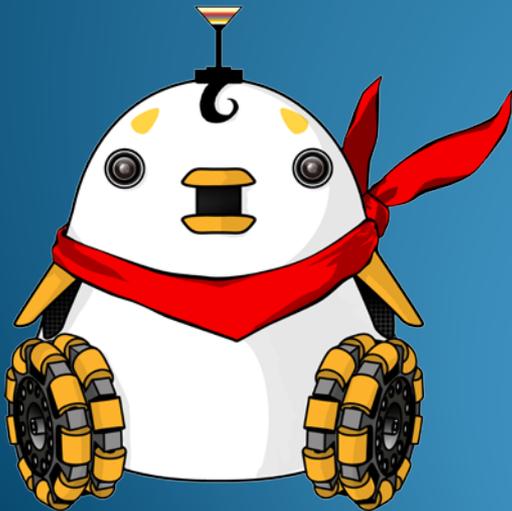


# 人工知能概論

第4回 探索(3)  
ゲームの理論



# Information



- このスライドは「イラストで学ぶ人工知能概論」を講義で活用したり，勉強会で利用したりするために提供されているスライドです。



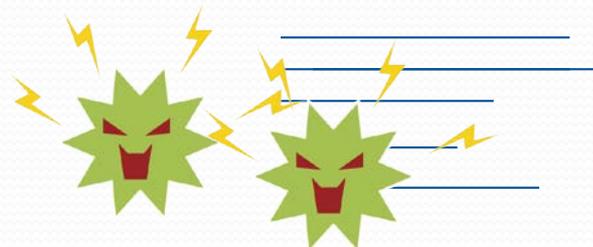
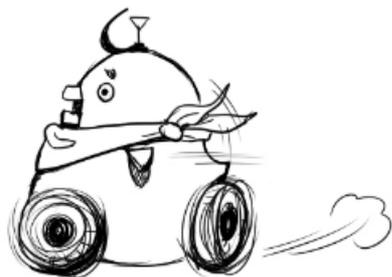
# STORY 探索(3)

- ホイールダック2号は一つ誤解をしていた。迷路の中ではとにかくまっすぐゴールに向かえばよいわけではない。迷路にはホイールダック2号を邪魔しようとする敵がいる。これとぶつかる则何かと面倒である。敵がどのように行動するのかを先読みしながら迷路を抜けなければならない。



# 仮定 探索(3)

- ホイールダック2号は自分と敵の行動に対する利得を知っている。
- ホイールダック2号は自らの行動に対する結果を確実に予測できるものとする。
- 敵は**合理的**(rational, p.46)に行動する。
- ホイールダック2号も敵も物理的につながっている場所・状態には意図すれば確定的に移動することができるものとする。



