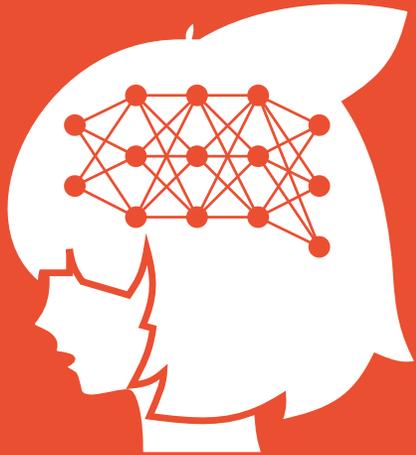


VRChat自律機械知能プロジェクト 中間発表



ML Shukai

げそん <GesonAnko>

ocha_krg

myxy

Earl Klutz (クルツ)

最初のお願い

- 本LTは視聴している皆さんを**Critic(直訳:映画の評論家)**と想定して作られています。
- **面白そうな点**や**よくわからない点**、**突っ込みどころ**など是非考えながら視聴してみてください！
そして後で**ツッコミ入れて**ください！

最初のお願い

- 本LTは視聴している皆さんを**Critic(直訳:映画の評論家)**と想定して作られています。

- **面白そうな点**や**よくわからない点**、**突っ込みどころ**など

是非考えながら視聴してください。

そして後で**ツッコミ人**。

でも素人質問は勘弁な！

VRChat自律機械知能プロジェクトとは

- VRChat上に自律機械知能を生み出し、その振る舞いを記録・解析することにより人間及び知能全体への理解を深める。
- また、それらを論文にまとめバ学会にて発表する。

第1目標

- **原始自律機械知能の作成とVRChat上への実装**
- 要件
- VRChat上で動作する原始自律機械知能の作成

第1目標

- **原始自律機械知能の作成とVRChat上への実装**
- 要件
- VRChat上で動作する原始自律機械知能の作成

やったぜ！

完了

第2目標

- **実装した原始自律機械知能を学習させVRChat上でワールド内の探査を行えるようにする。**
- 要件
- VRChat上で動作する原始自律機械知能の作成
- 原始自律機械知能を観察・分析するためのワールド作成

自律機械知能とは何でしょうか？

自律機械知能とは

- **自律的に動作する機械知能を持つ自動機械**

自律機械知能とは

- **自律的に動作する機械知能を持つ自動機械**

- **自律性**(自律的)とは

人間など知的存在の直接的介入無しに、
定められた目的を達成するための行動を生成し
動作し続けるもの(強化学習における**報酬**など)

自律機械知能とは

- 自律的に動作する**機械知能**を持つ自動機械

- **機械知能**とは

学習し何らかの知識やスキルを獲得する理論

情報処理アルゴリズムと**センサー**からなる機械

例：画像認識、世界モデルなど

自律機械知能とは

- 自律的に動作する機械知能を持つ **自動機械**

- **自動機械**とは

特定の動作や作業を自動的に行う機械

行動生成アルゴリズムと**実行媒体(アクチュエータ)**からなる機械

例：組立ロボット、茶運び人形など

つまるところ、原始自律機械とは？

最低限以下の4(5)つを持った自律的に動作する機械

- センサー
- 情報処理モデル
- アクチュエーター
- 行動アルゴリズム
- (自律性を生み出すための報酬(後ほど詳しく説明))

