

第7回 マルチエージェント

7th week : Multi-Agents

延原 肇 (知能機能システム専攻)

Hajime Nobuhara @ iit



授業構成

なぜマルチエージェント？

[Why multi-agents?](#)



エージェントとは？

[What is the concept of agents?](#)

マルチエージェントとは？

[What is the concept of multi-agents?](#)

マルチエージェントと強化学習

[Multi-agents and reinforcement learning](#)

マルチエージェントとゲーム理論

[Multi-agents and game theory](#)

最終課題についての説明

[About the final report](#)

2023年度スケジュール

[4/17] 第1回 ガイダンス (Guidance)

[4/24] 第2回 フラクタル (Fractal)

[5/08] 第3回 カオス (Chaos)

[5/15] 第4回 セルオートマトン (Cellular Automata)

[5/22] 第5回 パーコレーション (Percolation)

[5/29] 第6回 複雑ネットワーク (Complex Networks)

[6/05] 第7回 マルチエージェント (Multi-agents)

[6/12] 最終課題のため自習期間

[6/19] 最終課題のため自習期間

[6/26] 最終課題提出

授業構成

なぜマルチエージェント？

[Why multi-agents?](#)



なぜマルチエージェント？

Why multi-agents?

- やってみないとわからない（と思われる）
- 解析的に解くことが難しい（と思われる）ような問題に有用
The multi-agents are helpful for the problem that cannot be solved analytically.

ある宴会会場におけるテーブルの配置の問題

Table and food assignment problem in the Banquet.



それぞれの人がそれぞれの価値観・思惑で行動

Each participants do own behavior based on own evaluations and intentions.

あのエビチリを狙おう
Let's eat Ebichiri

あのパテがなくなりそうだ。
早く確保しないと・・・
Patty will be exhausted to eat,
so I need to keep some of them.

あのサラダが美味しいそう
That salad seems to be delicious



あそこで、カレーをよそって
いる人、早くどいてくれない
かな
I hope the person who is serving
own food will leave soon.

なぜマルチエージェント？

Why multi-agents?

この状態を解析的に解けるか？

Can you solve the following situation, I mean, can you optimize the comfortable degree of participants ?



なぜマルチエージェント？

Why multi-agents?

日本経済新聞 2008年10月28日

トイレ不足 深刻

首都直下地震で「トイレ不足」が深刻化する地域

都心7割の人使はず

防災会議調査

2008年10月28日

日本経済新聞 2008年10月28日

マグニチュード 7.3 直下地震発生

Magnitude 7.3 Earthquake occurrence

23区から 1,300万人が徒歩で帰宅

13 million people from Tokyo return home on foot

学校は子供の安全を考慮し開放されない

School is not open to consideration
of the safety of students.

日本経済新聞 2008年10月28日

首都直下地震で「トイレ不足」が深刻化する地域

都心7割の人使はず

防災会議調査

2008年10月28日

日本経済新聞 2008年10月28日

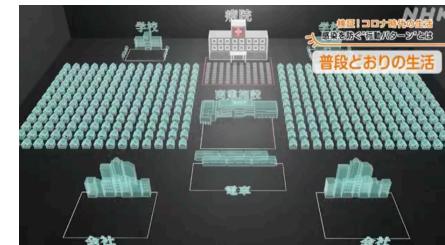
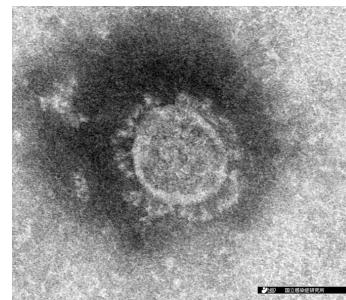
なぜマルチエージェント？

Why multi-agents?

2011年3月11日の東日本大震災に関する報告書（国土交通省）

Ministry of Land, Infrastructure and Transportation and Tourism report
about 2011 Tohoku earthquake

事例No	144	分類	緊急対応期	類型	帰宅困難者対策
事象等	渋滞・歩行帰宅者への対応				
実施者	東京都				
概要	首都圏では、震災発災後、公共交通機関が機能を失い、多数の帰宅困難者が発生した。国土交通省では、運転再開が困難な鉄道に代わる代替輸送をバス会社に要請した。しかし、一部の帰宅困難者は、自家用車による帰宅や歩行による帰宅を求め、都内の道路には、歩行者があふれ渋滞も発生した。				
○帰宅困難者に水やトイレをコンビニなどに支援要請	<p>帰宅困難者への支援として、首都圏の九都県市（埼玉県、千葉県、東京都、神奈川県、横浜市、川崎市、千葉市、さいたま市及び相模原市）は、震災当日の夕方、「災害時における帰宅困難者支援に関する協定」を結んでいるコンビニエンスストア、ファミリーレストランなどに対して、支援の要請を行った。要請を受けた各事業者では、急増する来店者への対応に加え、店内トイレの貸出などをを行い、帰宅困難者を支援した。</p> <p>（国土交通省「首都圏整備に関する年次報告」）</p>				



10

授業構成

なぜマルチエージェント？

Why multi-agents?



エージェントとは？

What is the concept of agents?

エージェントとは？(1)

What is the concept of agents?

世間一般での意味

(1) 代理人。代理店。

a person or company that represents another person or company, especially in business.



(2) 芸能人・俳優・音楽家・スポーツ選手などの代理業者。

本人に代わって、出演や入団などの仲介・交渉を行う。

someone who finds work for actors, musicians etc, or who finds someone to publish a writer's work.

(3) 諜報員。情報提供者。

someone who works for a government or police department, especially in order to get secret information about another country or organization.



エージェントとは？(2)

What is the concept of agents?

工学での意味

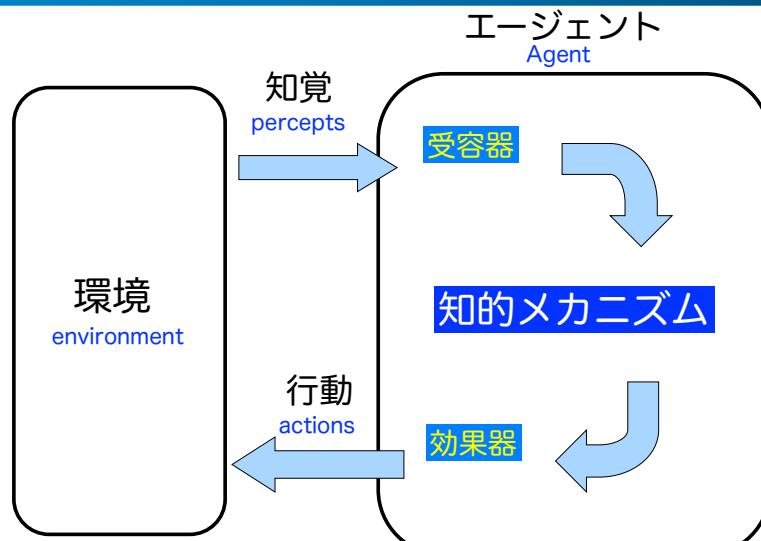
(4) マン-マシン-インターフェースの一。ユーザーの連続した操作を必要とせず、自律的に情報収集や状況判断を行い、適切な処理動作を実行できる機能。

また、そのソフトウェア。

autonomous entity which observes and acts upon an environment and directs its activity towards achieving goals

エージェントの基本フレームワーク

Agent overview



エージェントとは？(3)

What is the concept of agents?