# Contents

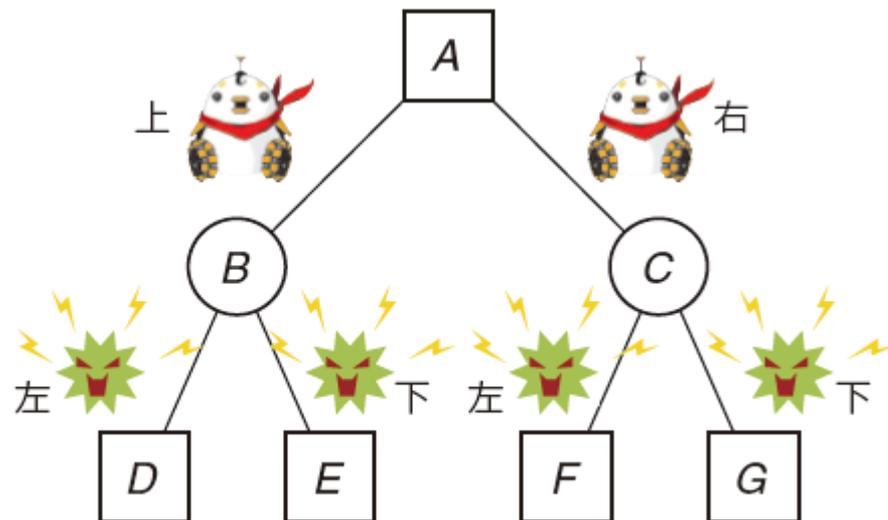
- 4.1 利得と回避行動
- 4.2 標準型ゲーム
- 4.3 展開型ゲーム

# 4.1.1 はじめに

- ホイールダック2号は  
どうすべきか？



図 4.2 敵をかわしてゴールに向かうタスク



ホイールダック2号	$C_W$	0	$D_W$	$D_W + C_W$
敵	$C_E$	$D_E$	0	$D_E + C_E$

図 4.3 図 4.2 のゲームの展開

表 4.1 図 4.2 のゲームの利得行列

	敵は左へ	敵は下へ
ホイールダック 2 号が上へ	$C_W, C_E$	0, $D_E$
ホイールダック 2 号が右へ	$D_W, 0$	$D_W + C_W, D_E + C_E$

# 利得行列

- プレイヤーが二人の場合には、各プレイヤーの行動を行列の行と列に書き、それぞれの交わるセルに、それぞれのプレイヤーが得る利得を書いたものを利得行列と呼ぶ。
- 一般の行列とは異なり、双行列(bimatrix)と呼ばれる。

表 4.1

図 4.2 のゲームの利得行列

プレイヤー2  
の行動

	敵は左へ	敵は下へ
ホイールダック 2 号が上へ	$C_W, C_E$	$0, D_E$
ホイールダック 2 号が右へ	$D_W, 0$	$D_W + C_W, D_E + C_E$

プレイヤー1  
の行動

左がプレイヤー1, 右がプレイヤー2の利得

## 4.1.2 ケース1:

### 敵はホイールダック2号を捕まえない

- ホイールダック<sub>2</sub>号にとっては上に移動すると敵にぶつかってしまうが、まだ利得が-5で済むため、上への移動を選ぶのが最適な選択となる。

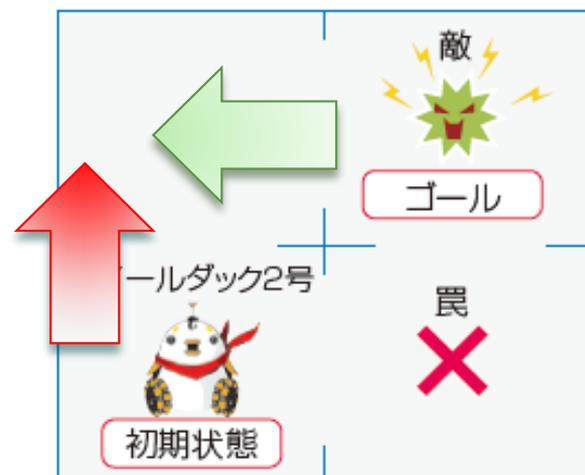


図 4.2 敵をかわしてゴールに向かうタスク

$C_W = -5, C_E = 3, D_W = D_E = -2$	敵は左へ	敵は下へ
ホイールダック 2 号が上へ	-5, 3	0, -2
ホイールダック 2 号が右へ	-2, 0	-7, 1

### 4.1.3 ケース2:

## 少しだけ敵のモチベーションが下がったら？

- ホイールダック2号は右に行くことで×印で-3のダメージを受けるだけで、ゴールにたどり着けるのである。
- 少しの利得の違いでとるべき行動が変化する。



図 4.2 敵をかわしてゴールに向かうタスク

$C_W = -5, C_E = 3-1, D_W = D_E = -2-1$	敵は左へ	敵は下へ
ホイールダック 2 号が上へ	-5, 2	0, -3
ホイールダック 2 号が右へ	-3, 0	-8, -1

主体の意思決定が混ざり合っって状況が決定する系を**ゲーム**と呼ぶ

# Contents

- 4.1 利得と回避行動
- 4.2 標準型ゲーム
- 4.3 展開型ゲーム

# 4.2.1 はじめに/ 標準型ゲーム

- ゲーム理論

複数のプレイヤーの意志決定を扱う理論

1944年ジョン・フォン・ノイマン, オスカー・モルゲンシュテルン  
「ゲームの理論と経済行動」

- 基本的な用語の定義

- **プレイヤー** 意志決定を行なう個々の主体. 複数存在する.
- **行動** プレイヤーの選択. 戦略と呼ぶこともある. 探索の作用素に相当.
- **利得** プレイヤーの行動の組み合わせに対して定義される数値. 結果に対する各プレイヤーの効用を示す. 大きいほうがより嬉しいとする.
- **合理的** 各プレイヤーは自分の利益を最大化しようと最大限の努力をする. 利己的(selfish)と呼ぶこともあるが, ニュアンス的に誤解がある. 英語ではrational.
- **均衡** 合理的な意思決定の結果として, 自ずと決まる全プレイヤーの行動の落ち着く先.