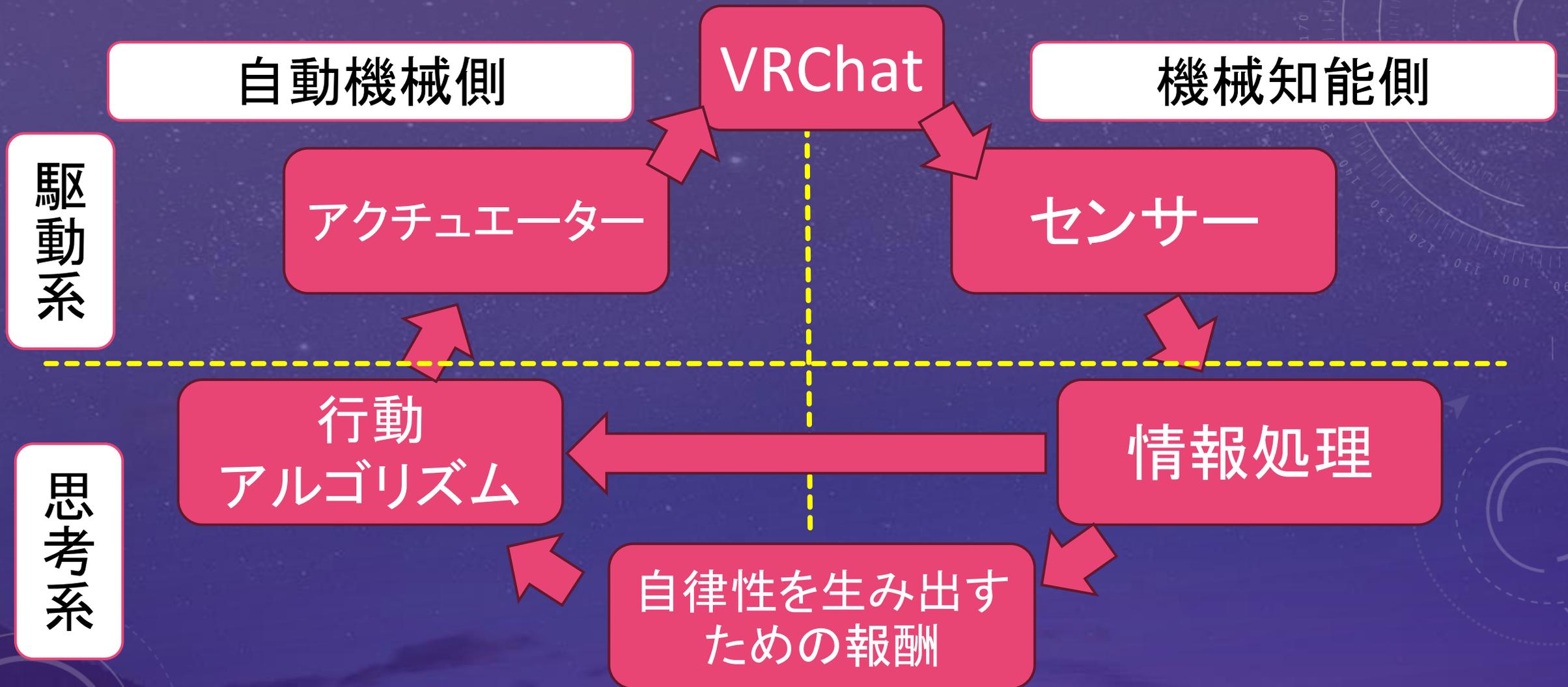
つまるところ、原始自律機械とは？

最低限以下の4(5)つを持った自律的に動作する機械

- センサー
- 情報処理モデル
- アクチュエーター
- 行動アルゴリズム
- (自律性を生み出すための報酬(後ほど詳しく説明))

これはなくても良いらしい
何でだろう？

VRChat上で動作する原始自律機械知能



VRChat上で動作する原始自律機

今回は
駆動系を環境(Environment)
として実装

自動機械側

VRChat

環境
(駆動系)

アクチュエーター

センサー

エージェント
(思考系)

行動
アルゴリズム

情報処理

自律性を生み出す
ための報酬

思考系をエージェント(Agent)
として実装

VRChat上で動作する原始自律機

自動機械側

VRChat

今回は
駆動系を環境(Environment)
として実装

環境
(駆動系)

アクチュエーター

Agentから見ると
駆動系はただの
世界環境に見えるように設計

サー

エージェント
(思考系)

