

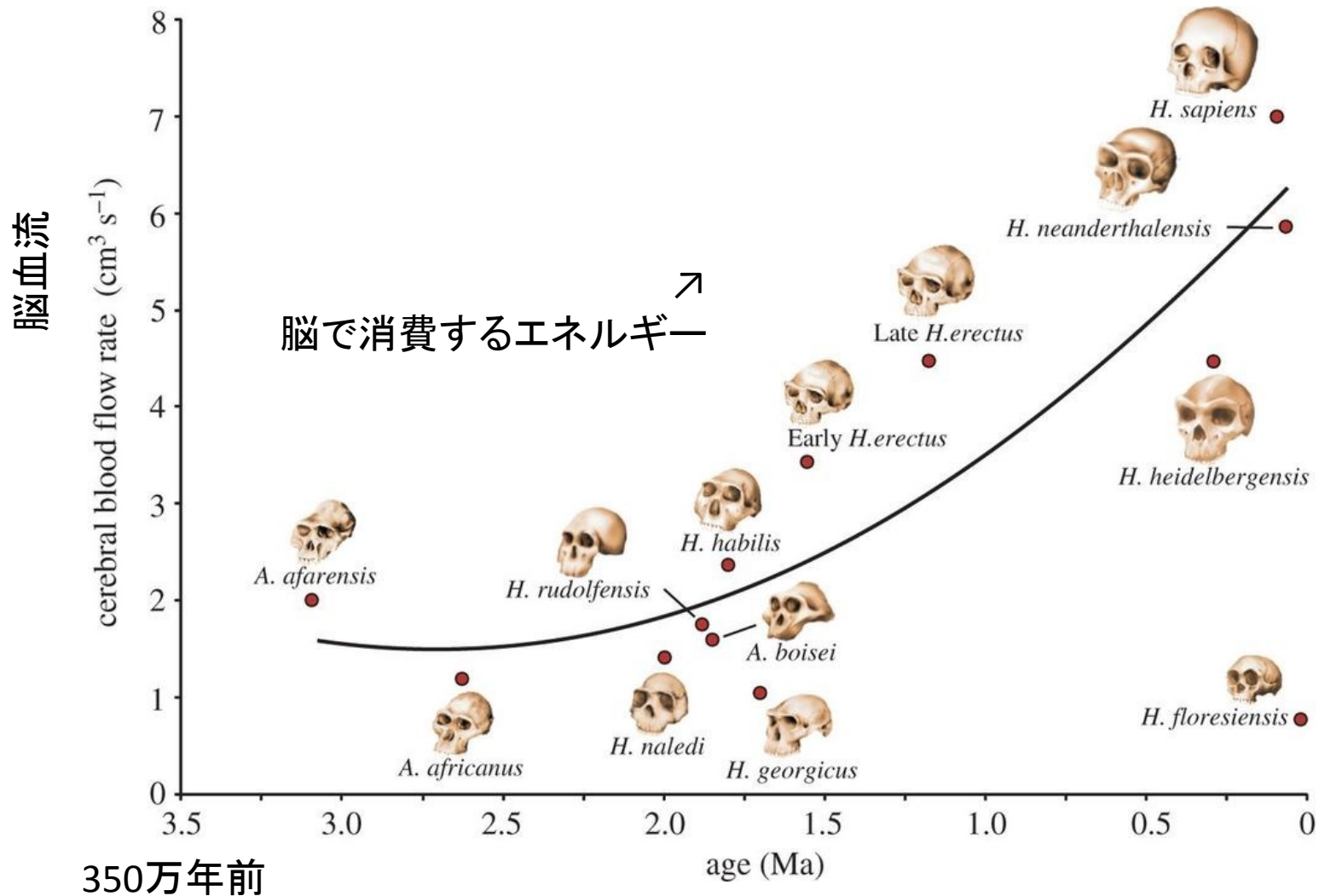
知能システム数理セミナーI

第3回 壮大な冒険の世界へ

知能情報研究室

橘完太

ヒトの知能の壮大な物語



状態認識と行動決定

状態 x_t
観測 z_t
行動 u_t

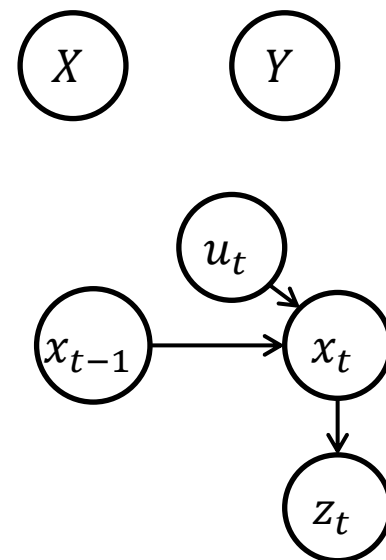
- 我々の知能は、
 - 状態の認識 $p(x_t | z_{1:t}, u_{1:t}) = Bel(x_t)$ と
 - 行動の決定 $Bel(x_t) \mapsto u_{t+1}$
- のためにある
- 我々には、真の状態は分からないので
 - 状態を、確率分布 $Bel(x_t)$ で持つ(信じる)
 - そして、確率アルゴリズムに従って、
行動 u_{t+1} を決定する。

状態認識

- センサ値＝状態(の一部)の観測＋観測誤差
- 観測モデル(センサモデル): この状態 x_t なら(条件)、この範囲のセンサ値が出る確率はこれくらい $p(z_t | x_t)$
- 状態＝前の時刻の状態＋制御＋制御誤差＋推定誤差
- 状態遷移モデル(システムモデル): この状態 x_{t-1} からこの行動 u_t をしたら、この範囲の状態になる確率はこれくらい $p(x_t | u_t, x_{t-1})$