# 4.2.4 支配戦略均衡

- 相手の行動が何であろうが、その行動をとった方が高い利得を得られる行動を支配戦略という。
- ゲームの状態は支配戦略均衡に至る(前提:支配戦略の存在+合理性)

表 4.4 支配戦略均衡を持つ利得行列

	敵は エネルギー供給装置へ	敵は 普通に休む
ホイールダック 2 号は エネルギー供給装置へ	4, 4	8, 1
ホイールダック 2 号は 普通に休む	1, 8	1, 1

エネルギー供給装置



# 4.2.5 ナッシュ均衡

- ナッシュ均衡：行動の組(ホイールダック<sub>2</sub>号の行動, 敵の行動) が互いに相手の行動に対する最適な反応となっている場合

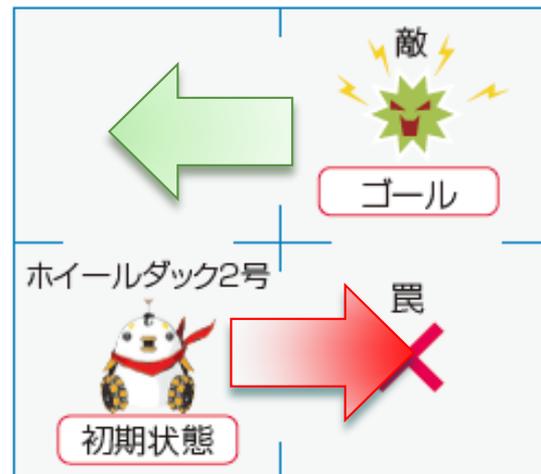


図 4.2 敵をかわしてゴールに向かうタスク

	敵は左へ	敵は下へ
ホイールダック 2 号が上へ	-5, 2	0, -3
ホイールダック 2 号が右へ	-3, 0	-8, -1



## 4.2.6 囚人のジレンマ

- ナッシュ均衡は必ずしも全体として良い状態に至るわけではない。

「白状したらおまえだけは助けてやるぞ！」

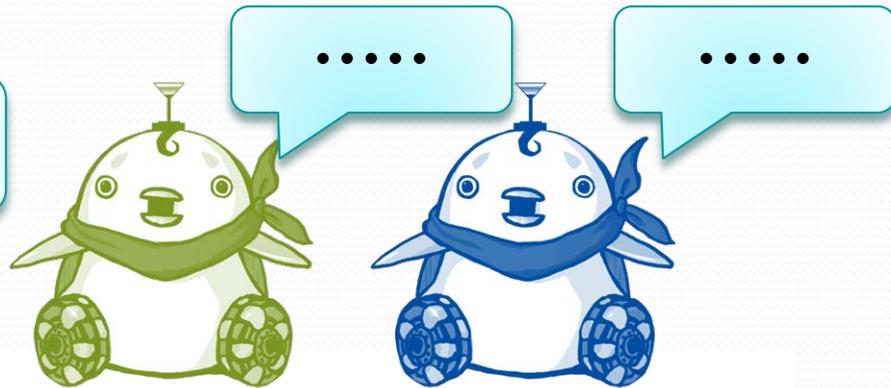


表 4.5

囚人のジレンマの利得行列

囚人 1 \ 囚人 2	自白	黙秘
自白	-3, -3	5, -5
黙秘	-5, 5	3, 3

## 4.2.7 ゼロサム・ゲーム



- ゼロサム・ゲームはプレイヤーの利得の総和が0になるゲームであり、特にプレイヤーが2名の場合はプレイヤー<sub>1</sub>の利得を $r$ とすると、プレイヤー<sub>2</sub>の利得は $-r$ となることになる。

**表 4.6** 賭けジャンケンの利得行列：ゼロサム・ゲームの例

	プレイヤー 2：グー	プレイヤー 2：チョキ	プレイヤー 2：パー
プレイヤー 1：グー	0	100	-100
プレイヤー 1：チョキ	-100	0	100
プレイヤー 1：パー	100	-100	0