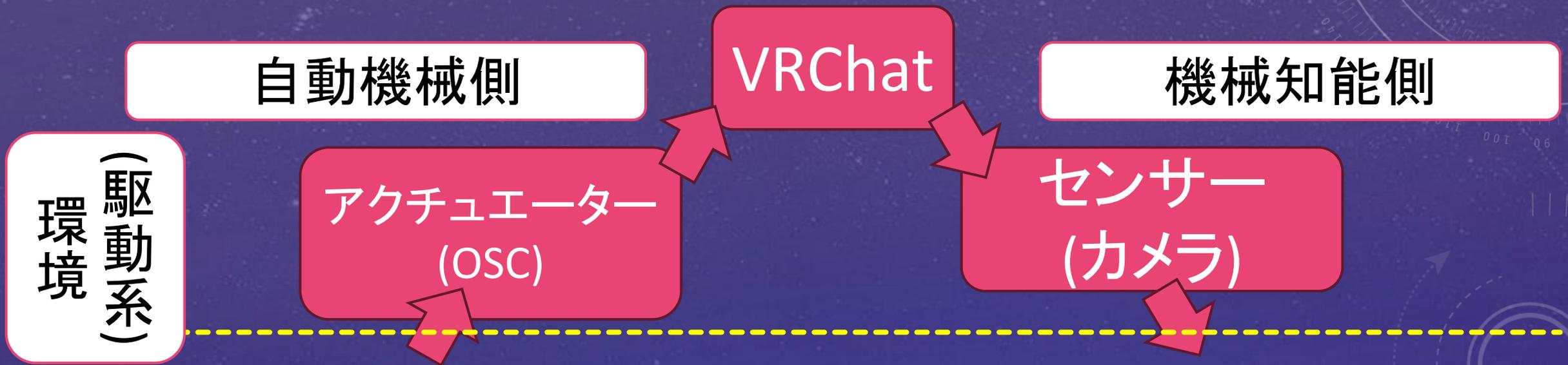
行動
アルゴリズム

情報処理

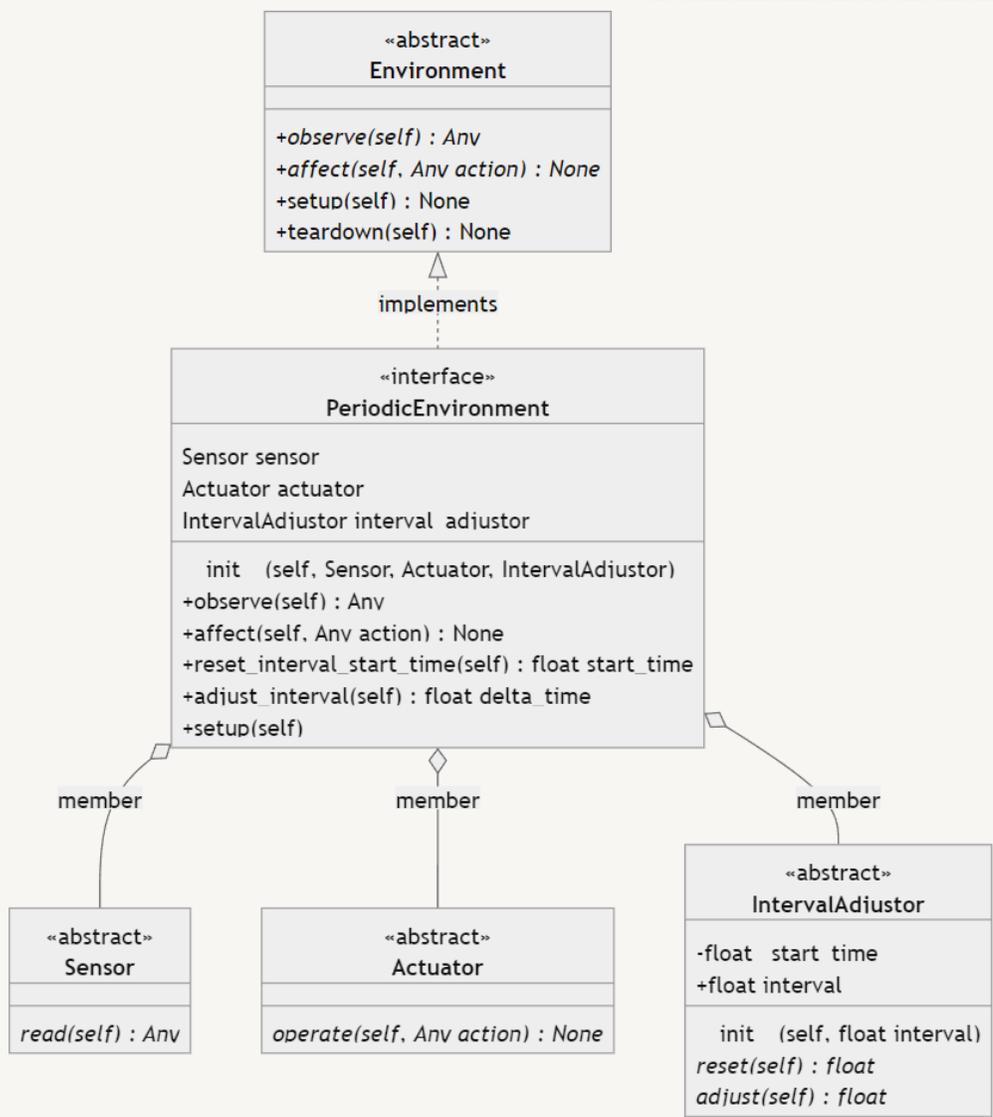
自律性を生み出す
ための報酬

思考系をエージェント(Agent)
として実装

Environmentクラスは単純(VRChat I/O)



Enviromentクラス図



- 環境とのやり取りの手続きを制定するクラス。

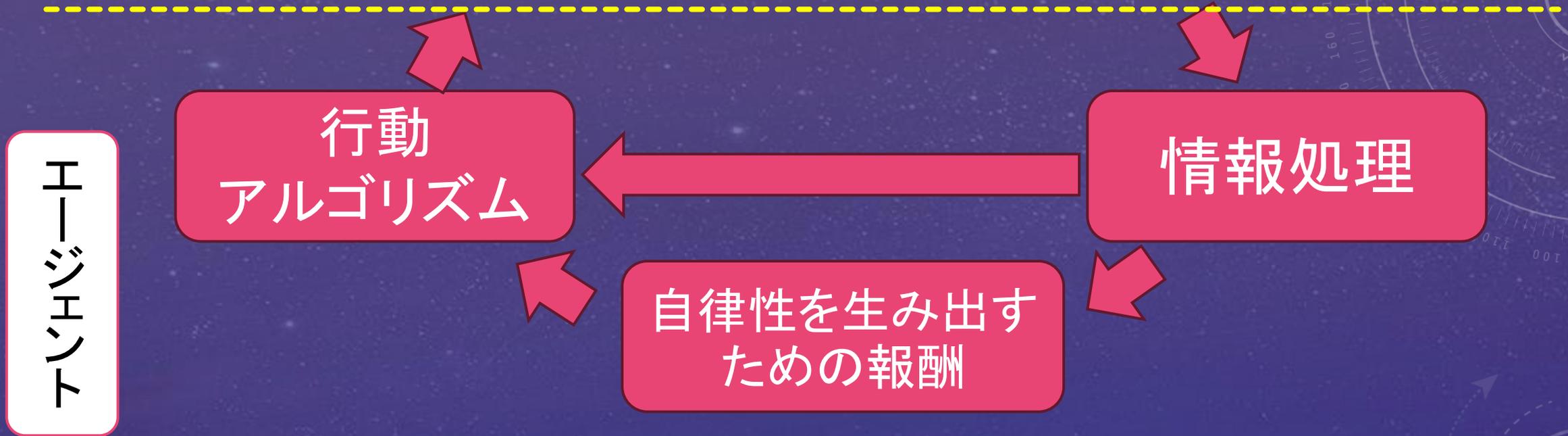
- sensor(センサー)