グラフィカルモデル

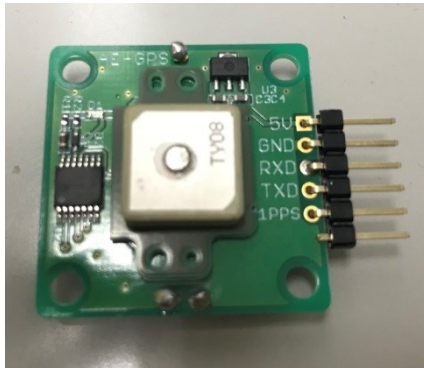
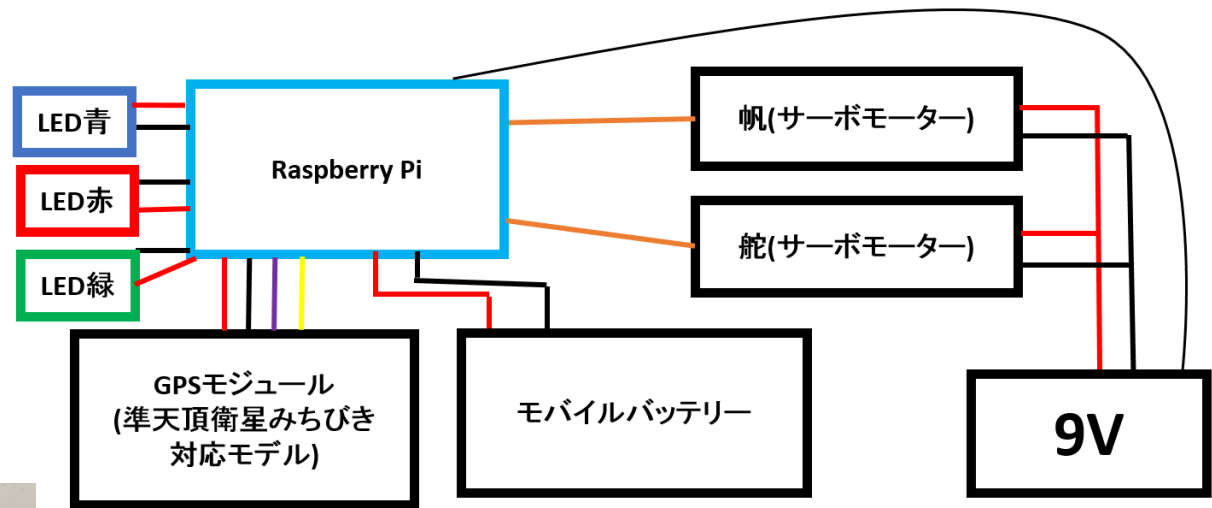
- 独立な関係なら結ばない
 - 弟にじゃんけんで勝つ日は降水確率が高い(?)
 - Y : 雨が降る、 X : 弟にじゃんけんで勝つ
 - $P(Y | X) = P(Y)$
- 依存関係は結ぶ
 - 舵を切ったら x_t に移動する
$$p(x_t | u_t, x_{t-1}) \neq p(x_t)$$
 - x_t にいたらセンサ値は z_t が出る
$$p(z_t | x_t) \neq p(z_t)$$



実用例 ロボヨット



実用例 ロボット



GPSモジュールで得られるセンサ値
時刻／緯度／経度／スピード／
方位角(※スピードが遅い時、怪しい)

観測からの状態推定

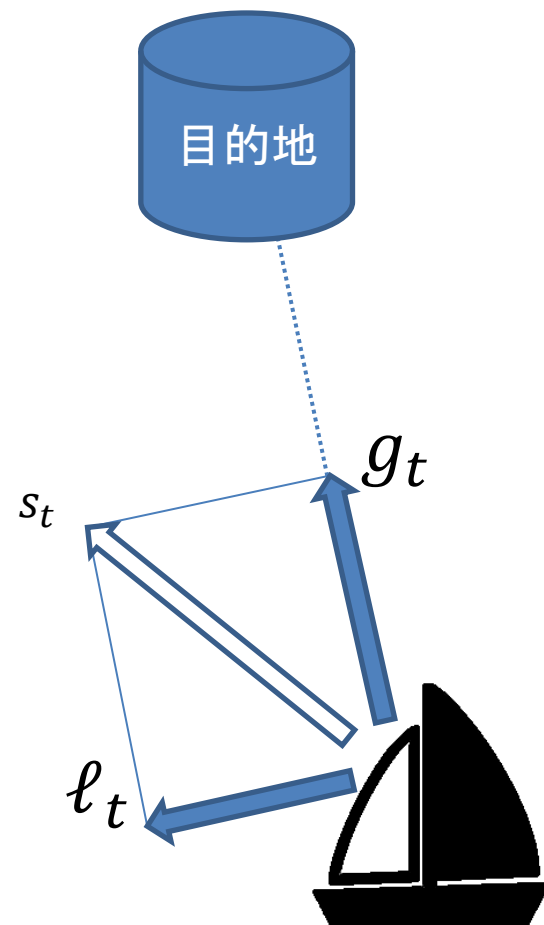
- 本来は、ベイズフィルタを使って、方位角など状態変数を推定すべき。
- H29卒論では、単純な手法でお茶を濁した。
- スピードが速い時 ($s_t > 0.3$)、方位角のセンサ値を信用する。
- スピードが遅い時 ($s_t \leq 0.3$)、少し前の緯度経度→最新の緯度経度の方向を方位角とする。