本授業における定義 Definition of “agent”, in this lecture

抽象化したモデル

エージェントはある環境をセンサである受容器を用いて知覚し、効果器を通して行動するもの

An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through effectors

環境はどこだ？

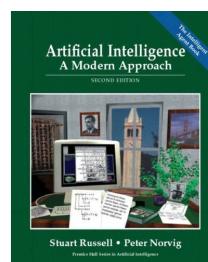


受容器はどこだ？

効果器はどこだ？

エージェントのタイプ

Russellの4分類



Artificial Intelligence
Stuart J. Russell, Peter Norvig
Pearson Education

反射エージェント

Simple reflex agents

内部状態に基づくエージェント

Agents with memory

ゴールに基づくエージェント

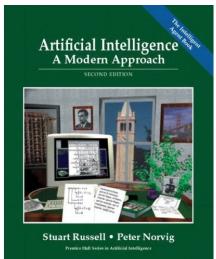
Agents with goals

効用に基づくエージェント

Utility-based agents

エージェントのタイプ

Russellの4分類



Artificial Intelligence
Stuart J. Russell, Peter Norvig
Pearson Education

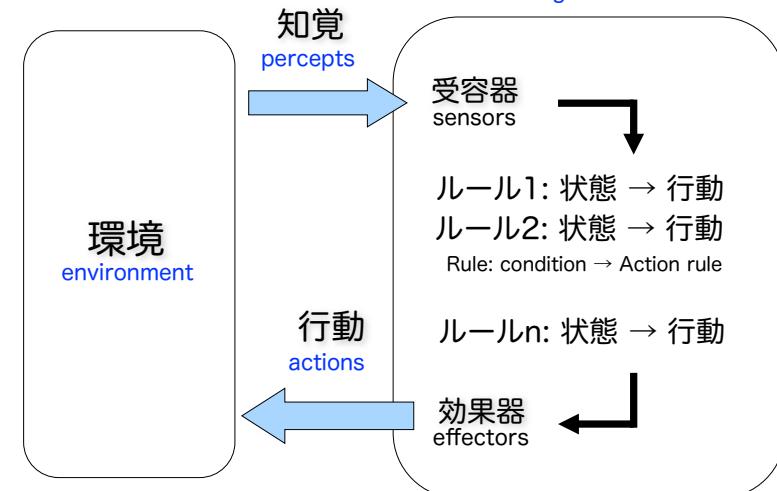
反射エージェント

Simple reflex agents

反射エージェント

Simple reflex agents

エージェント
Agent

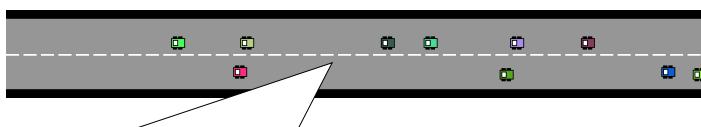


反射エージェントの例

Example of simple reflex agents

交通シミュレータ「ながれさん」の各車両

Transportation simulator: each car = simple reflex agent



ルール1： 前方に車両がなければアクセルON
Rule 1 : if there is no vehicle in front then Accelerator ON

ルール2： 前方に車両があればブレーキON
Rule 2 : if there is vehicle in front then Breaking ON

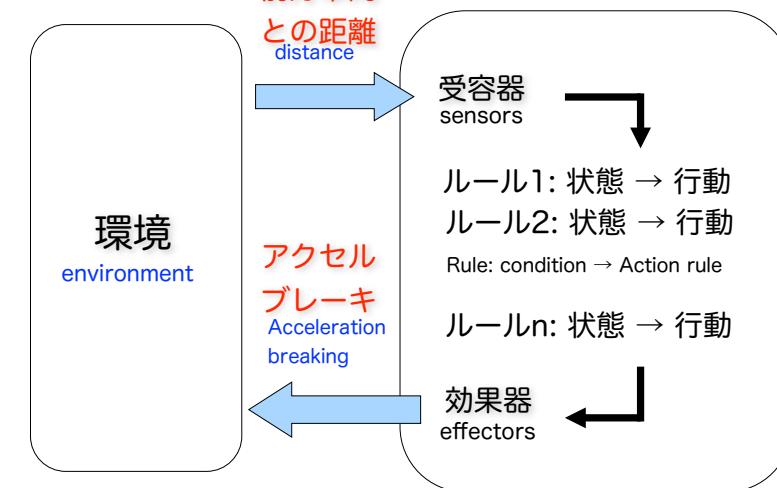
※実は、すでに反射エージェントの基本を学び、実装！

Actually, we have learned the simple reflex agents, already!

反射エージェント(再)

Simple reflex agents

車両
Car



ルールベースについて

About the rule-base

ルール1: 状態 → 行動

ルール2: 状態 → 行動

Rule: condition → Action rule

ルールn: 状態 → 行動

一言でいえば

「各状態に対応して行動が発火」

Action rule triggered by each state
(condition).

想定される困難 Assumed difficulties

①状態数の増加

Increase in the number of states

②明確な境界を持たない状態の発生

Occurrence of a state in which no clear boundary

「ファジィ集合」によるアプローチ

One of the solutions with respect to the above difficulties = fuzzy sets theory

(明確な境界をもたない状態を理解するための) アンケート調査

① 「暑い」と感じるのは何度?

Do you feel 'hot'? How many degrees?

②携帯電話を購入する際

「高い」と感じるのはいくら?

About cellular phone, do you feel 'expensive'? How much?



背の高さ

低い

高い

155cm 160cm 165cm 170cm 175cm

境界をぼかすためには?



背の高さを表すファジィ集合

帰属度

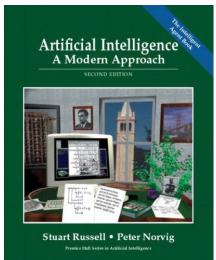
低い

高い

155cm 160cm 165cm 170cm 175cm

エージェントのタイプ

Russellの4分類



Artificial Intelligence
Stuart J. Russell, Peter Norvig
Pearson Education

反射エージェント

Simple reflex agents

内部状態に基づくエージェント

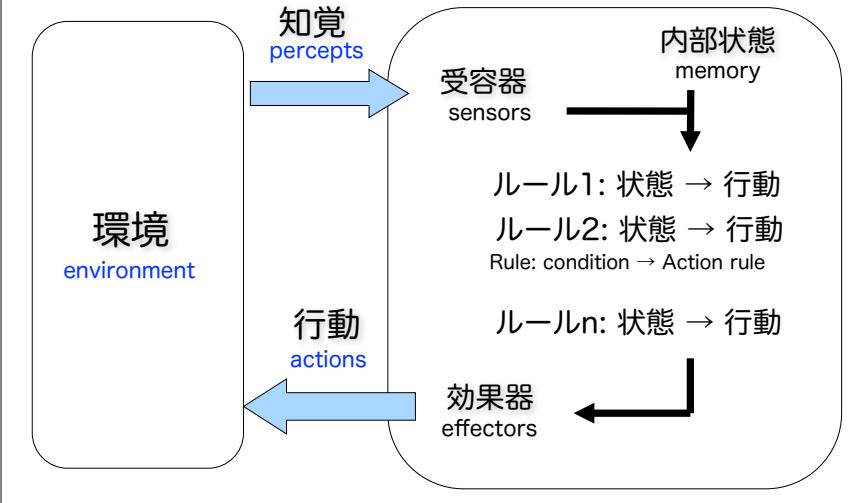
Agents with memory

内部状態に基づくエージェント

Agents with memory

エージェント
Agent

New!



内部状態に基づくエージェントの例

Example of agents with memory

感染シミュレータ「うつるんDeath」の人々

Infection simulator: each person = agent



内部状態： 自分は健康か、そうでないか
免疫を獲得しているか？

Memory = normal or diseased or immunity

ルール： 自分が病気に感染するか否か

Rule: infection or not

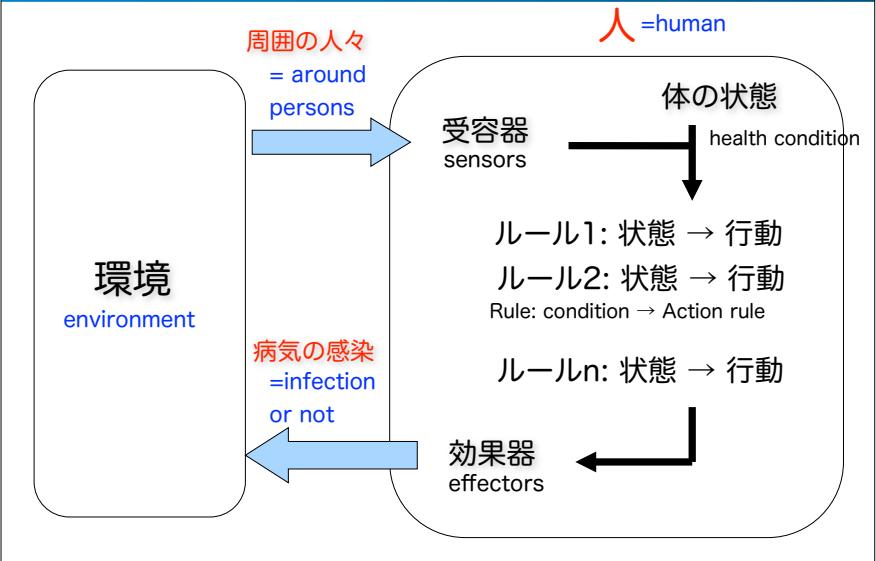
実はすでに基本のエージェントを学んでいた

Actually, we have learned the agents with memory, already!