- Actuator(アクチュエーター)

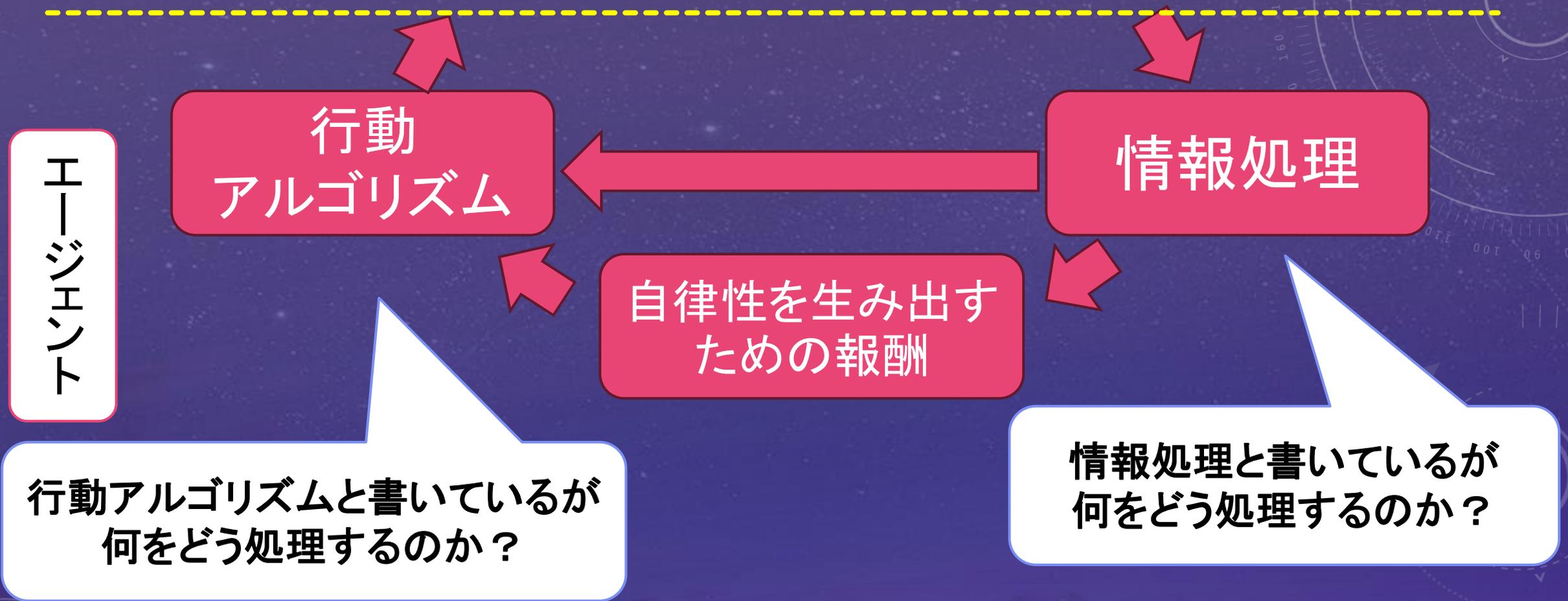
- InternalAdjustor

周期的な時間スケールを実装するための
帳尻合わせ機能

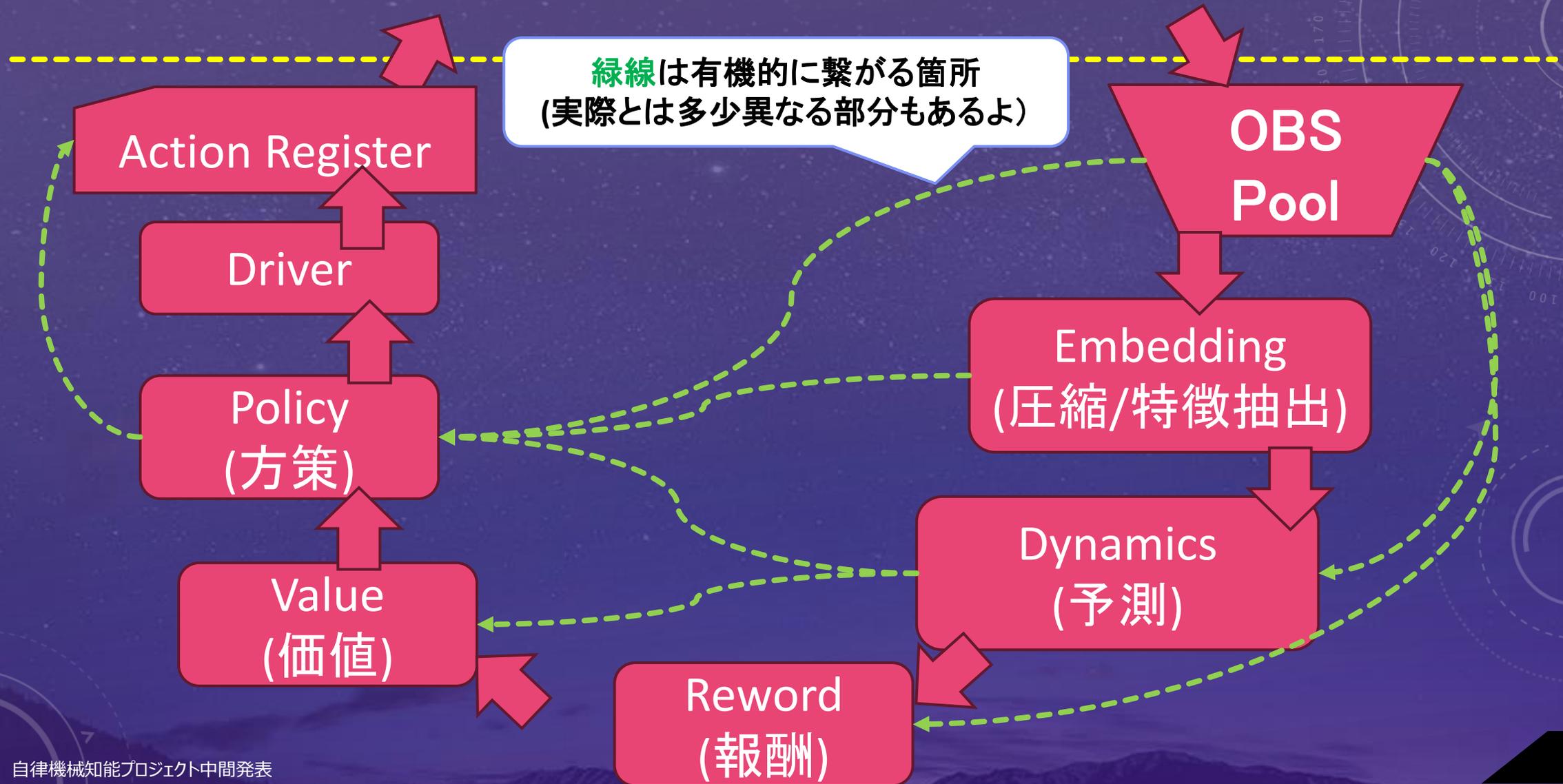
問題はエージェント、具体性に欠ける



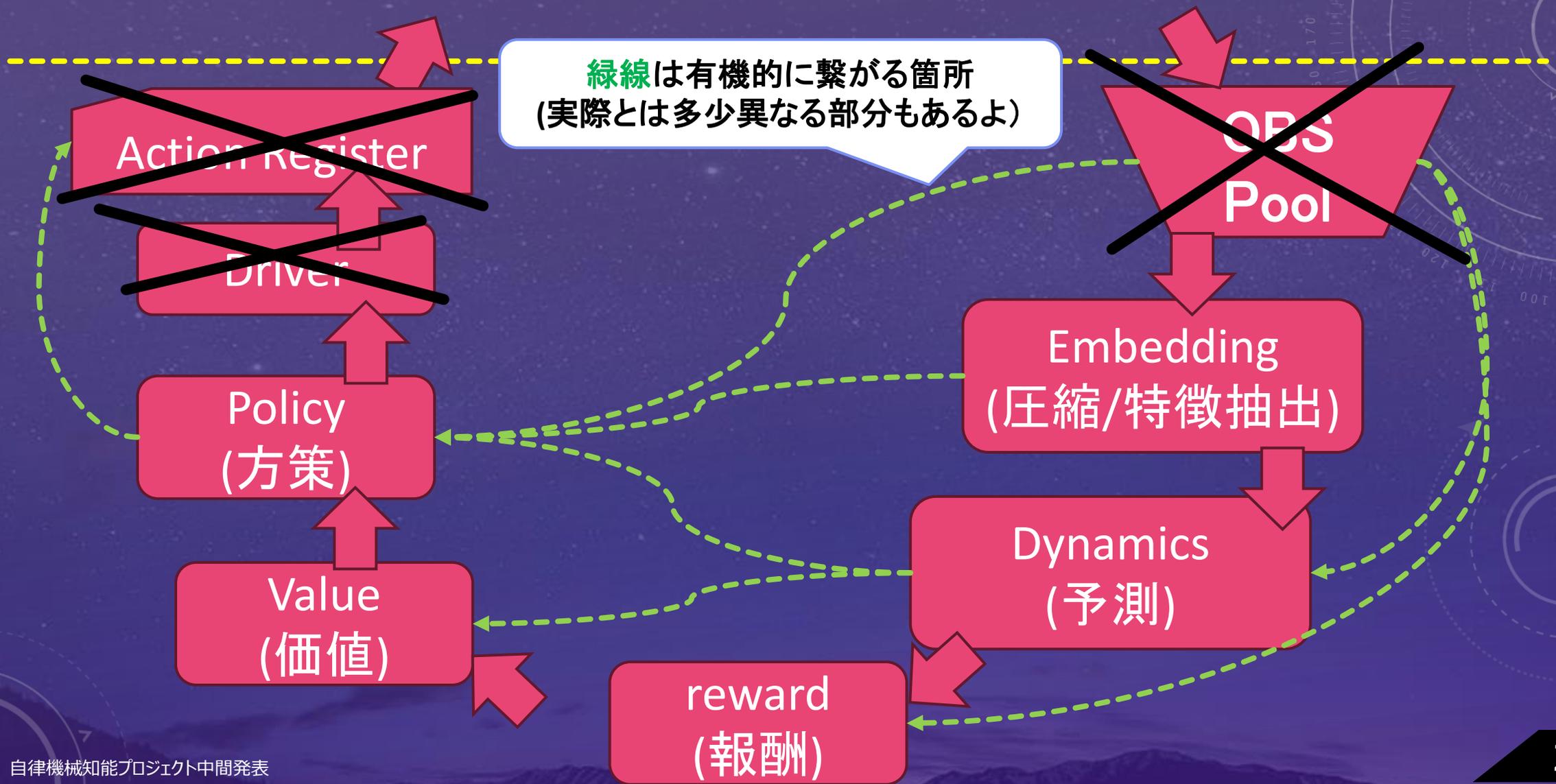
問題はエージェント、具体性に欠ける



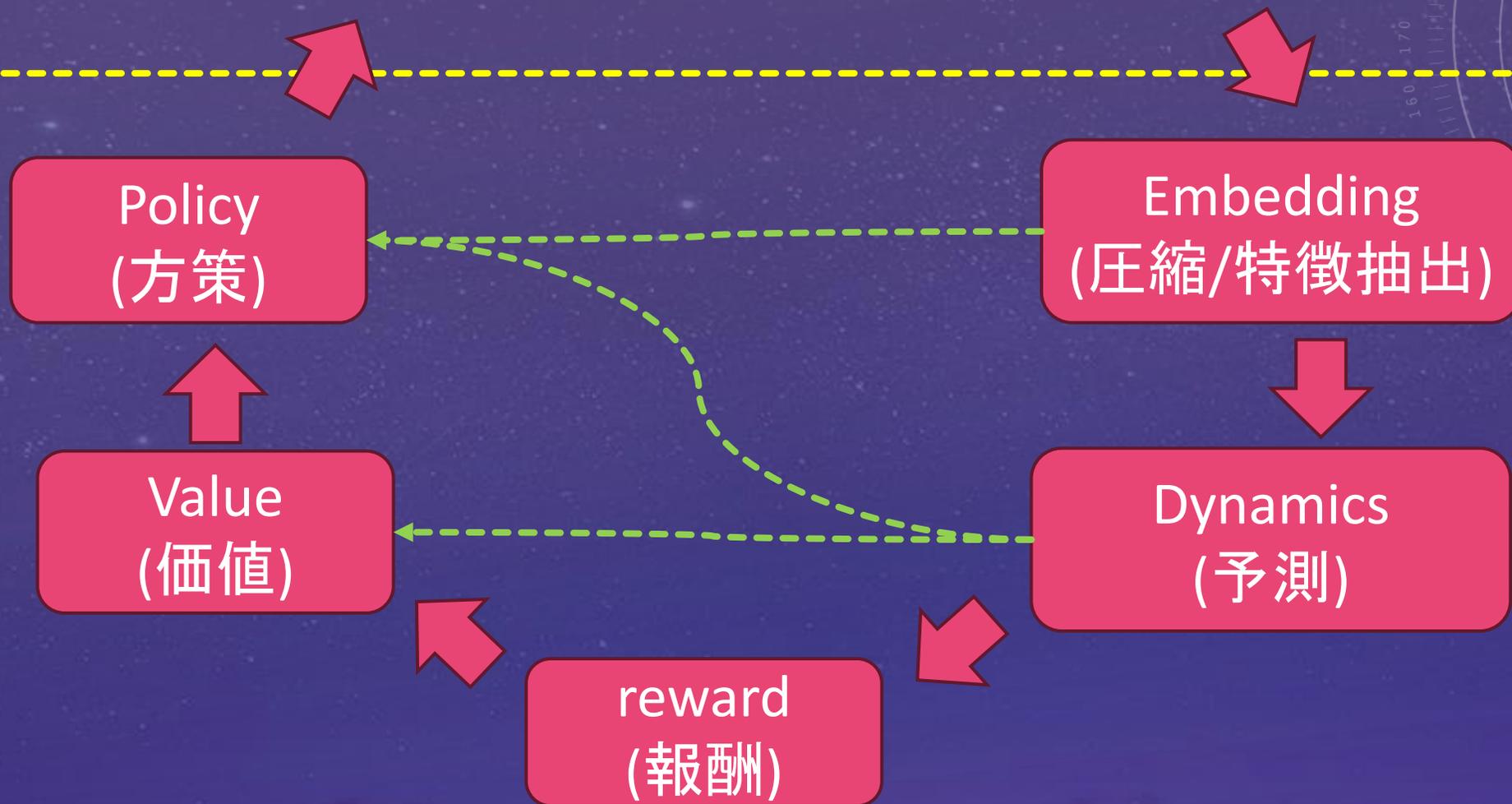
自律機械知能のエージェントをより具体的に図式化



ややこしいので一旦最低限の原始的自律機械知能に



原始自律機械知能のエージェントを図式化



原始自律機械知能のエージェントを図式化



これらをそれぞれ学習させていけば自律機械知能になる

Embedding



- 単純にカメラの観測を何らかの特徴量(埋め込み表現)に圧縮するアルゴリズム
- 基本的なCNN構造で設計

Dynamics



- 前方策、埋め込み表現、方策などから次の埋め込み表現を予測するアルゴリズム(Forward Dynamics)