行動選択のための状態空間の定義

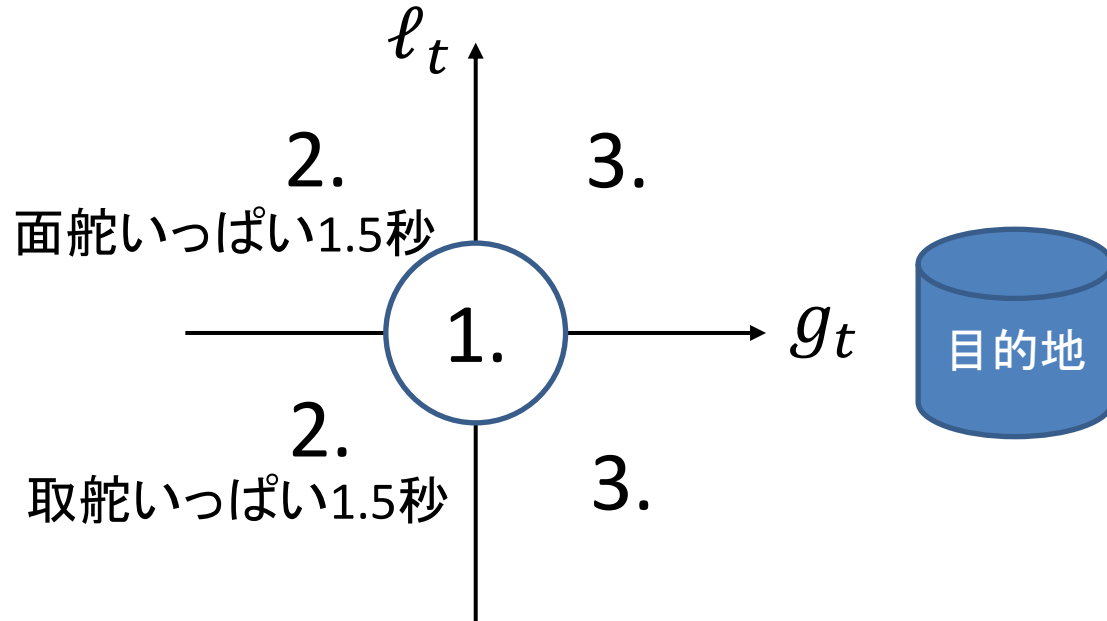
- 観測(緯度、経度、スピード、方位角)
- 状態(緯度、経度、スピード、方位角)
- としてもよいが、
- 池Aで得た状態空間と行動選択の知識
 $Bel(x_t) \mapsto u_{t+1}$ を、緯度・経度が異なる湖Bで
使えない。(池Aでも他の水域には使えない)
- 例えば、
状態(目的地への相対速度)
とすれば、他の水域でも知識を利用できる。

行動選択のための状態変数

- 状態変数2つ $x_t: (g_t, \ell_t)$
 - 速度の目的方向成分 g_t
 - 速度の目的地と直交成分 ℓ_t
- H29では、If-then rule方式で、行動 u_{t+1} を決定
- 行動は舵と帆の操作の組み合わせ



制御ルール If-then rule

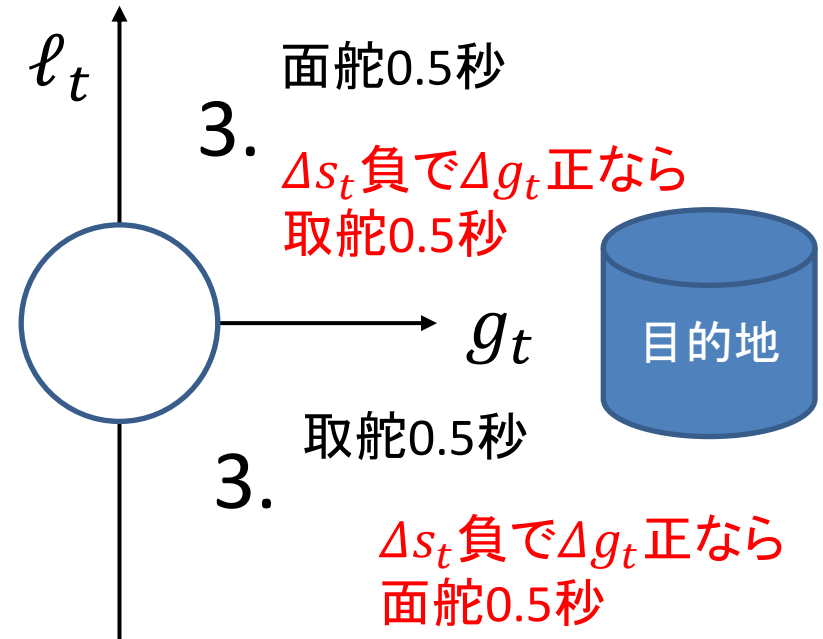


If Then

1. 方向はともかく、まず進む。
2. (目的地から遠ざかる時) 舵をいっぱい切る。
3. 帆と舵をちょっとずつ調整する。(次頁)

制御ルール If-then rule

- 帆と舵をちよつとづつ調整する。
($s_t = \sqrt{g_t^2 + \ell_t^2} > 0.3 \ \&\& \ g_t > 0$ のとき)
- $\Delta s_t = s_t - s_{t-1}$ と $\Delta g_t = g_t - g_{t-1}$ でさらに場合分けて、
舵と帆を微調整
- デッドゾーン(後述)の
対策
- 詳細は、SailH30.py

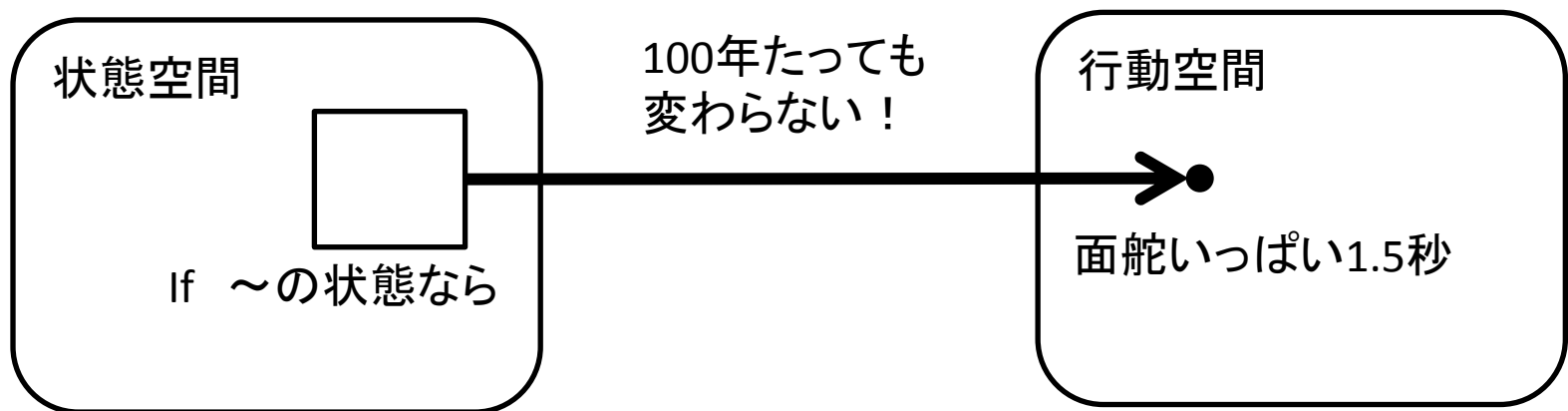


上記（お茶を濁した）コーディング の結果

- 次回5月8日のセミナーをお楽しみに！
- しかし、このコーディングの最大の欠点は、...

