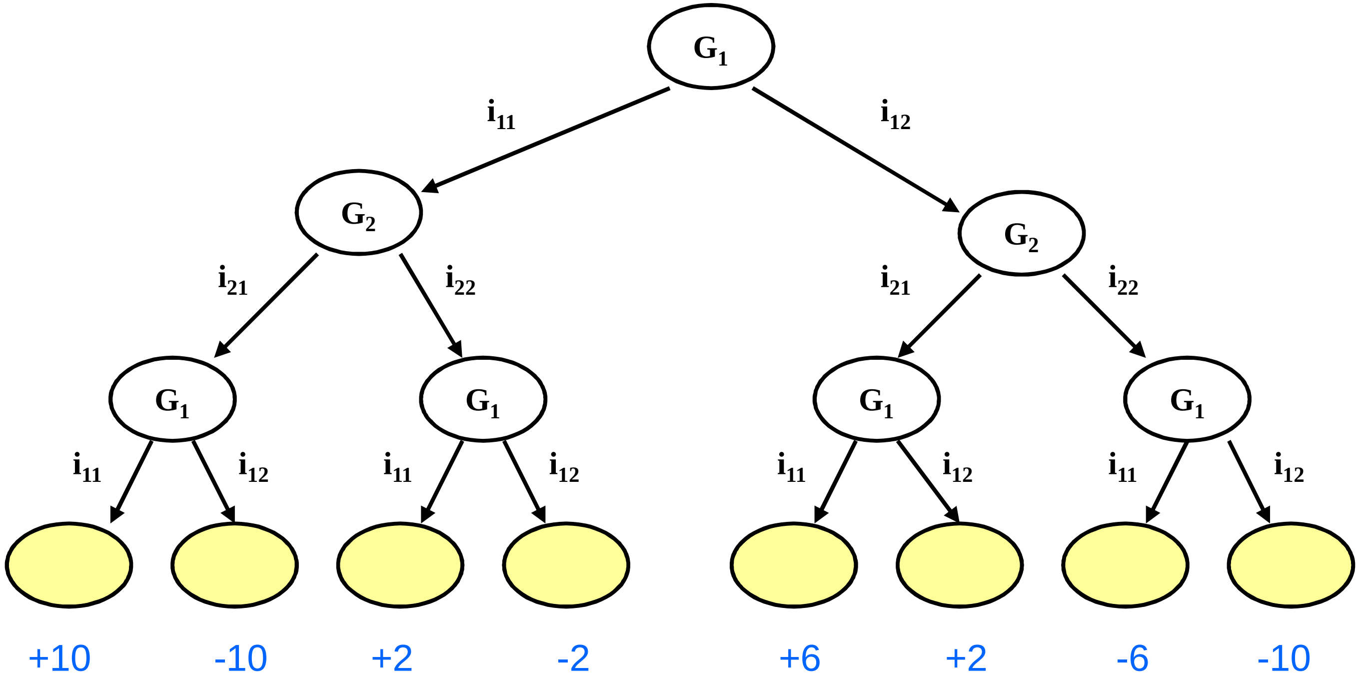
<u>Ogólny model gry *n*-osobowej:</u>

Niech **M** = $\{m_1, m_2, .. , m_k\}$ oznacza zbiór ruchów dopuszczalnych dla danej gry.

W zbiorze **M** można wyróżnić *n* podzbiorów $M_1, M_2, .. , M_n$. Zbiór M_i jest zbiorem dopuszczalnych ruchów dla gracza G_i $(i = 1, 2, .. , n)$.

W ogólnym przypadku podzbiory M_i nie muszą być rozłączne.

Przykładowe drzewo dla gry dwuosobowej:



Wyznaczenie strategii gry

Dla gracza G_1: $M_1 = \{i_{11}, i_{12}\}$

Dla gracza G_2: $M_2 = \{i_{21}, i_{22}\}$

Każda partia kończy się po trzech ruchach (gracz G_1 wykonuje dwa ruchy, G_2 tylko jeden). W każdym posunięciu gracz może wybrać jeden z dwóch możliwych ruchów.