Time=50
normal=1127
sickness=872

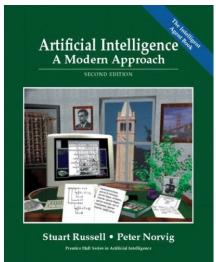
内部状態に基づくエージェント (再)

Example of agents with memory



エージェントのタイプ

Russellの4分類



Artificial Intelligence
Stuart J. Russell, Peter Norvig
Pearson Education

反射エージェント

Simple reflex agents

内部状態に基づくエージェント

Agents with memory

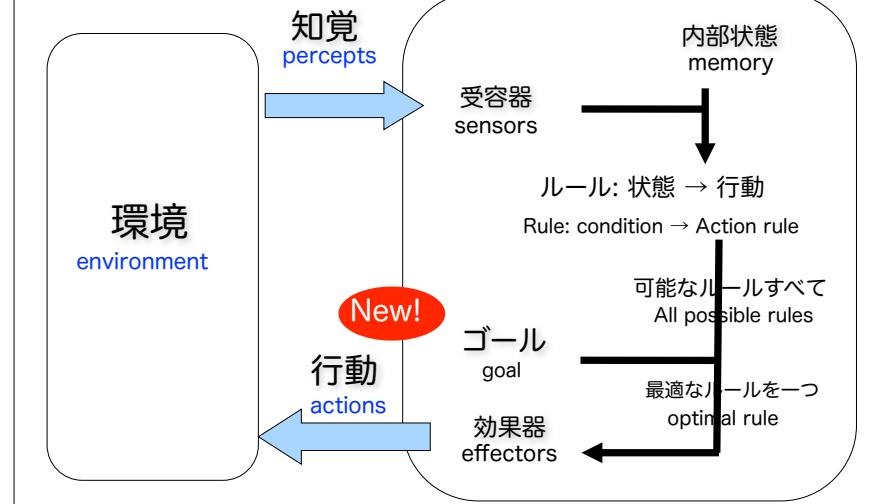
ゴールに基づくエージェント

Agents with goals

ゴールに基づくエージェント

Agents with goals

エージェント Agent



ゴールに基づくエージェントの例

Example of agents with goals

※今週の演習で、このシミュレーションを作成します
Exercise of this week, we will try to implement this simulation

食べ物を探し捕獲する

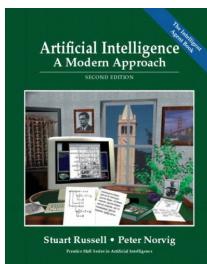
Looking for foods



近くにいる敵を倒す
Defeat the around persons (enemy)

エージェントのタイプ

Russellの4分類



Artificial Intelligence
Stuart J. Russell, Peter Norvig
Pearson Education

反射エージェント

Simple reflex agents

内部状態に基づくエージェント

Agents with memory

ゴールに基づくエージェント

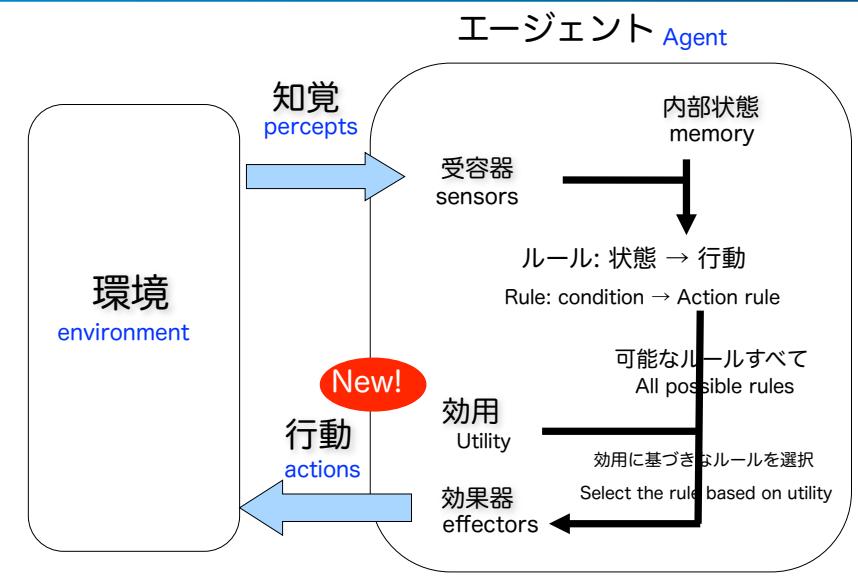
Agents with goals

効用に基づくエージェント

Utility-based agents

効用に基づくエージェント

Utility-based agents



自分の直感を信じてよいか？

Is it better to believe your Intuition?

以下の2つのクジのうち、どちらかを1回だけひけるとしたらどちらを選択する？

※直感で答えてください。 Which lottery will you play?

クジ1:

確率 0.4 : 10,000円
確率 0.6 : 2,000円

Lottery type 1
Prob. = 0.4: JPY 10,000
Prob. = 0.6: JPY 2,000

10回に6回は2,000円になってしまう！

クジ2:

確率 0.1 : 6,000円
確率 0.9 : 5,000円

Lottery type 2
Prob. = 0.1: JPY 6,000
Prob. = 0.9: JPY 5,000

10回に9回は確実に5,000円ゲット！

人間の直感を考慮しながら

効用を取り扱うのは非常に難しい

It is quite difficult to treat with the utility and human-intuition.



効用に基づくエージェントの例 (1)

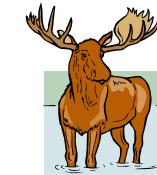
Example of utility-based agents

①今晚の食糧がない場合はどちらを選択するか？

① We have no food tonight

②3日分の食料が確保されている場合は？

② We have enough food (for three days)



残りの弾は1発

You have only one bullet (one shot)



効用評価が異なれば

選択も異なる



遠くにいる大物
(リスク大)
Big game but long distance,
that is, high risk.

近くにいる小物
(リスク小)
Small game but short distance,
that is, low risk.

おまけのシナリオ：ちょっと考えてみてください

Extra scenario: Please think about it.

①ある日、突然、警察に捕まる

One day, suddenly, you get caught by the police



②密室での長時間の取り調べ開始

Long interrogation in a closed room

③弁護士到着

A lawyer comes

正義の味方と思っていた弁護士のセリフ

The lawyer said...

99.9%の確率で有罪（懲役 数ヶ月）

99.9% = Not guilty

0.1%の確率で無罪

99.9% = Not guilty

選択肢1:

罪を認める

*無罪にもかかわらず

示談金、約30万円

選択肢2:

無罪を主張する

数ヶ月に渡る裁判

1000件に1件しか無罪を勝ち取れない
有罪の場合、数ヶ月の懲役

自分の直感を信じてよいか？

Is it better to believe your Intuition?