制御ルールが

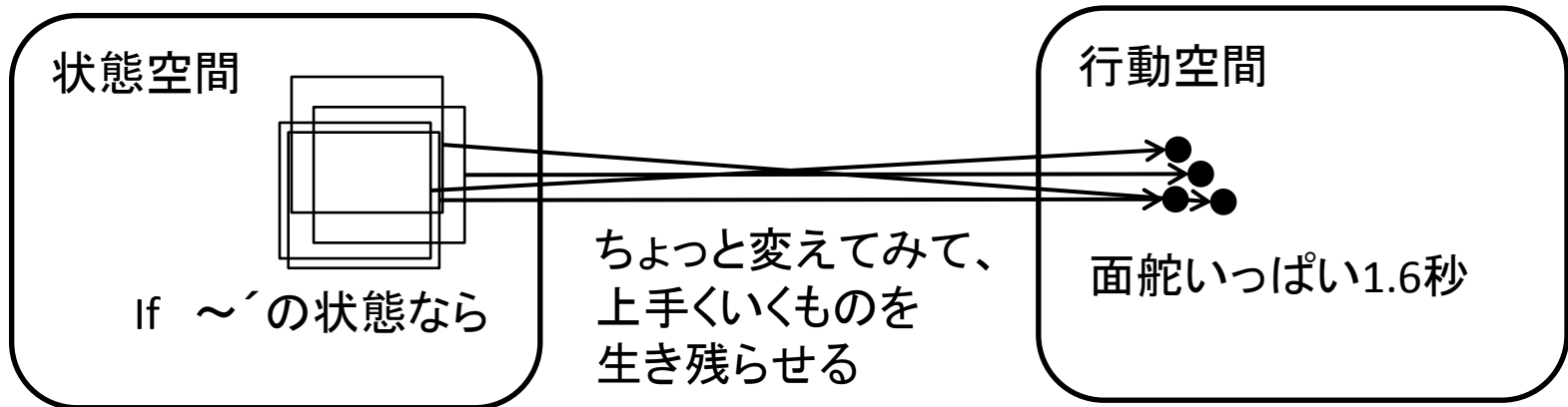
- 現時点のIf-thenルールは、卒論生A君の脳で考えたまま。学習で変わるわけではない。
- たとえ100年間、データを与え続けても、少しも賢くならない。