双行列で書く必要がない

## 4.2.8 ミニマックス戦略

- 相手がプレイヤーの利得を最小化してくることを前提としながら、自らの利得を最大化する戦略
- ナッシュ均衡を実現する。

表 4.7 ゼロサム・ゲームでの利得行列とミニマックス戦略

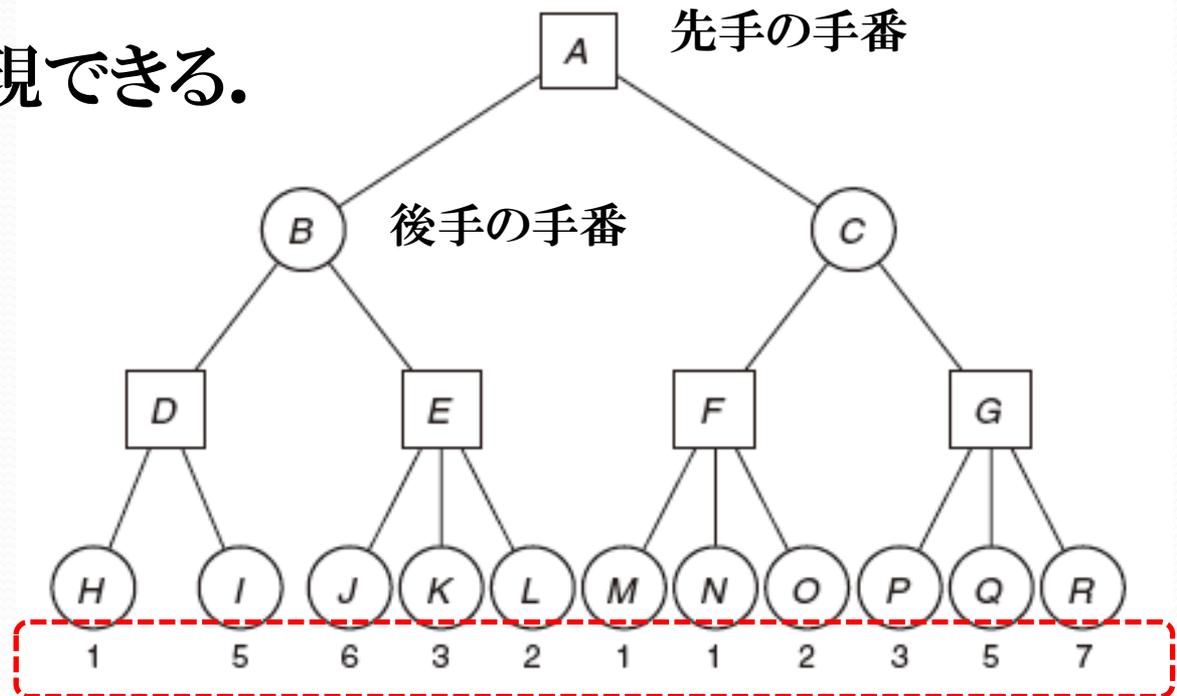
ホイールダック 2号 \ 敵	$b_1$	$b_2$	$b_3$	$b_4$
$a_1$	7	2	5	1
$a_2$	2	2	3	4
$a_3$	5	3	4	4
$a_4$	5	2	1	6

# Contents

- 4.1 利得と回避行動
- 4.2 標準型ゲーム
- 4.3 展開型ゲーム

# 4.3.1 展開型ゲーム

- 実際のゲームは一度きりの意思決定ではなく、多段階の意思決定を含む。このようなゲームを展開型ゲームという。
- オセロやチェスなど多くのゲームは展開型ゲームでモデル化できる。
- ゲーム木で表現できる。