やさしい新しい経済学 神経経済学(1)脳の特性から行動説明 大阪大学准教授
田中沙織
[新刊販促案]

小・中・大 価格： 売印 判型： リカルト

近頃は経済学は、できるだけ単純に世の中の経済現象を説明しようとしました。そのため、意思決定する人間は、自己的利益の追求と完全な情報の計算を待ち構え、関係に関する完全な情報を持っていると仮定しました。一方、現実の世の中では奇遇やお金に多くの人が参加します。複数のスーパーの商品価格をいちいち比較して買い物するのは一苦労です。

このような理論上の人間と、現実の人間との「ギャップ」が大きな問題として取り扱う動きが、20世紀後半に始まりました。例えば、人を得てすることよりも、損をすることが得をします。だから宝くじが売れるのですが、そのような人の心理学的な特徴が経済に関する行動にどう影響するのか、実験やアンケートを使って探るのが「行動経済学」です。

さらに、人間の癖やパターンはどうして生じるのかを、生物学的な立場から理解しようとする動きも出てきました。その一つが、これまでブラックボックスとして扱ってきた「脳」の特性から説明しようとする「神経経済学」という分野です。

ランチの店を決めるとき、林の売買を行ふとき、その他さまざまな場面において、私たちは時に注意深く、時に無意識に、意思決定をしています。神経経済学では、人間がこのような意思決定をするときの脳の動きに注目します。

脳の特定の部位にどのような回路で情報が伝わるのか、また内部物質の動きがどうなるのか、脳を傷つけずに調べます。そこで活躍するのが、磁気を使って脳の活動を調べる「fMRI(機能的磁気共鳴画像法)」という手法です。

神経経済学の目的一つは、経済行動に関する脳の仕組みを解明することで、癖やパターンを含めた人間の行動を説明できる経済理論を作り上げることです。

たなか・さおり 76年生まれ。奈良先端科学技術大学院大博士。専門は神経経済学

日本経済新聞 2013年6月5日朝刊より

得よりも損の印象が強く
記憶に定着

Impression of “bad” is stronger
than “good”

希少な事象の発生確率
を大きめに見積り

The probability of occurrence of
rare events...we tend to estimate it
higher than actual probability.



それでも
ボクは
やってない

監督・脚本 周防正行

周防正行



会議で闘う

負けない
私は
負けない



岩波書店

エージェントのタイプ



反射エージェント

Simple reflex agents

内部状態に基づくエージェント

Agents with memory

ゴールに基づくエージェント

Agents with goals

効用に基づくエージェント

Utility-based agents

高機能化 highly-functional

授業構成

なぜマルチエージェント？

Why multi-agents?



エージェントとは？

What is the concept of agents?

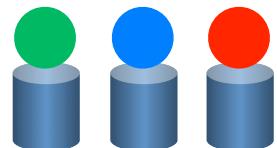
マルチエージェントとは？

What is the concept of multi-agents?

1) 分散環境と 2) 相互作用

Distributed environment and Interaction

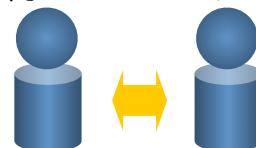
分散環境（個性がある）



Distributed environments: means that each agent has own AI

相互作用：

エージェント同士がコミュニケーションする



Interaction: each agent has communication with other agents, that is, each other.

マルチエージェントとは？

What is the concept of multi-agents?

(シングル) エージェント
(single) agent



(マルチ) エージェント
(multi) agents



1980年台初頭

分散人工知能の1分野として登場

The first concept is proposed as distributed AI, in 1980.

研究スタイル： 主に3つの項目を研究

1) 分散環境 Distributed environments

2) 相互作用 Interaction

3) ミクロ・マクロループ Micro-Macro loop

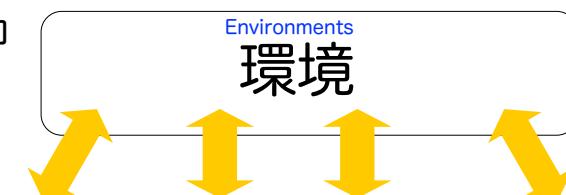
3) ミクロ・マクロループ

Micro-Macro loop

例：新人君達の参入で研究室の方向性が変化？

マクロ
Macro

Environments
環境



エージェント群

Agents

ミクロ

Micro



授業構成

なぜマルチエージェント？

Why multi-agents?



エージェントとは？

What is the concept of agents?

マルチエージェントとは？

What is the concept of multi-agents?

マルチエージェントと強化学習

Multi-agents and reinforcement learning

強化学習とは？ Reinforcement learning

環境の完全なモデルを持たない場合の

学習に関する研究分野

To research about the learning algorithm under the incomplete model of the environment

Q. その他の学習方法との違いは？

What is the difference between other learning methods, i.e., followings:

ニューラルネットワーク Neural networks

遺伝的アルゴリズム Genetic algorithm

サポートベクターマシン Support vector machine

エージェントと強化学習

Multi-agents and reinforcement learning

エージェントの学習・進化といっても・・・

究極的にはエージェントの持つルールベースをいじること

Learning process of multi-agents = modification of rule-base

ルール価値の修正

Rule-value modification

強化学習※ reinforcement learning

ファジィ理論（主観・あいまいさの導入）

ルールの生成・削除

Rule generation/delete

遺伝的アルゴリズム Genetic Algorithm

遺伝的プログラミング Genetic Programming

強化学習とは？ (Reinforcement learning)

ニューラルネットワーク

遺伝的アルゴリズム

サポートベクターマシン

教師つき学習

Supervised Learning

外部信号のみによるチューニング

正しい入力と出力のデータセットが必要

We need to prepare the correct input and output datasets

① (教師ありなので) 試行錯誤なし No try and error

② 即時報酬 Immediate reward

強化学習とは？ (Reinforcement learning)

② 即時報酬 Immediate reward



例：ピアノのレッスン

Ex. Piano lesson

間違えたらすぐに修正

ミスのない教師による常時監視

If you make a mistake, suddenly, the teacher give you correct answer.

強化学習とは？ (Reinforcement learning)

強化学習の特徴

① 試行錯誤的であること

Heuristic (try and error)

② 報酬が時間遅延を伴って発生

Delayed rewards



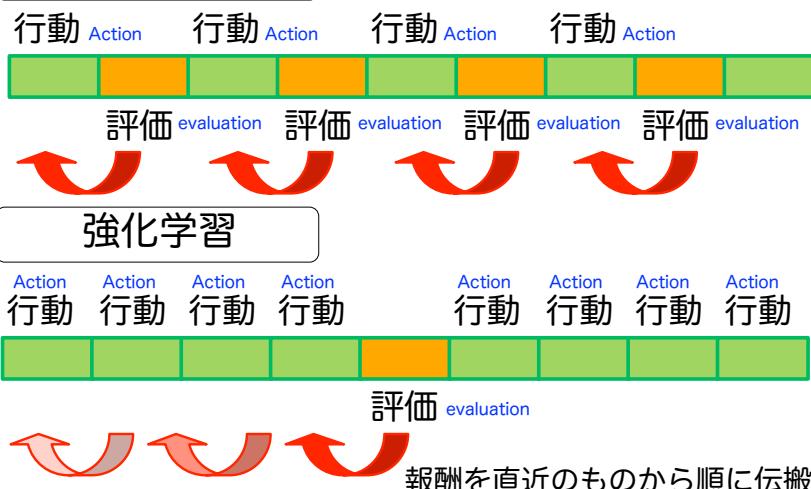
例：サッカーチーム Example: Soccer team

- 複数のプレーが連動した上でゴールが決まる
- ゴールのときに報酬途中の各プレイヤーの動きに對して明確な報酬を設定できない

The goal is achieved through the sequential/several players behavior, therefore, we cannot define the rewards about each behavior, clearly.