聡明なアイデア

- 強化学習 : reinforcement learning
- 進化学習 : neuro-evolution

<https://youtu.be/Aut32pR5PQA>

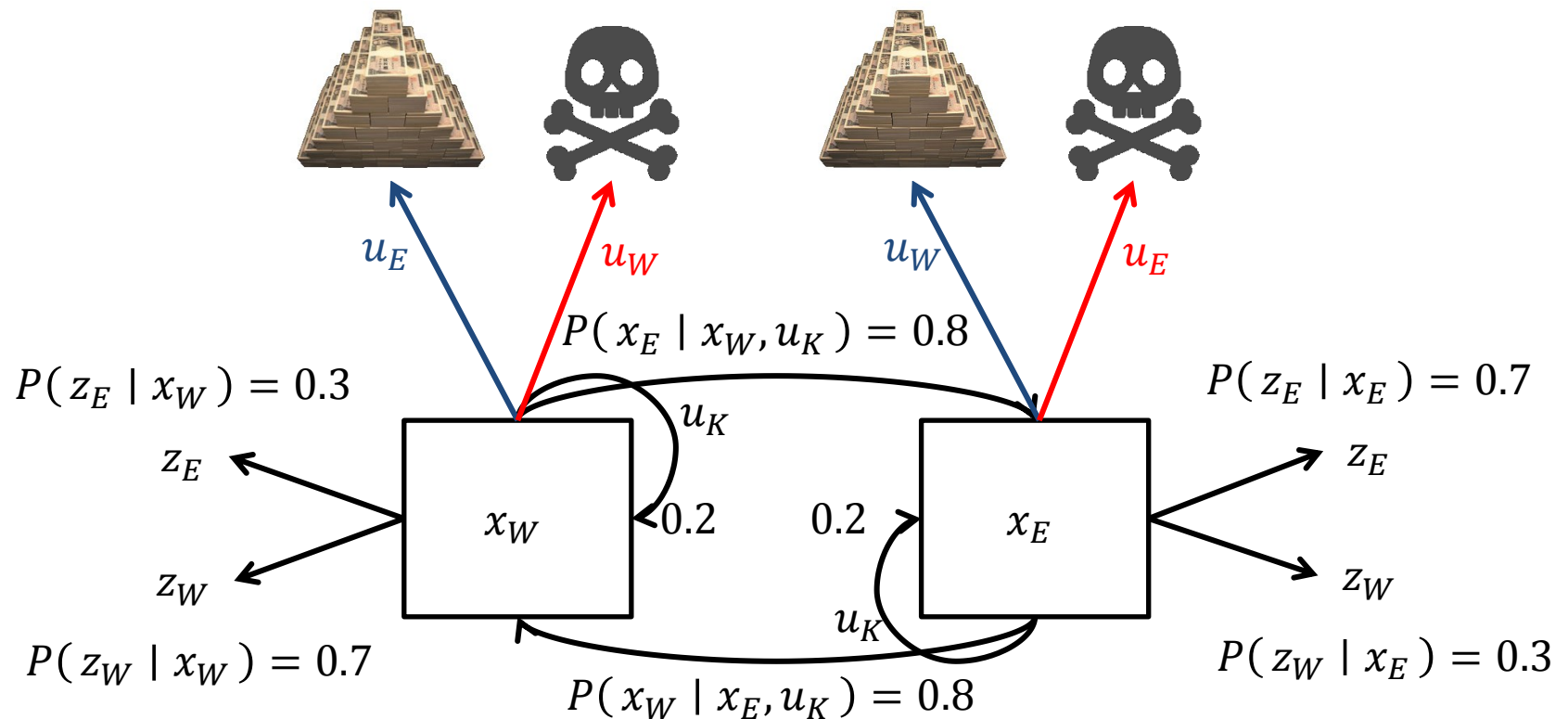


強化学習

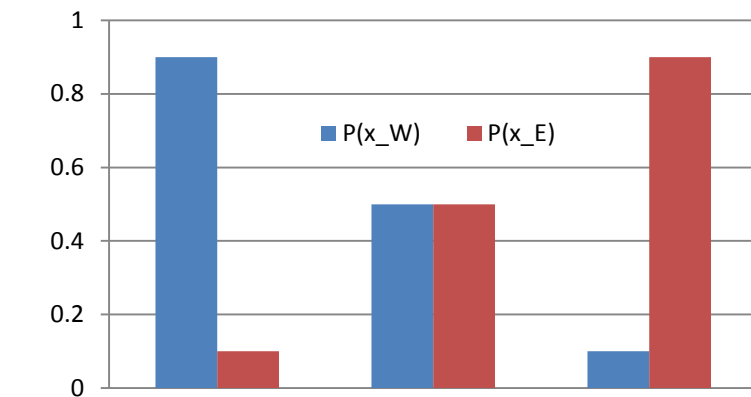
- 行動
 - u_E : 東扉から入る
 - u_W : 西扉から入る
 - u_K : 壁ドン→唸り声を聴く
- 状態
 - x_E : トラが東扉の近くにいる
 - x_W : トラが西扉の近くにいる
- 観測
 - z_E : 唸り声が東扉の方から聞こえる
 - z_W : 唸り声が西扉の方から聞こえる



状態・行動・報酬



強化学習

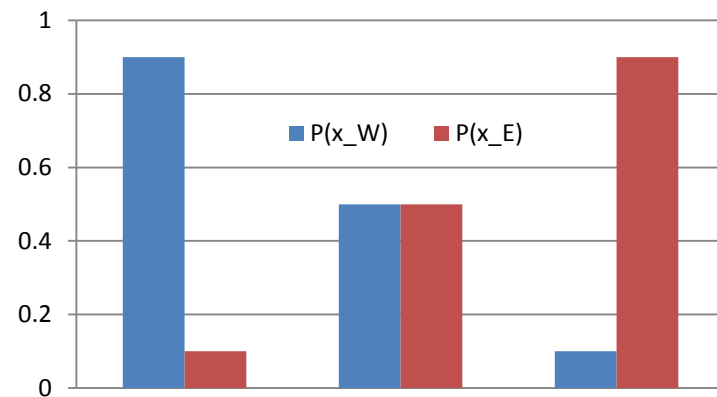


学習前の行動選択

u_E 0.3 0.3 0.3

u_K 0.3 0.3 0.3

u_W 0.3 0.3 0.3



学習後の行動選択

u_E 0.9 0.0 0.0

u_K 0.0 0.9 0.0

u_W 0.0 0.0 0.9

自動運転の複雑さ

低

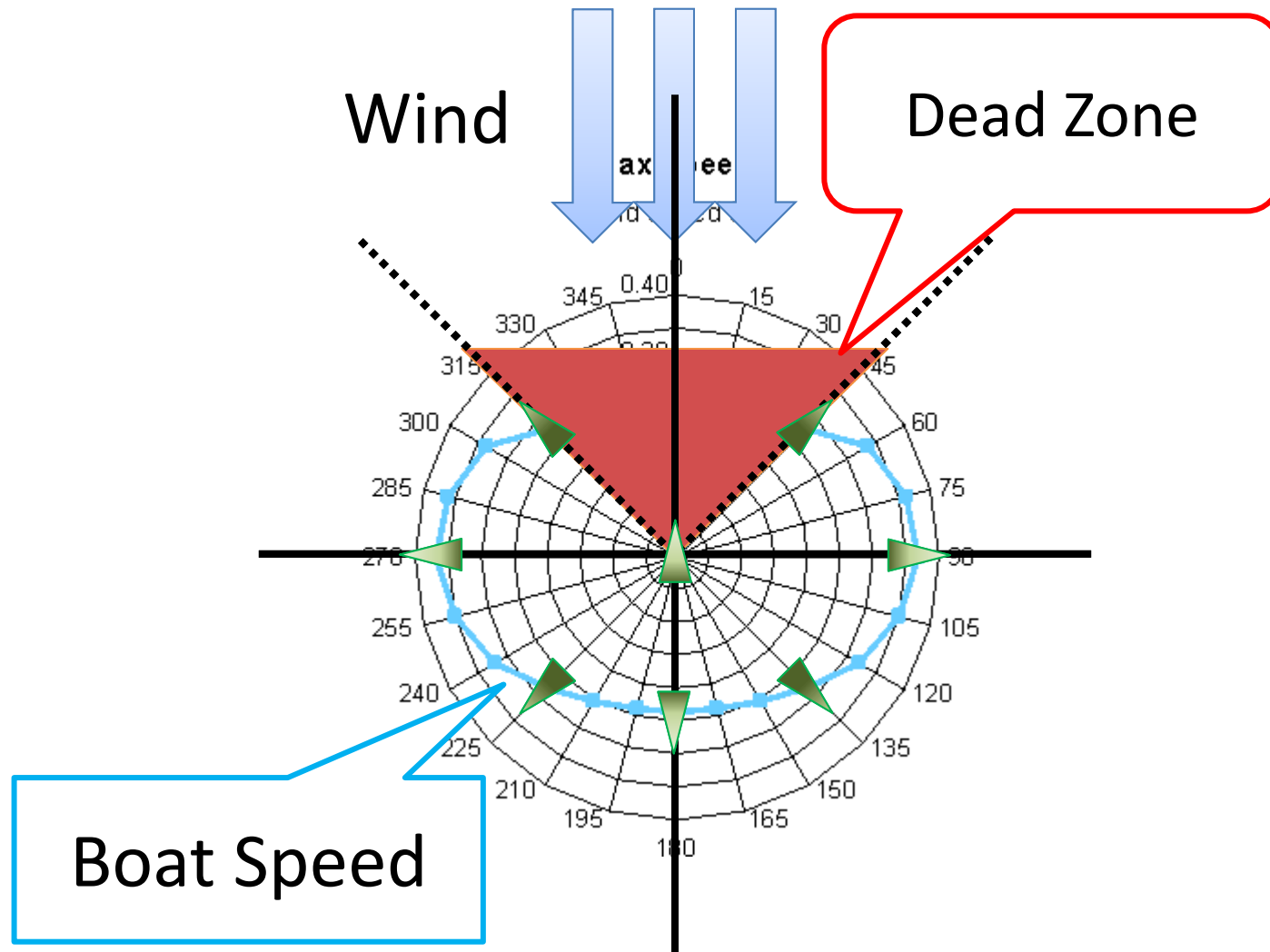
高



動力を制御できる

動力の風を
制御できない

Wind direction and Sailboat speed



I. Head for the favored side

Boat A started near the 'wrong' end and sailed away from the expected shift for a while, so she will end up near the back of the fleet.