Gracz G_1 posiada cztery różne strategie:

$S_{11} = i_{11}\, i_{11}$ $S_{12} = i_{11}\, i_{12}$ $S_{13} = i_{12}\, i_{11}$ $S_{14} = i_{12}\, i_{12}$

Gracz G_2 posiada dwie różne strategie:

$S_{21} = i_{21}$ $S_{22} = i_{22}$

Na tej podstawie możemy utworzyć macierz gry **S**:

	S_{21}	S_{22}
S_{11}	+10	+2
S_{12}	-10	-2
S_{13}	+6	-6
S_{14}	+2	-10

Gracz G_1 dąży do znalezienia takiej strategii, która zapewni mu maksymalną wygraną (wartości wypłat ze znakiem +).

Przeciwnik G_2 również dąży do znalezienia takiej strategii, która zapewni mu maksymalną wygraną (wartości wypłat ze znakiem -).

Strategia optymalna dla gracza G_1 zostanie osiągnięta gdy:

$$\alpha^* = \max_{i} \{ \min_{j} S_{ij} \},$$

gdzie S_{ij} – element macierzy S (i - indeks wiersza, j- indeks kolumny)

W tym celu dla każdego wiersza określamy wartość minimalną i wybieramy taki wiersz, w którym znalezione minimum jest największe:

	S_{21}	S_{22}	
S_{11}	+10	+2	+2
S_{12}	-10	-2	-10
S_{13}	+6	-6	-6
S_{14}	+2	-10	-10

$\alpha^* = S_{11}$

Wyznaczenie strategii gry

Strategia optymalna dla gracza G_2 zostanie osiągnięta gdy:

$$\beta^* = \min_{j} \{ \max_{i} S_{ij} \},$$

W tym celu dla każdej kolumny określamy wartość maksymalną i wybieramy taką kolumnę, w której znalezione maksimum jest najmniejsze:

	S_{21}	S_{22}
S_{11}	+10	+2
S_{12}	-10	-2
S_{13}	+6	-6
S_{14}	+2	-10
	+10	+2

$$\beta^* = S_{22}$$

Wyznaczenie strategii gry

Optymalna dla gracza G_1 strategia $\alpha^* = S_{11}$ nazywa się strategią maksyminową (lub maksyminimalną). Zapewnia ona graczowi G_1 wygraną +2 (dolna wygrana).

Optymalna dla gracza G_2 strategia $\beta^* = S_{22}$ nazywa się strategią minimaksową (lub minimaksymalną). Funkcja wypłaty wynosi w tym przypadku +2 (przegrana gracza G_2). Strategia ta określa górną przegraną gracza G_2.

Jeżeli zachodzi równość:

$$\alpha^* = \beta^*$$

to gra posiada punkt siodłowy, a uzyskana w tym przypadku funkcja wypłaty nazywa się wartością gry.

Wyznaczenie strategii gry

Gry posiadające punkt siodłowy nazywamy grami o pełnej informacji (lub grami zamkniętymi). W takich grach gracze dokonują analizy macierzy gry, następnie wybierają swoje strategie optymalne i na tym właściwie gra się kończy.

Z punktu widzenia teorii gier gry o pełnej informacji są mało ciekawe, gdyż ich rozwiązanie jest trywialne (każdy gracz stosuje tylko jedną strategię – tzw. strategia czysta).

W praktyce bardzo rzadko spotykamy się z grami o pełnej informacji. Najczęściej gry dwuosobowe nie posiadają punktu siodłowego (nazywamy je grami o niepełnej informacji lub grami otwartymi).

Wyznaczenie strategii gry

W wielu przypadkach reprezentacja macierzowa jest nieprzydatna przy wyznaczaniu strategii optymalnych np. gdy nie można zbudować pełnego drzewa gry lub drzewo jest znacznych rozmiarów.

Optymalną strategię można wówczas znaleźć poprzez zastosowanie rekurencyjnych funkcji wartościujących wierzchołki drzewa gry.

W drzewie gry wyróżniamy dwa typy wierzchołków:

- na poziomie gracza
- na poziomie przeciwnika.

Na początku funkcje wartościujące przypisane są jedynie wierzchołkom końcowym (liściom drzewa).