強化学習とは？ (Reinforcement learning)

NNなど



強化学習の歴史

History of Reinforcement Learning

1955 1960 1965 1970 1975 1980 1985 1990 1995 2000

TD学習
[Sutton 88]

動的計画法
[Bellman 57]

Q学習
[Watkins 92]
Sarsa
[Rummery 94]

学習分類子システム
[Holland 78] バケツリレー・アルゴリズム
[Holloand 85]

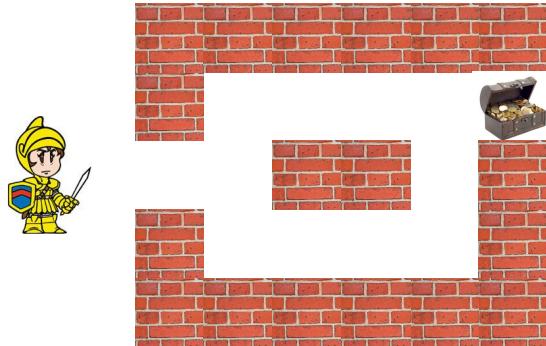
モンテカルロ法
[Holland 68] Profit Sharing
[Grefenstette 88]

LMS法
[Holland 60]

1984年 サンタフェ研究所設立
コンピュータによるシミュレーション能力向上

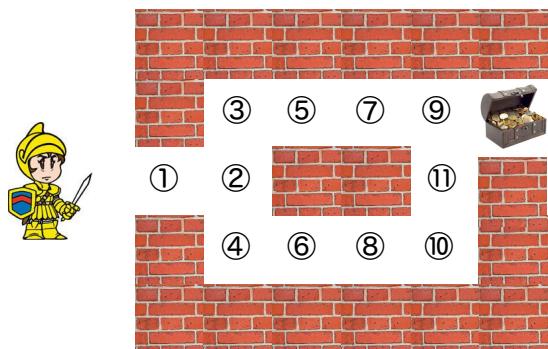
強化学習の登竜門 TD学習

TD (Temporal Difference) 学習 のエッセンスをRPGダンジョンで概説



ステップ2: 行動状態を定式化

Step2: Initialization of actions



状態①：右

状態②：上下左

状態③：下右 ···



State ①

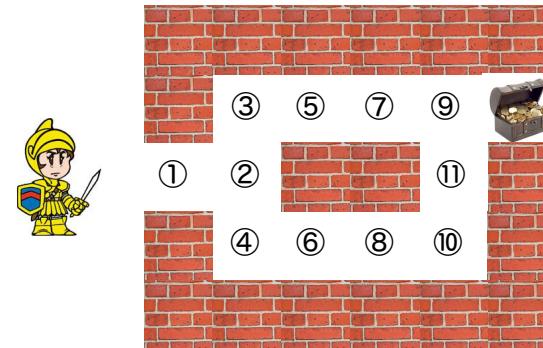
State ②

State ③

TD Learning

ステップ1: (位置)状態を定式化

Step1: Initialization of states

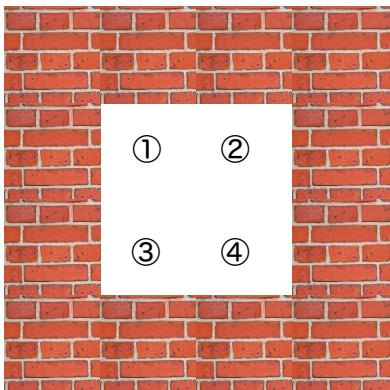


ステップ3: 状態と行動の組み合わせ

Step3: combination of states and actions

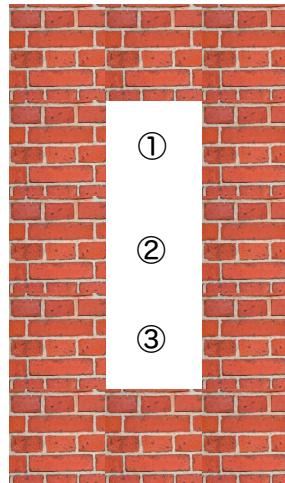
番号 (number)	位置状態 (states)	行動候補 (actions)
1	①	右
2	②	上
3	②	下
4	②	左
5	③	下
6	③	右
7	④	上

本当にわかっているかどうか確認①

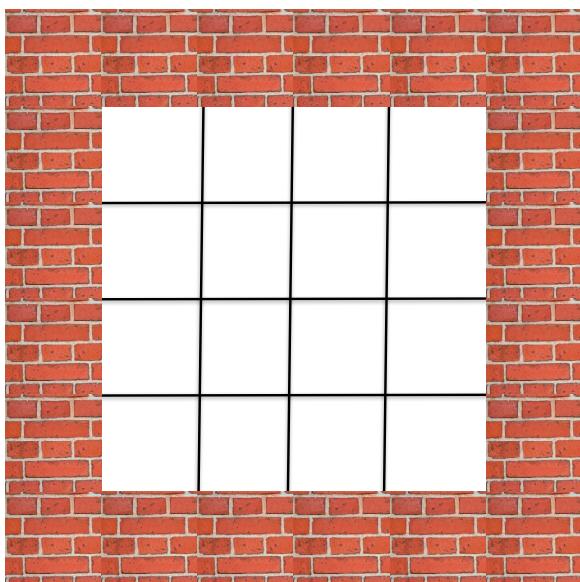


- (1) 状態数は？
- (2) 各状態の行動の数は？
- (3) 状態-行動の組み合わせの数は？

本当にわかっているかどうか確認②

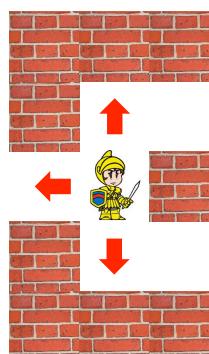


- (1) 状態数は？
- (2) 各状態の行動の数は？
- (3) 状態-行動の組み合わせの数は？



一応、なんとなく動くことはできるが・・・

Then, the agent can move, somehow, but...



どの選択肢をとるべきか?
Which choice should the agent take?

各状態に対する「価値」を定義する必要有
We need to assign the value to each state

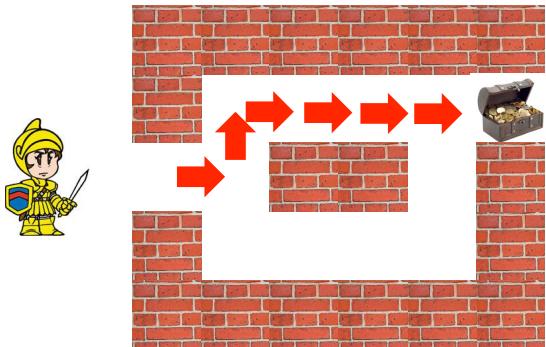
但し、この「価値」はわからない
※学習して獲得する！
We have no idea about the value, that's why, leaning!

価値の高い順に遷移してゆけば最適戦略
Based on the values which obtained by learning, we determine how to do the transition, and it means the optimal strategy.

TD学習

TD Learning

理想：下図のように推移するように各状態に価値を与える
Purpose: we try to obtain the state values that the agent can move as following.

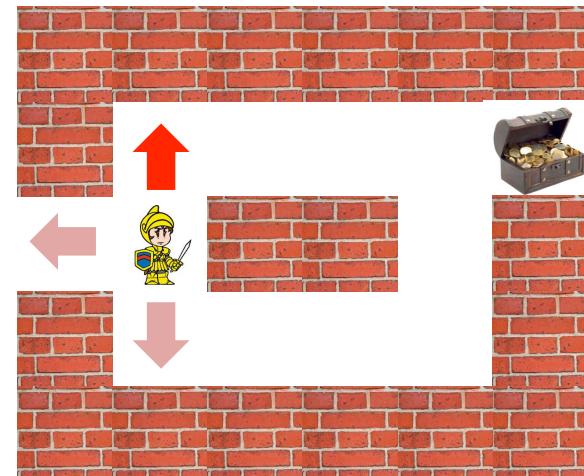


TD学習とは 「状態価値更新の方法論」

TD learning = an algorithm for updating state values

下記の位置で「左」「下」よりも「上」に移動してほしい

For example, the following position, we hope that the agent will choose upper direction, not left and lower directions



ステップ3: 状態と行動の組み合わせ

Step3: combination of states and actions

番号 (number)	位置状態 (states)	行動候補 (actions)
1	①	右
2	②	上
3	②	下
4	②	左
5	③	下
6	③	右
7	④	上

3,4よりも2
を選択して
ほしい

TD学習の基本更新式

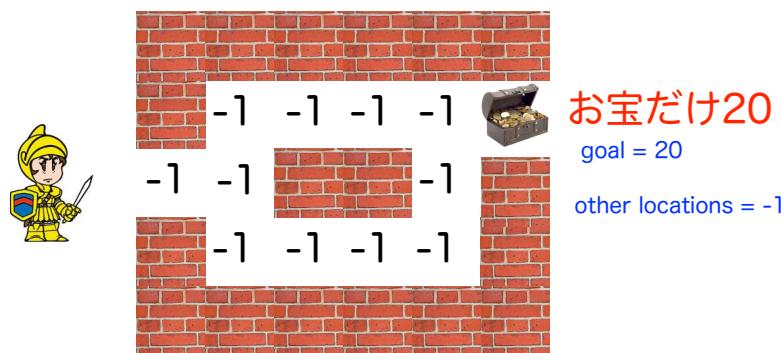
Update process of TD Learning

$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

↓ ↓ ↓ ↓
Current state value Next state reward Next state value Current state value

↓
Next state value

ステップ4: 報酬設定 Step4: Initialization of rewards



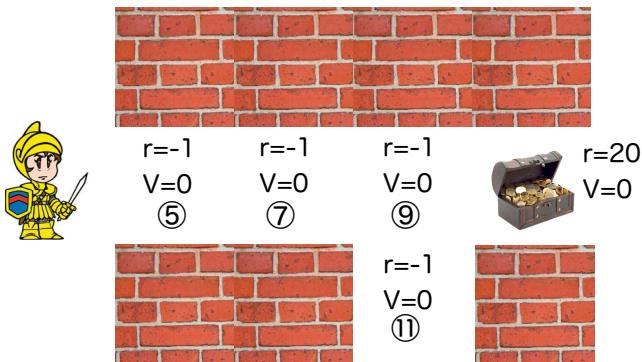
※現実世界でも、この報酬設定は試行錯誤！

In the real situation, we will assign the reward, that is, heuristic.