最終的に先手が得る利得

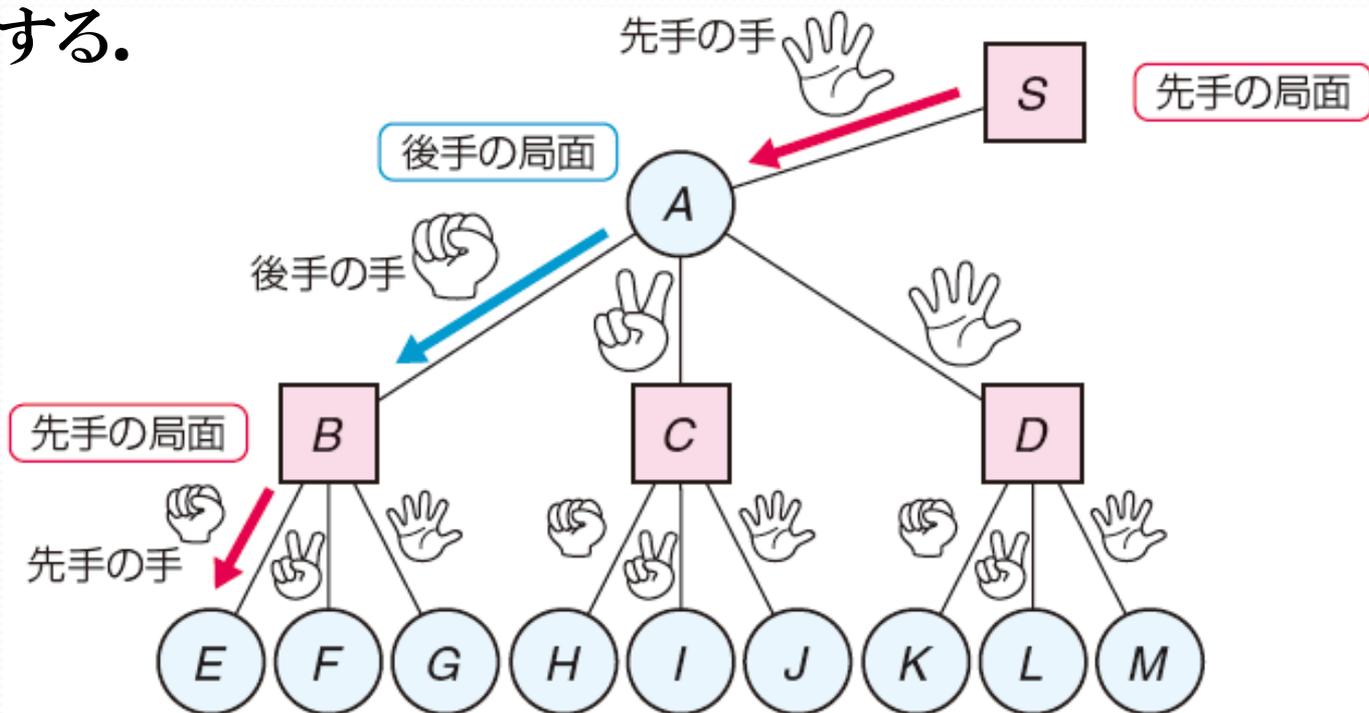
図 4.4

ゲーム木の例

# 演習問題4-1 交互ジャンケン

1. 順番にジャンケンを出すゲームをする。相手に勝つ手を出すと、自分にその指の本数分加点される。(負けた場合、あいこの場合得点の変化はない。)

自分がまず初期状態からパーを出した状態からスタートし、初期状態→相手→自分の一往復で終了する際のゲーム木は以下のようになる。ゲーム木の葉ノードに評価値を記入せよ。ただし評価値は評価値 = 自分の得点 - 相手の得点とする。