Wyznaczenie strategii gry

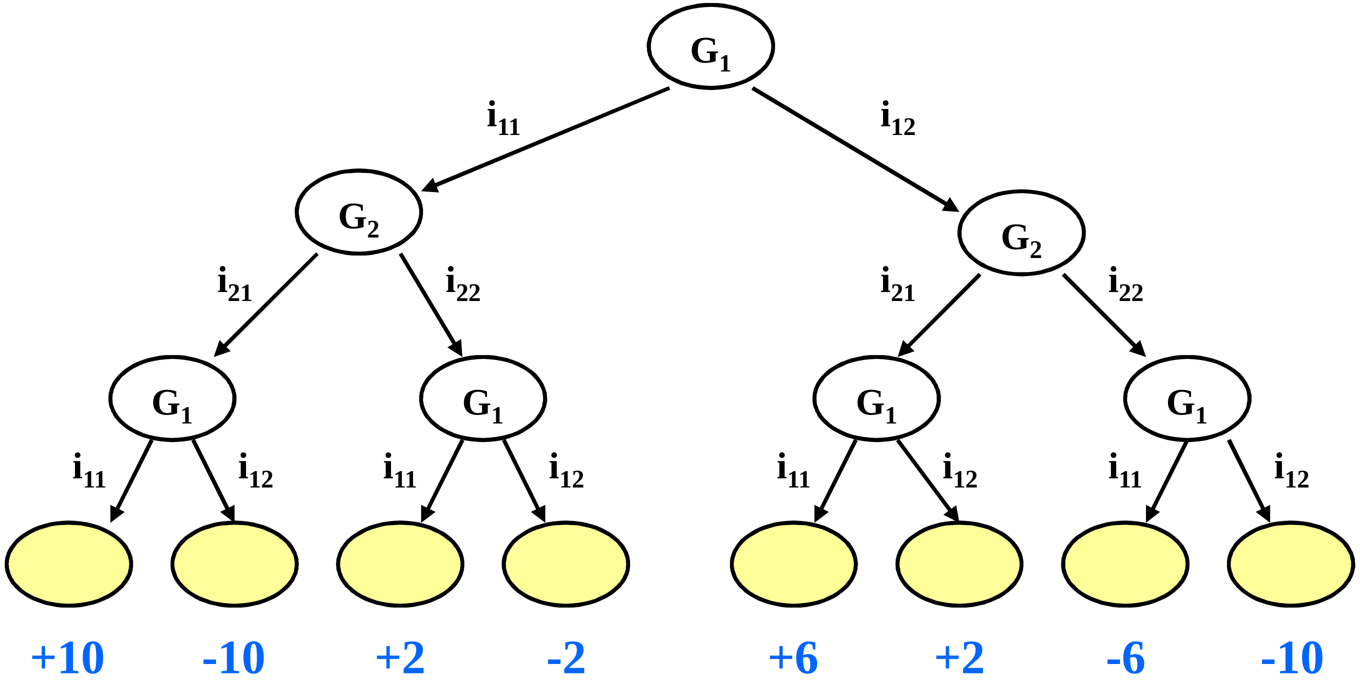
Wierzchołkom na poszczególnych poziomach drzewa gry (od liści do korzenia) przypisywane są kolejne wartości zgodnie z następującym algorytmem:

- wierzchołkom na poziomie gracza nadawana jest maksymalna wartość potomków
- wierzchołkom na poziomie przeciwnika nadawana jest minimalna wartość potomków