※各ブロックに報酬が割り当てられているので
即時報酬になっているが、割り当てなければ遅延報酬

(時間短縮のため) お宝手前で学習してみよう

In order to save the time, we will see only the learning process near the goal.



$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$



1エピソード目

更新前

状態⑤ r=-1 V=0	状態⑦ r=-1 V=0	状態⑨ r=-1 V=0	お宝 r=20 V=0
--------------------	--------------------	--------------------	-------------------

$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

$$\alpha = 0.1, \gamma = 0.9$$

更新後

状態⑤ r=-1 V=-0.1	状態⑦ r=-1 V=-0.1	状態⑨ r=-1 V=2	お宝 r=20 V=0
-----------------------	-----------------------	--------------------	-------------------



2エピソード目

更新前

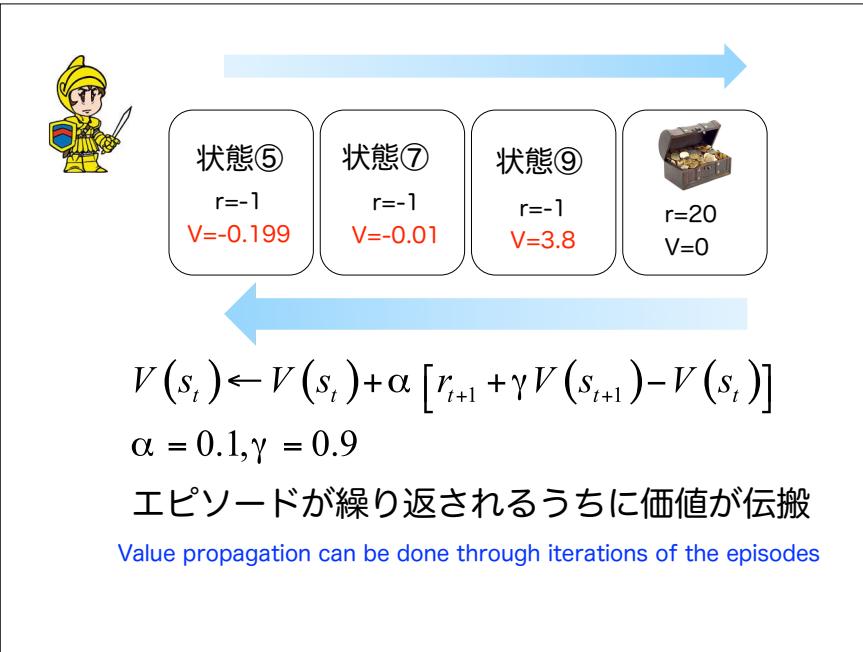
状態⑤ r=-1 V=-0.1	状態⑦ r=-1 V=-0.1	状態⑨ r=-1 V=2	お宝 r=20 V=0
-----------------------	-----------------------	--------------------	-------------------

$$V(s_t) \leftarrow V(s_t) + \alpha [r_{t+1} + \gamma V(s_{t+1}) - V(s_t)]$$

$$\alpha = 0.1, \gamma = 0.9$$

更新後

状態⑤ r=-1 V=-0.199	状態⑦ r=-1 V=-0.01	状態⑨ r=-1 V=3.8	お宝 r=20 V=0
-------------------------	------------------------	----------------------	-------------------

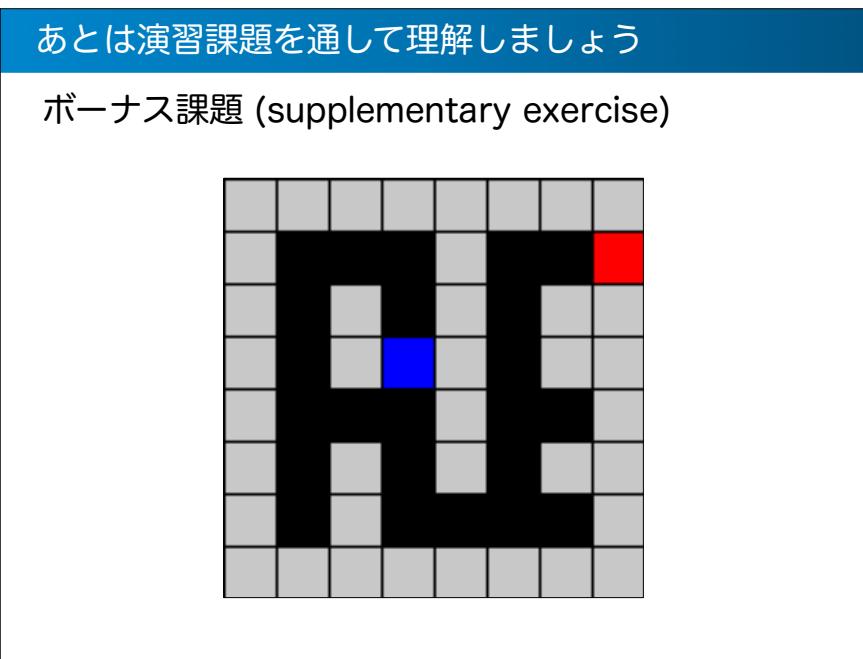


TD学習後の価値

After the TD learning, we obtained the value of each state-action.

番号 (number)	位置状態 (states)	行動候補 (actions)	価値 (value)
1	①	右	
2	②	上	
3	②	下	
4	②	左	
5	③	下	
6	③	左	
7	④	上	

各位置・行動に対応する価値がTD学習により決定



機械学習しっかり学びたい人は
澁谷先生の授業を履修してみてください。

さらに、良い書籍もあります！

discrete mathematics
応用事例とイラストでわかる
離散数学
カンタンな数学でAIも理解できる?
第2版 延原 肇(著)

6.1 はじめに 183
6.2 単純で難しい分割問題 185
6.3 数え上げの原理 190
6.4 順列と二項係数 194
6.4.1 順列 194
6.4.2 二項係数 198
6.5 順列と二項係数の応用事例 199
6.5.1 応用事例1：検索エンジンと順列 199
6.5.2 応用事例2：情報推薦と二項係数 200
6.6 AIと計算の複雑さ 201

73

スタート

-1	-1	-1	-1	-1
-1		-1		-1
		-1		
-1	-1	-1	-1	20

ゴール

図 6.11 迷路探索問題における報酬設定

授業構成

なぜマルチエージェント？

Why multi-agents?

エージェントとは？

What is the concept of agents?

マルチエージェントとは？

What is the concept of multi-agents?

マルチエージェントと強化学習

Multi-agents and reinforcement learning

マルチエージェントとゲーム理論

Multi-agents and game theory



ゲーム理論 (1)

Game theory

2人以上のプレイヤーの意思決定を分析する理論

Game theory is a study of strategic decision making, especially, it is "the study of mathematical models of conflict and cooperation between intelligent rational decision-makers"



オスカー・モルゲンシュテルン



再登場!