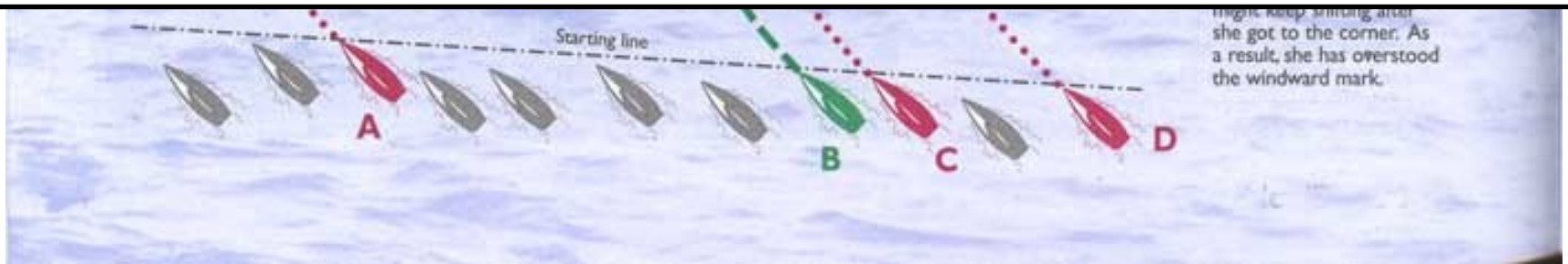


Boat B may have started a little too far down the line, but she was able to tack early and head to the right. Though B initially passed behind Boat D, she caught her by sailing farther into the persistent shift. B also gained on Boat C by tacking shy of the layline so she wouldn't over-

Targeting North mark,

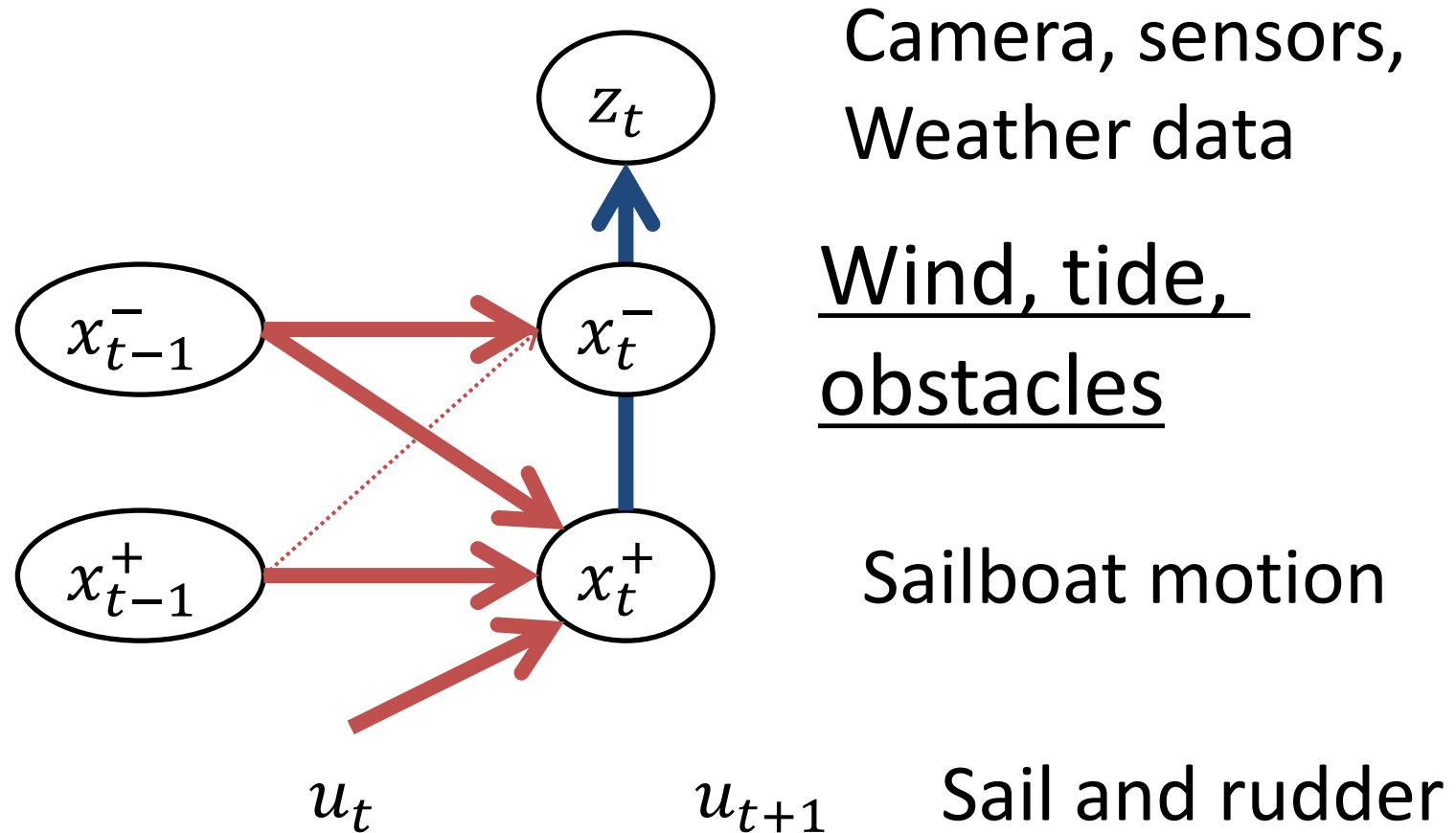
Wind is expected to shift from North to NNE.

