

IMPLEMENTACIÓN MONTECARLO

Mientras Recursos

$n \leftarrow \text{Seleccionar}(T)$

def Seleccionar(T)

si terminal(T) or TieneHijo SiExpandir(T)

return T

else

return Seleccionar(hijo con mayor UCT(nt))

será un nodo interior del árbol y todos sus hijos estarán expand.

Necesitamos calcular sobre: (+)

→ hijos sin expandir [] → (cuando este este vacío)

→ Diccionario con [m] = hijo

→ U de visitas del nodo (nv)

→ Vector Victorias [Empates, J1, J2]

→ padre

Desde aquí solo el hijo con mayor UCT ya que puedo calcular el UCT de un nodo.

UCT(nt) = $\frac{Q(nt)}{N(nt)} + 2cp \sqrt{\frac{\lg(N(nt.padre))}{N(nt)}}$

U de veces q es visitado

U de de nuevo es ese nodo

U de padre, nv

U de 1/2

Nodo n ⇒ nodo que quiero expandir

Mientras Recursos

$n \leftarrow \text{Seleccionar}(T)$

$nt \leftarrow \text{expandir}(n)$

valor ← Simular(a1)

gana j1 = 2

gana j2 = 1

empate = 0

def expandir(nt)

Si no tiene hijos

return nt

else

acción ← elección Aleatoria (hijos sin expandir)

// Realizamos la acción q es crear un nuevo nodo

new ← realiza la acción (acción) en nt.

eliminar acción hijos sin expandir

// Añadir en el diccionario la nueva acción

dict[acción] = new

return new

→ política para elegir hijos del vector: random

(+)

→ Tablero (info de a quien le toca jugar, etc).

Tablero

va siendo aleatorio hasta que:

→ gana j1

→ gana j2

→ empate

TABLAS ⇒ Bucle (?)

Si se llega a ese máximo ⇒ se empató.

⇒ Alg número para estimar la longitud.

def Simular(a1)

// Depende de nuestro tablero y nuestros valores.

Mientras Recursos

$n \leftarrow \text{Seleccionar}(T)$

$nt \leftarrow \text{expandir}(n)$

valor ← Simular(a1)

nt ← Retropropagación(ut, valor)

Actualizar Recursos → por tiempo

// Se sale del bucle (x iteraciones en x seg)

def Retropropagación(ut, valor)

nt.victoria[valor] += 1

nt.nv += 1

aux ← nt.padre

Mientras aux <> none

aux.nv += 1

aux.victoria[valor] += 1

aux ← aux.padre

return nt



logramos así una estimación de los hijos

Ahora para seleccionar la jugada usaremos una heurística (Greedy (el de + victorias), el de UCT ↑ ...)