# 4.3.3 ミニマックス法

- 先手は後手が「最も先手が低い利得になる」ような手をとるであろうことを前提にして、自らが高い利得を得られる行動をとる。

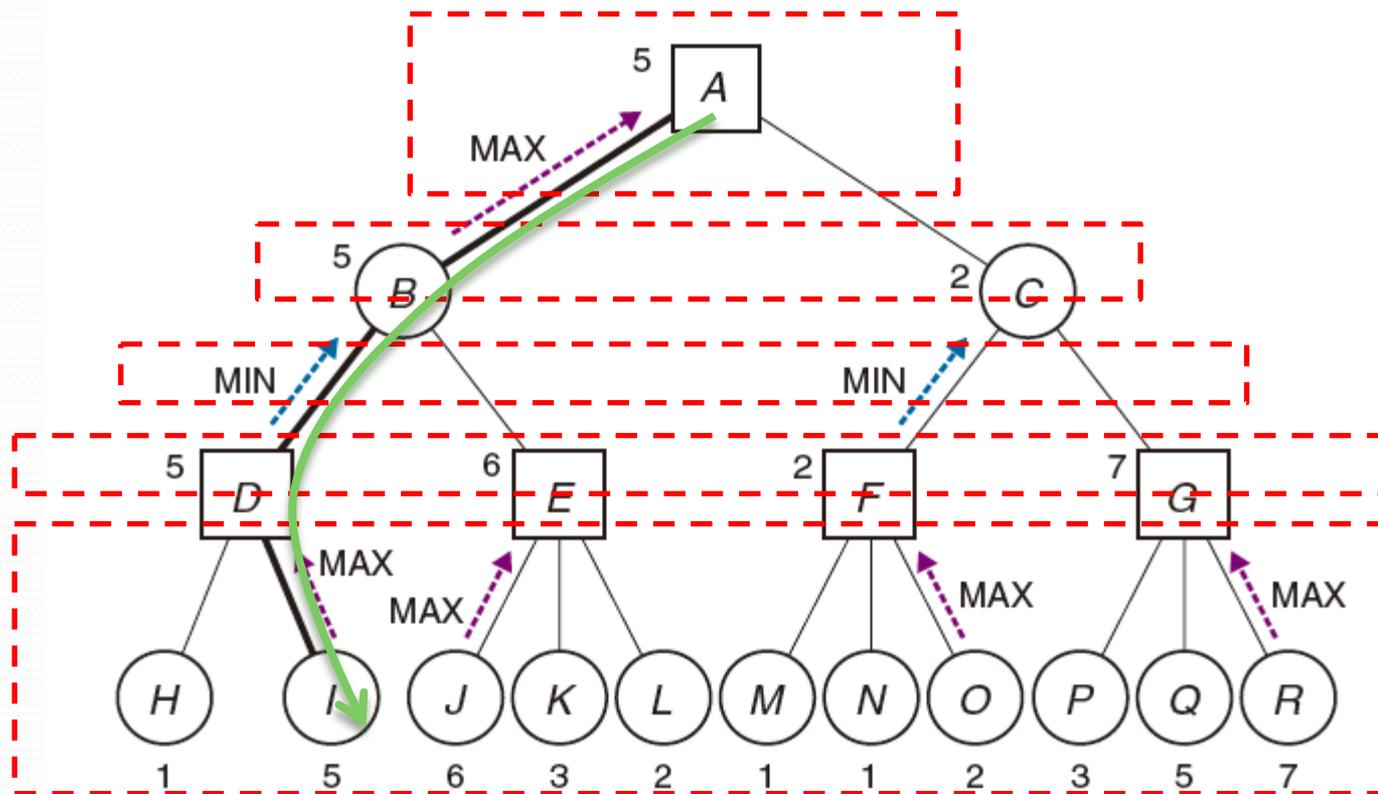


図 4.5 ミニマックス法の解探索例



# アルファ・ベータ法

- ミニマックス法では盤面の局面を先読みすればするほど、良い手を選択し、ゲームを有利に進めていける。
- しかし、探索しすぎるとゲーム木の探索空間が膨大になる。
- ミニマックス法の性質を生かして、不必要な探索を避ける(サボる)ことができる。

pruning = 枝刈り

- アルファ・ベータ法 ( $\alpha\beta$  pruning)
  - $\beta$ カット( $\beta$  pruning) 評価値最小化局面の枝刈り
    - 後手がわざわざ評価値の大きな手を打たないことを利用して、**先手の行動(作用)をカット**
  - $\alpha$ カット( $\alpha$  pruning) 評価値最大化局面の枝刈り
    - 先手がわざわざ評価値の小さな手を打たないことを利用して、**後手の行動(作用)をカット**