Where should you start?

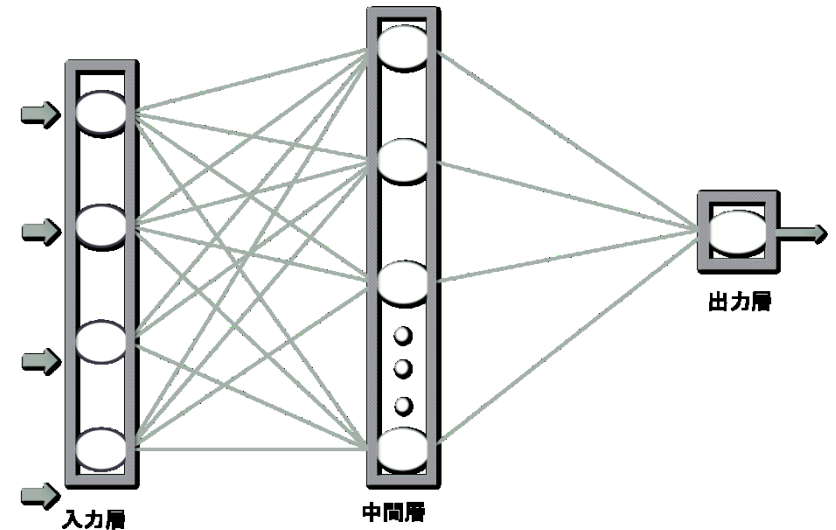
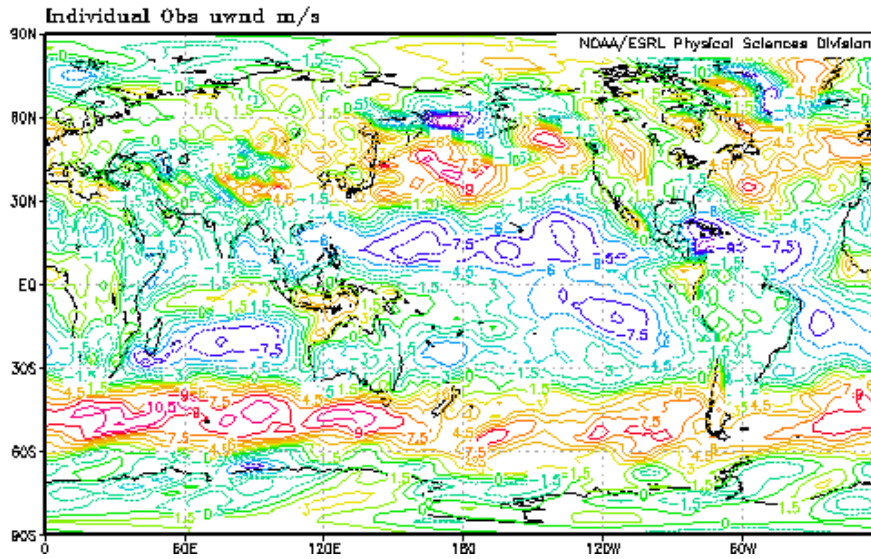


might keep sailing until she got to the corner. As a result, she has overstood the windward mark.