PolicyとValue



• 強化学習の報酬

報酬に基づいて行動する際の価値と次の方策(行動)を計算するアルゴリズム

今回は**PPO(Proximal Policy Optimization, 2017)**を採用

Reward



• 強化学習の報酬

エージェントがある行動を選択し、その行動によって環境がどのように反応するかを評価するための指標。

ゲームなどの場合は**スコア**などになるがVRChatの場合**明確な指標はない**ため今回は**好奇心ベースの内発的報酬**を採用

自律性を生み出すための報酬の種類

- **好奇心ベース**

内発的報酬で多くの現実的な環境設定で実装可能

- **エネルギーベース(食欲など)**

外発的動機付けで物理的な体を持つ場合実装可能

- **社会性ベース**

外発的動機付け、現状実装不能

- **その他外部報酬**(一般的な強化学習タスクはこれ)

好奇心ベース(CURIOSITY)の内発的報酬とは

myxyさんの資料より引用

強化学習

環境とエージェントのインタラクションにより学習する

- 環境に対してエージェントは**アクション**を行う
- エージェントは環境から**報酬**を得る

最終的に得られる報酬を最大化するのが目的



強化学習の課題

一般に環境から得られる報酬は**粗もしくはは存在しない**場合
報酬がほとんど得られない環境では、エージェントは報酬を得られない
= 学習ができないままほぼランダム(未学習の方策)で動くことになる。



環境から報酬(外発的報酬)が得られない場合

- エージェント自体の内部で内発的報酬による学習が必要

- 新奇性の高い観測をする行動に対して報酬を与える

(Bellemare et al., 2016; Lope et al., 2012 Poupart et al., 2006)

- 環境の変化が予測しやすい行動に対して報酬を与える

(Houthoof et al., 2016; Mohamed & Rezende, 2015; Schmidhuber, 1991; 2010; Singh et al., 2005; Stadie et al., 2015)

など方法があり前者を実現する機構を**Curiosity(好奇心)**と呼んでいます。

Curiosity(基本)

- 現在時間 t の観測 O^t とアクション a^t 、および次フレームの観測 O_{t+1} から次フレームの $t+1$ の観測予測 $O_{t+1} = F(O_t, a_t)$ を得る順方向モデル F を学習
- 次フレーム $t+1$ の実際の観測 O_{t+1} と \hat{O}_{t+1} の誤差を報酬とする

Curiosity(基本) 概略図

Observation

観測 O_t

Action
行動 a_t

**Forward
Model**

Prediction

予測 \hat{O}_{t+1}

Observation

観測 O_{t+1}

\ominus

Reward

報酬 R_t

Curiosity(基本) 概略図

Observation

観測 O_t

Action
行動 a_t

**Forward
Model**

Prediction

予測 \hat{O}_{t+1}

Observation

観測 O_{t+1}



Reward
報酬 R_t

予測と実際の観測が違えば違うほど
大きな報酬となる考え方だよ！

Curiosityの課題(摂動問題)