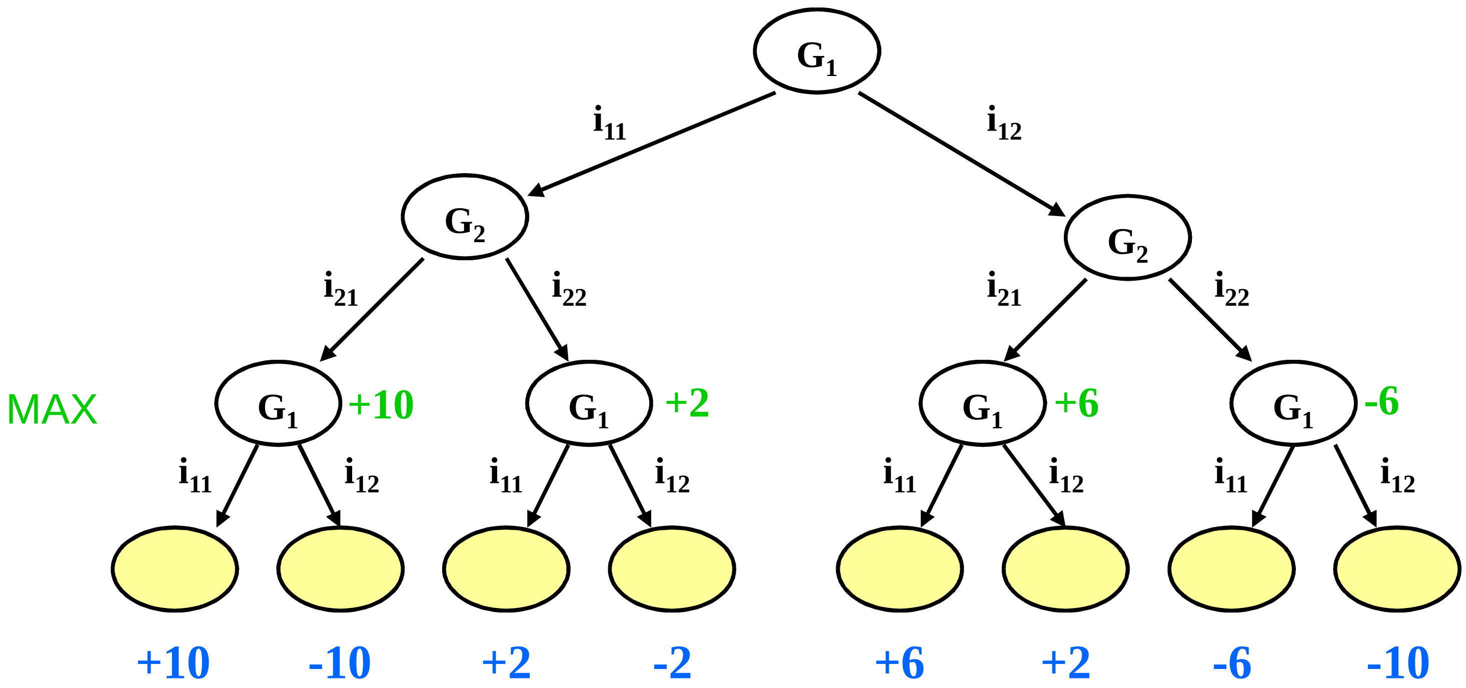
Powyższy algorytm zakłada, że każdy gracz w kolejnych krokach dąży do wybrania jak najlepszego ruchu.

Działanie algorytmu ilustruje następujące drzewo gry:

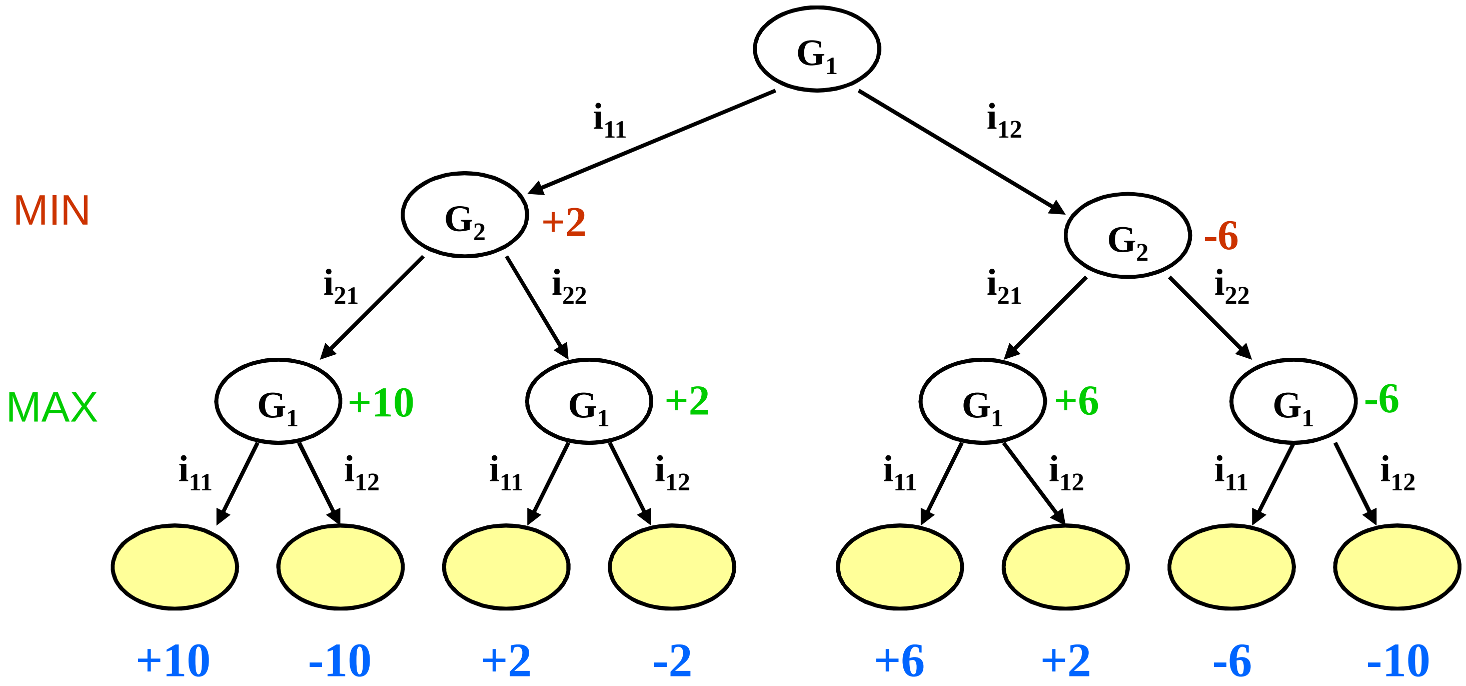
etap 1 – budowa drzewa gry i ocena wierzchołków końcowych