# βカット (β pruning)

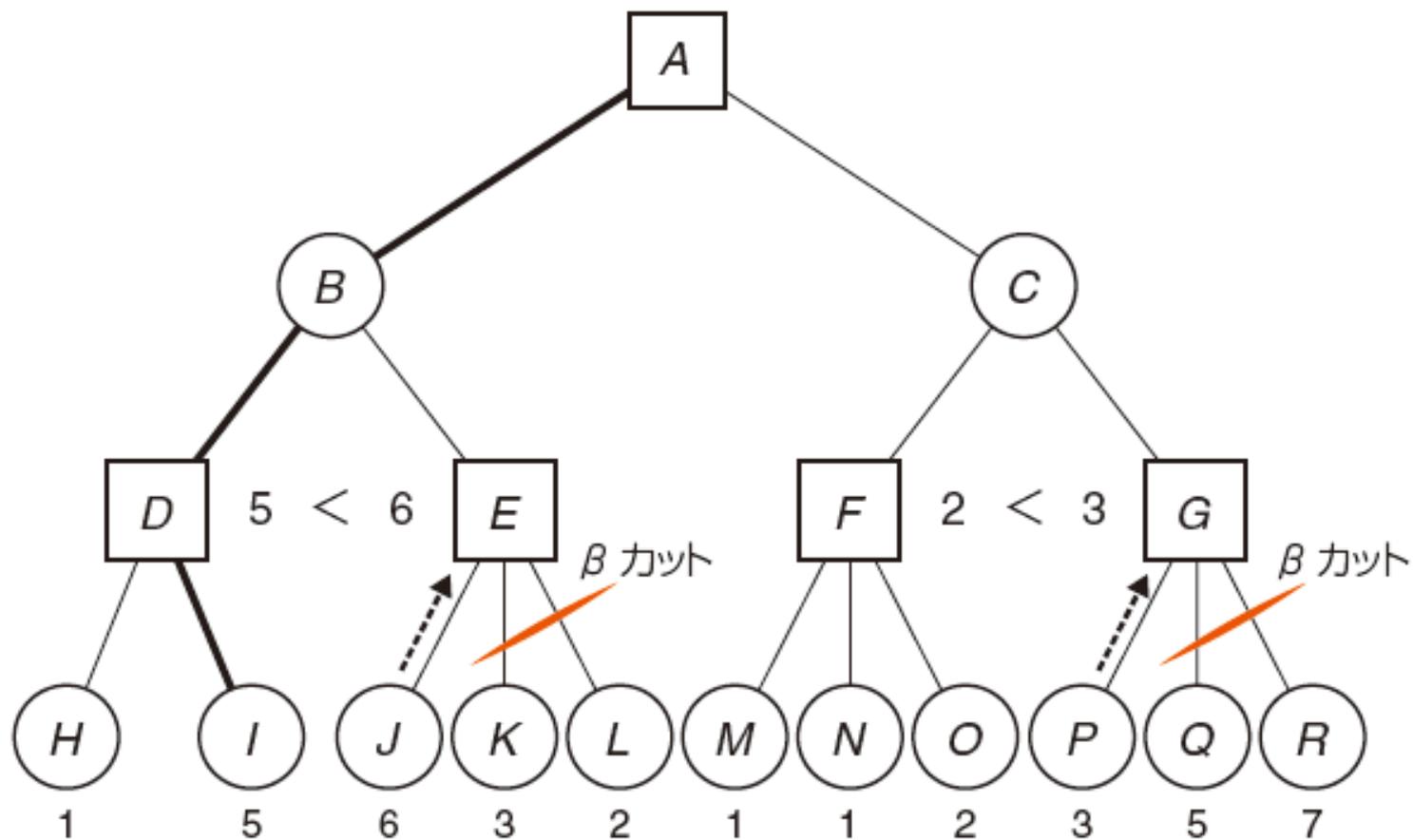


図 4.6

$\beta$  カットの例

# $\alpha$ カット ( $\alpha$ pruning)

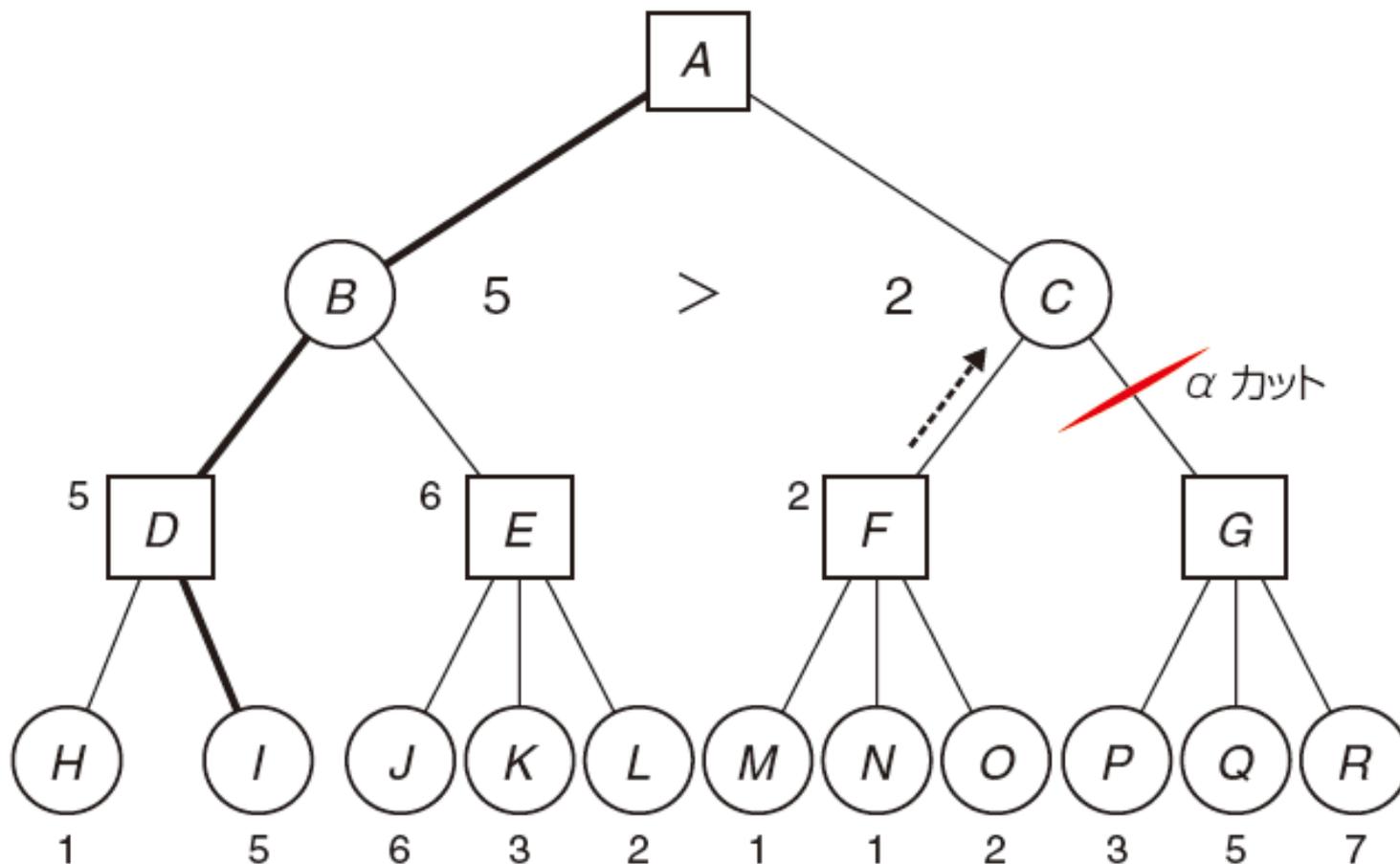


図 4.7  $\alpha$  カットの例



# 演習4-4 現実のゲーム

- オセロはゲーム木で表現される種類のゲームである。
- オセロのゲーム木を表現し、勝利のための必勝法を計算したい。
- 初手黒が置き、その後、交互に置いていく事を考えると、オセロの状態空間の大きさ(葉ノードの数)はどれくらいになるか？概算を求めよ。
- 盤面は  $8 \times 8$  である。
- 近似として一回に置ける手は平均して5箇所程度であるとしてよい。



# 第4章のまとめ

- プレイヤー, 利得行列や合理的な行動といったゲーム理論における基本用語を学び, ゲーム理論の対象となるゲームとは何かを知った.
- 支配戦略均衡やナッシュ均衡といった標準型ゲームにおける均衡概念の基礎について学んだ.
- 展開型ゲームとそのゲーム木による表現について学んだ.
- 展開型ゲームがゼロサム・ゲームであった場合について効率的に解を求めるミニマックス法を学んだ.
- ミニマックス法において解の探索を効率化する $\alpha$  カットと $\beta$  カットについて学んだ.

# 宿題

1. 章末問題の1を答えよ(演習4-1~4-3を含んでいる)。

答えは教科書の巻末に与えられているので、宿題としては特に(5)が答えのようになることの「理由説明」を中心とする

2. 予習問題：第5章は「動的計画法」が重要なポイント

図5.5の初期値からアルゴリズム5.1によって、どのように図5.10という結果が得られるか、考えてみよ(次回講義の一つのポイント)