フォン・ノイマン

1944年 「ゲームの理論と経済行動」 が出発点

ゲーム理論 (2)

Game theory

ゲームの代表的な的な形式: types of games

同時ゲーム

Simultaneous



交互ゲーム

Sequential



ゲーム理論とマルチエージェント

「囚人のジレンマ」がよく登場

prisoner's dilemma



A, B ともに相談できない環境

Each prisoner is in solitary confinement with no means of speaking to or exchanging messages with the other.

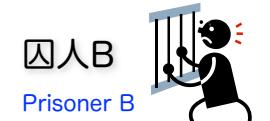
A, B ともに「自白する」（裏切り）あるいは

「自白しない」（協調）の選択肢がある

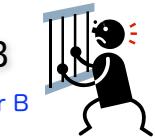
Each prisoner is given the opportunity either to betray the other, by testifying that the other committed the crime, or to cooperate with the other by remaining silent

囚人のジレンマ

囚人Aから見た損得表



囚人B
Prisoner B



自白する
defects



囚人A
prisoner A

自白しない
cooperation

懲役2年
2 years

懲役5年
5 years

自白する
defects

懲役1年
1 year

懲役3年
3 years

囚人のジレンマ

囚人Bから見た損得表



囚人A
Prisoner A

自白しない
cooperation

懲役2年
2 years

懲役5年
5 years

囚人B
Prisoner B

自白しない
cooperation

懲役1年
1 year

懲役3年
3 years

自白する
defects

懲役1年
1 year

懲役3年
3 years

均衡点：お互いに相手の戦略に合わせて損をしないように繰り返し決定を行っていった場合に収束する点のこと

繰り返し囚人のジレンマ

Iterated Prisoner's Dilemma

囚人のジレンマを、複数回繰り返し行う
(以降IPD: Iterated Prisoner's Dilemmaと略す)

1回きりの囚人のジレンマと異なる点
「自分が自白（裏切り）した場合、相手からの報復を受ける可能性がある」

If two players play prisoners' dilemma more than once in succession and they remember previous actions of their opponent and change their strategy accordingly



ルールの確認

1試合あたり、5回の試行を実施。1回あたりの試行では、200回の囚人のジレンマを繰り返す
opponent

		相手が協力 cooperation	相手が裏切り defects
our side	自分が協力 cooperation	利得：3点 3 points	利得：0点 0 points
	自分が裏切り defects	利得：5点 5 points	利得：1点 1 points

試合例 (example)

Strategy 1

戦略1: 常に裏切る

Always defects

Strategy 2

戦略2: 奇数回目で協力

偶数回目で裏切り

Odd term: cooperation

Even term: defects

	1回目	2回目	3回目	4回目	5回目
戦略2 (Strategy2)	0	1	0	1	0
戦略1 (Strategy1)	5	1	5	1	5

戦略2の利得は2, 戦略1の利得は17

ちょっと演習 (exercise)

ではやってみましょう！

ステップ1. 自分の戦略を組み立ててください。

目標は「どんな相手でも、最も得点が

高くなるような戦略を立案」

相手によって最適戦略が変化するので、最善手はない

注意① ランダムや確率を導入しないでください

注意② 常に各回で戦略が確定するようにしてください。

注意③ 相手の戦略の戦略に基づく場合は

過去の履歴のみ参照できます。

※遅出しの、じゃんけんは禁止

ではやってみましょう！

ステップ2. 3人グループになりリーグ戦で
対戦してください

ステップ3. 対戦後、自分の利得を合計してください

3人いない場合は、TAさんがヘルプします。

結果発表

獲得合計点 0点～10点

獲得合計点 11点～20点

獲得合計点 21点～30点

獲得合計点 31点～40点

獲得合計点 41点～50点

最後に集計

今年は、どのような戦略が強かったのか？

・・・皆さんで検討しましょう

繰り返し囚人のジレンマの興味深い研究結果

Some research reports on IPD

1979年 ロバート・アクセルロッドは多数の識者
に対し、IPDの戦略を募集した

Axelrod invited academic colleagues all over the world to devise computer
strategies to compete in an IPD tournament.

13コの戦略+ランダム戦略の計14戦略
による総当たりの試合を実施
13 strategies and random strategy, total 14 strategies.

1試合あたり、5回の試行を実施。

1回あたりの試行では、200回の

囚人のジレンマを繰り返す

1 Game = 5 trials, and each trials is 200 times IPD.



繰り返し囚人のジレンマの興味深い研究結果

Some research reports on IPD

どんな戦略が優勝したでしょうか？

Whish one is the most excellent (strongest) strategy?

	相手が協力	相手が裏切り
自分が協力	利得：3点	利得：0点
自分が裏切り	利得：5点	利得：1点

その後、もっと多くの戦略を募集

After the tournament, he call for strategies, again.

がしかし、またもラポポートの戦略が優勝

However, Rapoport strategy is the winner, again.

試合成績を解析

- TFTのような戦略が上位半分
- 下位半分の大半は、自分から裏切る戦略
- 最下位はランダム戦略



なぜ？

優勝した戦略は？



トロント大学の心理学者

アナトール・ラポポートの「しっぺ返し」戦略
BASIC言語のプログラムでわずか4行

The winning deterministic strategy was tit for tat, which Anatol Rapoport developed and entered into the tournament. It was the simplest of any program entered, containing only four lines of BASIC,

ポイントとなる考え方

- ① 自分から決して相手を裏切らない
- ② 相手に裏切られても一度だけ裏切ってあとは根にもたない

When the opponent defects, on the next move, the player sometimes cooperates anyway, with a small probability (around 1–5%). This allows for occasional recovery from getting trapped in a cycle of defections.

試合例その(1) : example (1)

TFTどうしの試合の場合 (TFT vs TFT)

	相手が協力	相手が裏切り
自分が協力	利得：3点	利得：0点
自分が裏切り	利得：5点	利得：1点



お互いに協調してゆくので
200回 × 3点 = 600点獲得

試合例その (2): example (2)

TFT 対 常時裏切り戦略
(TFT vs Always defects)

	相手が協力	相手が裏切り
自分が協力	利得：3点	利得：0点
自分が裏切り	利得：5点	利得：1点

TFT 0 1 1 1 1 1 1 1

裏切り 5 1 1 1 1 1 1 1

お互いに、むなしい裏切りの応酬が続いた結果
TFT 199点、裏切り204点

満員電車はなぜケンカが起きやすい？

恒久的な人間関係が期待される場合

忍耐 > 自己

職場, 研究室, サークル, 学校のクラス等

一時的な人間関係で終わる場合

忍耐 < 自己

旅行先, 満員電車など。

旅行先では（無意識のうちに）クレームが出やすくなる
満員電車において、普段のストレスが爆発しやすい

試合結果の総括

構成メンバーにTFTが多ければ、お互いに
協調して得点を高くすることができる

If almost strategies of tournament are TFT, they will cooperation each other, and obtain higher score than other strategies.

※世の中もTFTであれば、うまくやつていけるか？



学生食堂などはTFT?



海の家は裏切り戦略でも
やっていける？

授業構成

なぜマルチエージェント？

[Why multi-agents?](#)



エージェントとは？

[What is the concept of agents?](#)

マルチエージェントとは？

[What is the concept of multi-agents?](#)

マルチエージェントと強化学習

[Multi-agents and reinforcement learning](#)

マルチエージェントとゲーム理論

[Multi-agents and game theory](#)

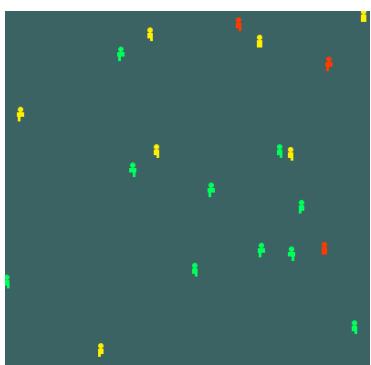
最終課題についての説明

[About the final report](#)

最終課題の説明の前に・・・

今週の課題について
About the report of this week

人工社会シミュレータ: ジロー君
Multi-agents simulator: jiro-kun



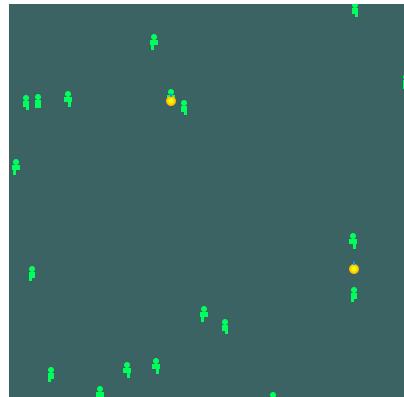
「体力（寿命）」と時間の概念の導入
This simulator employs the concept of life-time of each agent.