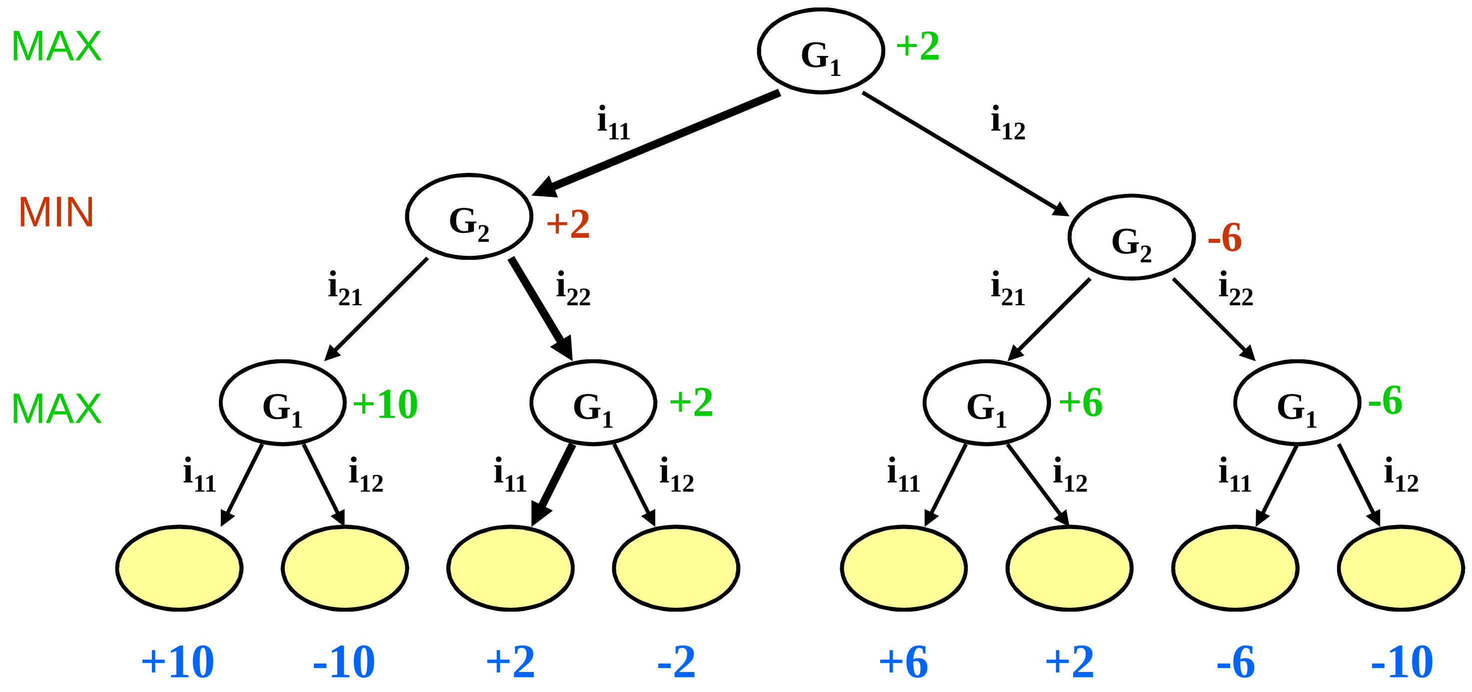
etap 2 – ocena wierzchołków na poziomie gracza (maksymalna wartość potomka)



etap 3 – ocena wierzchołków na poziomie przeciwnika (minimalna wartość potomka)