Uncontrollable state variables x_k^-



複素ニューラルネットワークによる 風の予測

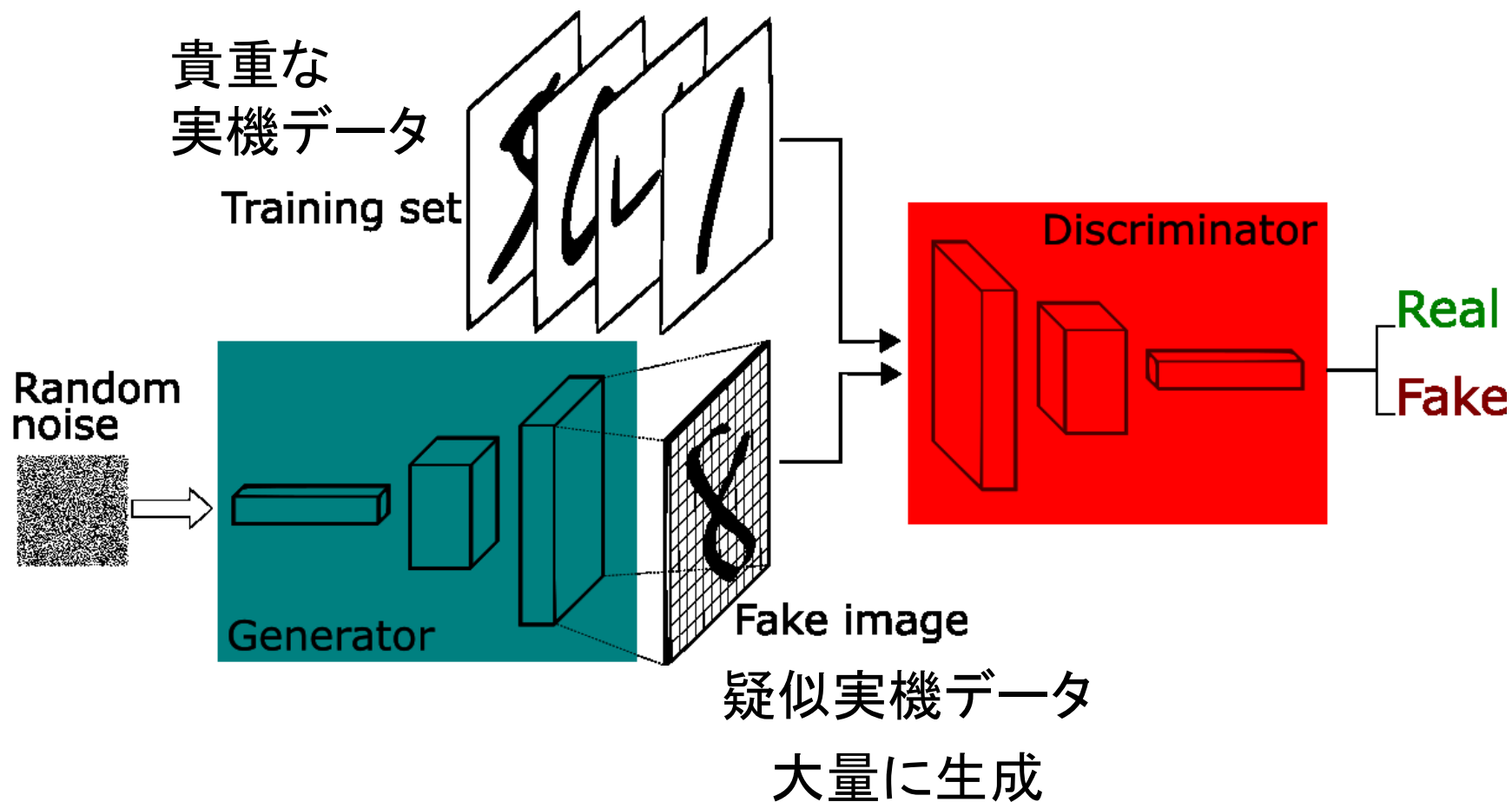


- 予測結果は、
- 次々回5月15日のセミナーをお楽しみに！

さらに大きな面白い問題！

- 実機データはなかなか取得できない
 - ハードウェア準備
 - ソフトウェア準備
 - 湖までのアクセス（大雪で渋滞、公園封鎖）
 - 気象（良い風が吹くか）
- 実機データと見分けがつかない疑似データを生成！（Generative Adversarial Network）

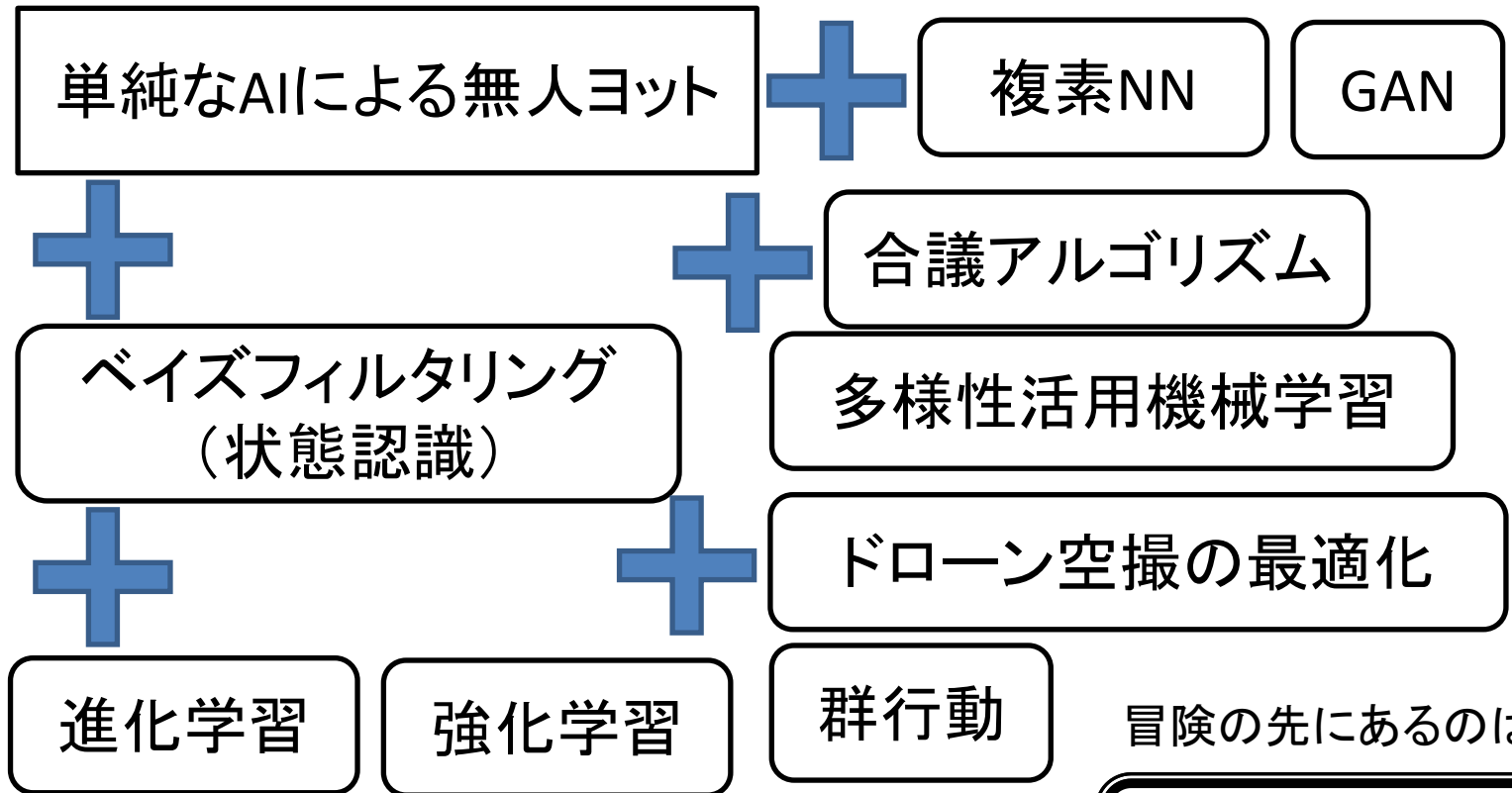
GANによるデータ拡張



さらにさらに面白いのは

- 固定の障害物回避
- 動く障害物（人やロボが操縦する船舶）回避
→「心の理論」、ゲーム理論、不完全情報
ゲーム
- 次回、次々回のセミナーをお楽しみに！

壮大な冒険の続きをつくるのは...



• あなたです！

“知能”の理解

最終発表会課題

- 5月22日（当日の指摘を受けて6月5日）
- 内容：
- ベイズフィルタ（カルマンフィルタ、パーティクルフィルタ）や機械学習...について調べ、面白い応用先について実装してくる。
- 検索ワード：“scikit-learn”，
“UCI Machine Learning Repository”
- 実装の目的（どこが面白いか）、概要、結果
- & 将来の展望：実装の技術を発展させて
東京オリパラ2020でどんな貢献ができるか