人工社会シミュレータ: イチロー君
Multi-agents simulator: ichiro-kun



何もない状態

人工社会シミュレータ: サブロー君
Multi-agents simulator: saburo-kun



捕食（食べ物）の導入
This simulator employs the concept of foods (helpful for survival of agents)

課題1: 人工社会シミュレータの改良

条件1:

季節（春夏秋冬）の概念を導入し、季節によって食物の発生頻度を変化させる機能

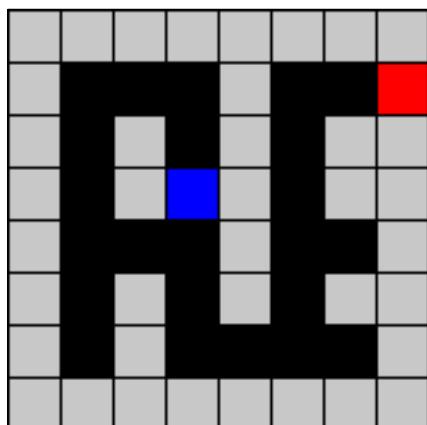
条件2:

各エージェントの体力が、ある一定値以上に
なったら増殖する機能

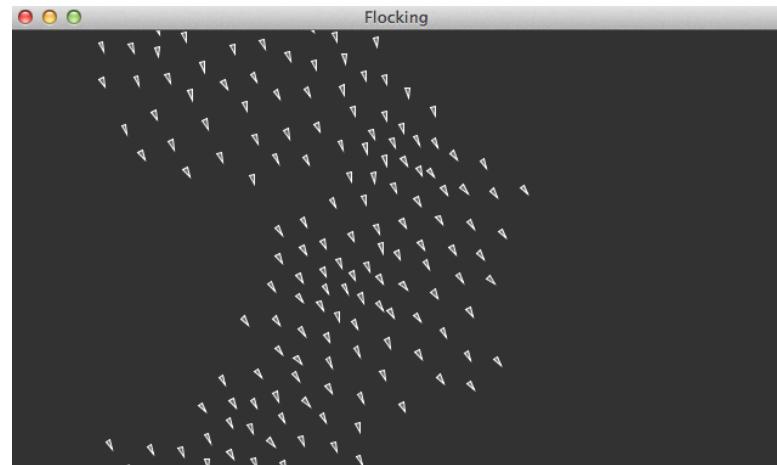
条件3:

2種類のエージェントを存在させ、異なるエージェント同士がケンカし、体力が減少する機能

ボーナス課題 (supplementary exercise)



課題2: boidシミュレーション



最終課題の説明の前に・・・
成績評価について

About evaluation of learning results

成績評価について (1)

About evaluation of learning results

授業への参加： 1回あたり2点 (合計 $2 \times 7 = 14$ 点)

注意： **理由なく3回以上欠席した場合は
単位取得不能です。絶対にダメです！
来年がんばってください。**

Participation in the class: 2×7 classes = 14 points

成績評価について (2)

About evaluation of learning results

第2回： 12点

第3回： 13点

第4回： 12点

第5回： 13点

第6回： 10点

第7回： 17点

合計： 89点 ※ボーナス含まず。出席点含む

成績評価について (3)

About evaluation of learning results

最終課題： 最高30点



成績評価について (4)

About evaluation of learning results

出席点+課題点+最終課題点 の合計得点

90点以上 A+

80点以上 A

70点以上80点未満 B

60点以上70点未満 C

その他は D

最終課題について

About final report

最終課題について (1)

About final report

提出期限： 2023/6/26 15:00

課題：

スマートインフォメディアに関連した方法論、
シミュレータ、作品等を自由に作成

Final repot theme:

Please create simulator, or method (algorithm), theory,
related to the smart info-media system.

使用言語はなんでもOK！

You can use any programming language, whatever you want.

最終課題について (2)

About final report

1) 映像で提出すること！

Submit your report as movie.

2) 映像は最長 2分間まで。2分超えたら失格。

The length of movie = maximum 2min!

Videos over 2 minutes will not be evaluated.

3) フォーマットはOneDriveで再生できるもの。

後日、全員が閲覧できるようにするので注意

We strongly recommend the video format that can be play by OneDrive,
we mean, well known video format.

Because, all students of class will watch your video, later.

最終課題について (3)

About final report

4) 提出方法:

※office 365のOneDriveにアップロード

Please upload your movie to OneDrive (Office 365)

※アップロードした動画が再生できることを確認

Please confirm the uploaded movie

※ファイルの共有のリンクを

テキストに貼り付けてmanabaから提出

Please paste your file sharing links in text and submit via manaba

提出がない場合には単位は出ません！

※期限は厳守しましょう！

Please be careful about the deadline of final report.



相互評価について

それぞれの学生からの提出期限： 2023/6/26 15:00

↓ 延原の方で、相互評価環境の整備

相互評価期間： 2023/07/01～2023/07/07

↓ 延原の方で、相互評価環境結果の集約

相互評価期間： 2023/07/14 結果発表！

受講生全員が他の受講生の動画を閲覧・評価

相互評価（作品1から10まで）

5段階評価で評価してください。以下の点数は目安です。

- 5点・・・素晴らしい。最優秀作品に相応しいと判断されるもの
- 4点・・・なかなか良い。上位20%くらいには位置すると判断されるもの
- 3点・・・普通
- 2点・・・イマイチ評価できない
- 1点・・・全く評価できない

作品1について評価してください。**(選択必須)**

1.1

- 1. 1点
- 2. 2点
- 3. 3点
- 4. 4点
- 5. 5点

他の受講生に対する相互評価を行わないと単位が出ません！

なぜ、このような相互評価をするのか？

最終課題について (4)

(伝説の) 最優秀作品賞

東本遼太郎 君

MazeCreators

-迷路を作るエージェント-

最終課題について (4)

昨年の最優秀作品賞

加藤 敦也 君

モジホコリシミュレータ

最終課題について (4)

ユーモア賞もあります

最終課題について(5)：おまけ

さらに最優秀作品賞

2011年11月22日 18時46分 投稿



交通事故・渋滞シミュレータつくってみた

再生: 398,060 | コメント: 21,940 | マイリスト: 8,145

ニコニコ技術部 カテゴリ前日総合順位: 156位 (過去最高: 1位)

ツイート 2,588

8+1 4

いいね! 58

共有▼

最後に、授業アンケート(TWINS) に協力をお願いします



筑波大学
University of Tsukuba



延原 嘉



スマホ版



English



残り約30分



ポータル設定



ログアウト



HOME



履修



成績



学生情報



学生検索



科目検索



休憩・カジュー



アンケート



掲示



課外活動

アンケート作成 アンケート回答・集計結果照会

新着情報 あなたの宛の新着情報はありません。

MYスケジュール 2020年5月

Sun	Mon	Tue	Wed	Thu	Fri	Sat
1	2					
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30
31						

2020/05/26(Tue)

5限: 知能機能システムコアステディ
5限: 知能機能システムコアステディ

在学生へのお知らせ Explanatory Meeting of Master's/Doctoral Program in Education, University of Tsukuba

【アンケート回答・集計結果照会】 > 回答対象一覧

【アンケート回答・集計結果照会】/回答対象一覧

2020年度

【授業アンケート】 一般アンケートへ

タイトル	実施期間	状態
		回答対象アンケートはありません

【一般アンケート】 授業アンケートへ

タイトル	実施期間	状態
		回答対象アンケートはありません

【担当教員アンケート】

タイトル	実施期間	状態
OAL5503 スマートインフォメディアシステム特論 (スマートインフォメディアシステム特論/延原 嘉/月 5,月6)	2020/6/1 - 2020/6/22	回答画面
OAL5503 スマートインフォメディアシステム特論 (スマートインフォメディアシステム特論/延原 嘉/月 5,月6)	2020/6/1 - 2020/6/22	回答画面