Ostateczne rozwiązanie:



Twierdzenie minimaksowe

Fundamentalnym twierdzeniem teorii gier, dotyczącym gier dwuosobowych o sumie zerowej, jest twierdzenie minimaksowe udowodnione przez von Neumanna (1928):

Każda skończona gra dwuosobowa o sumie zerowej ma co najmniej jedno rozwiązanie, które określa wartość gry i optymalne strategie graczy.

Strategia minimaksu

- utwórz drzewo gry o maksymalnej głębokości
- przypisz wierzchołkom końcowym drzewa (liściom) wartości zgodnie z przyjętą funkcją oceny np. +1, -1, 0 (w przypadku dużych drzew użyj funkcji heurystycznej)
- wierzchołkom na kolejnych poziomach (od liści do korzenia) nadaj wartości zgodnie z zasadą: maksymalna wartość potomka dla gracza oraz minimalna wartość potomka dla przeciwnika
- po osiągnięciu korzenia wybierz strategię, która zapewnia największe zyski

Algorytm cięć α-β

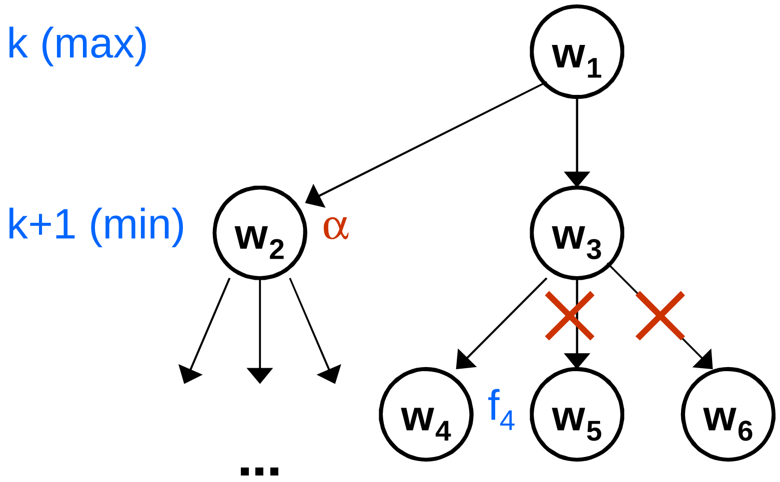
Algorytm cięć α-β stanowi modyfikację algorytmu mini-maksowego. W algorytmie tym węzły nie wpływające na wartość przypisywaną ich przodkom są eliminowane z dalszej analizy (znaczne ograniczenie drzewa gry).

W algorytmie występują dwa rodzaje cięć:

- **cięcie** α - dla wierzchołków na poziomie gracza ocenę wierzchołków potomnych można zakończyć, gdy wartość będzie niższa niż dotychczasowe maksimum α
- **cięcie** β - dla wierzchołków na poziomie przeciwnika ocenę wierzchołków potomnych można zakończyć, gdy wartość będzie wyższa niż dotychczasowe minimum β

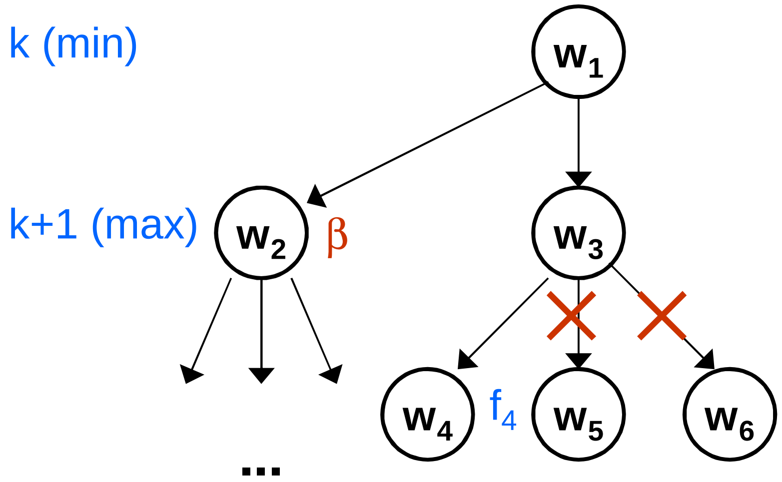
Wyznaczenie strategii gry

Cięcie α:



jeżeli $f_4 <= \alpha$, to wierzchołków w_5 i w_6 nie musimy rozpatrywać

Cięcie β:



jeżeli $f_4 >= \beta$, to wierzchołków w_5 i w_6 nie musimy rozpatrywać

Zastosowanie heurystyki do oceny wierzchołków

Pomimo pełnej informacji o przebiegu gry i skończonej liczbie możliwych ruchów jakie mogą wykonać gracze w danej sytuacji utworzenie pełnego drzewa gry i wyznaczenie na tej podstawie strategii optymalnych dla graczy jest w wielu przypadkach niemożliwe.

Szacuje się, że w szachach możliwych jest około 10^{120} różnych gier.

Dla gry „kółko i krzyżyk” na 9 polowej planszy liczba różnych partii wynosi 9! = 362880.

W takich przypadkach stosowane jest podejście heurystyczne do wyznaczenia optymalnej strategii gry.

Aby określić najlepszy ruch w danej sytuacji każdemu możliwemu wyborowi przyporządkowuje się pewne oszacowanie liczbowe.

W dowolnym posunięciu należy wybrać tę strategię ruchu, której odpowiada największa (lub najmniejsza) wartość oszacowania liczbowego.

Należy dążyć do tego, aby optymalne strategie poszczególnych ruchów prowadziły do optymalnej strategii pełnej.

Przykład – „kółko i krzyżyk”

Oznaczmy pola planszy zgodnie z rysunkiem:

1	2	3
8	9	4
7	6	5

Niech o_i (i = 1, 2, ... , 9) oznacza funkcję oceny kratki z numerem i. Na początku gry przyjmuje się:

- $o_9 = 3$ (środek planszy)
- $o_1 = o_3 = o_5 = o_7 = 2$ (narożniki planszy)
- $o_2 = o_4 = o_6 = o_8 = 1$ (pozostałe pola)

Wyznaczenie strategii gry

Z każdym polem związany jest zestaw kierunków:

- pole środkowe (9) posiada 4 możliwe kierunki – poziomy, pionowy i dwa diagonalne (po przekątnej)
- pola narożne (1, 3, 5, 7) posiadają 3 możliwe kierunki – poziomy, pionowy i jeden diagonalny
- pozostałe pola (2, 4, 6, 8) posiadają 2 możliwe kierunki – poziomy i pionowy.

Kierunek poziomy oznaczamy przez *h*, pionowy przez *v*, natomiast diagonalny przez *d*.

Dla dowolnym kierunku może wytworzyć się jedna z 6 typowych sytuacji (każdej przypisano ocenę liczbową):

x	x		40
o	o		20
x			10
o			5
			1
x	o		0

Wyznaczenie strategii gry

Mając tablicę ocen typowych sytuacji na poszczególnych kierunkach można wyznaczyć oceny liczbowe dla każdego pola w dowolnym momencie gry. Ocenę o_i wykonania ruchu na pole i obliczmy jako sumę ocen sytuacji na wszystkich kierunkach związanych z danym polem.

Należy określić najlepszy ruch dla gracza G_1 (stawiającego krzyżyk):

1	2	3 X
8	9 X	4
7 O	6 O	5 X

$o_1 = h + v + d = 10 + 5 + 10 = 25$
$o_2 = h + v \ = 10 + 0 = 10$
$o_4 = h + v = 10 + 10 = 20$
$o_5 = h + v + d = 20 + 10 + 10 = 40$
$o_8 = h + v = 10 + 5 = 15$

Jako podsumowanie:

Mistrz świata	– 2800 punktów
5 ruchów „do przodu”	– 1500 punktów
dodatkowy poziom (od 5 do 10)	– + 200 punktów

Deep Blue analizował 200-1000 milionów pozycji w ciągu sekundy.