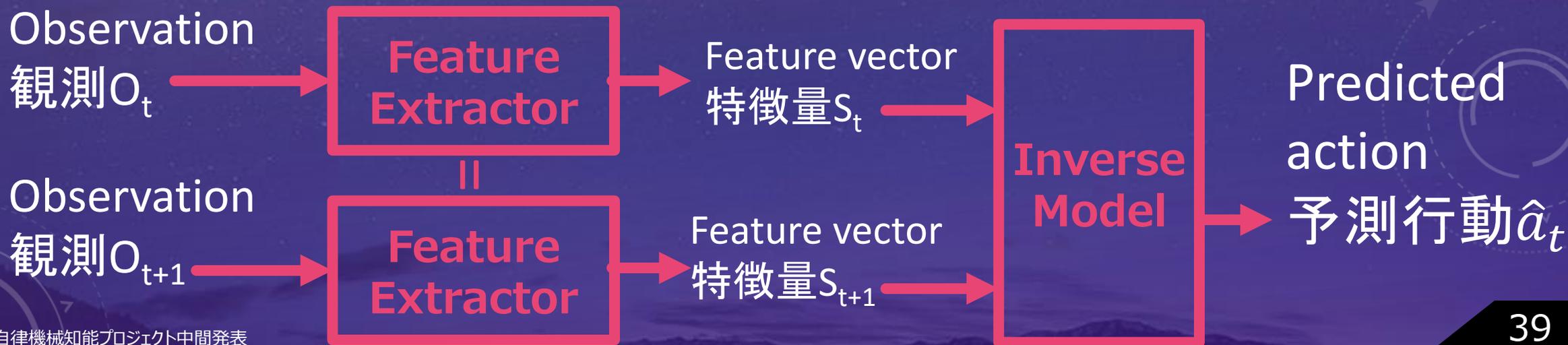
環境における**本質的に予測困難な摂動**に対して報酬を与えてしまう問題がある。

- 砂嵐の映るテレビ
- 木の葉の動き
- 揺れる水面

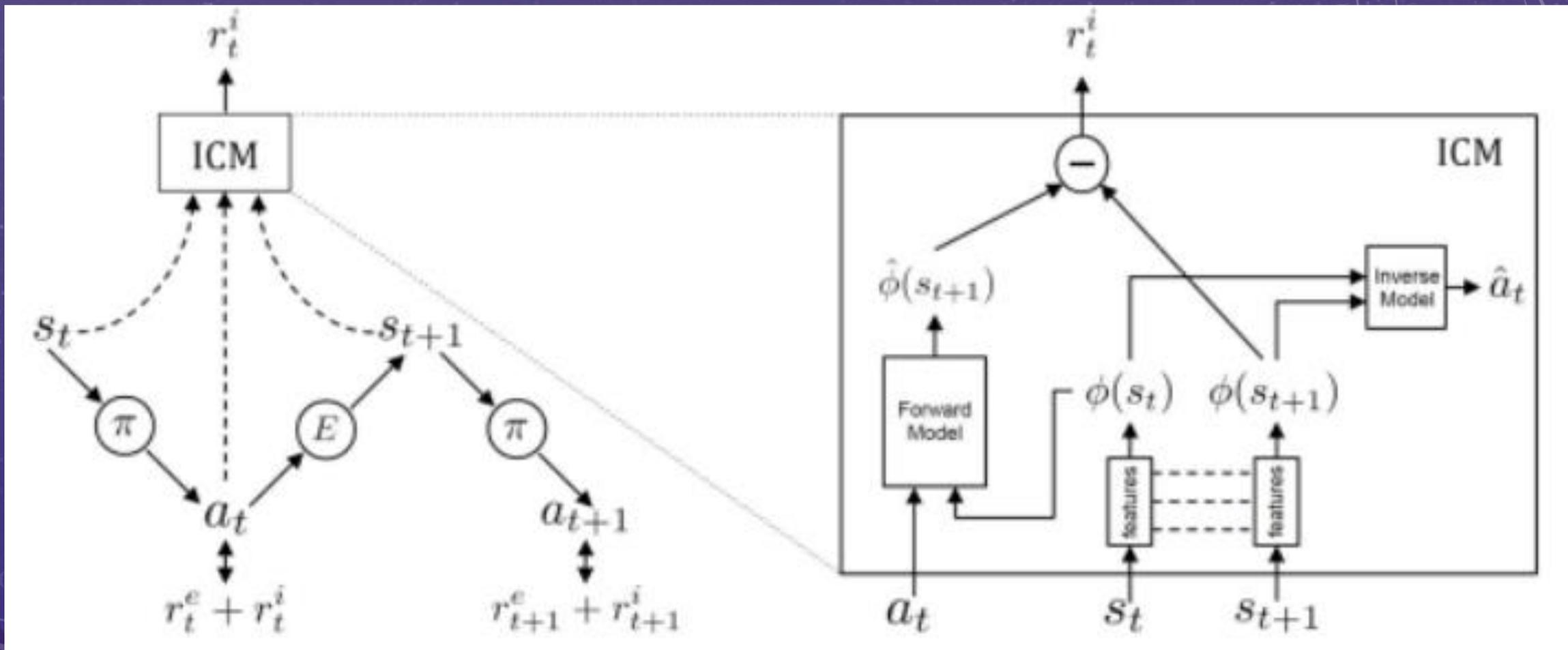
等の予測困難な摂動は実環境にあふれている。

解決策

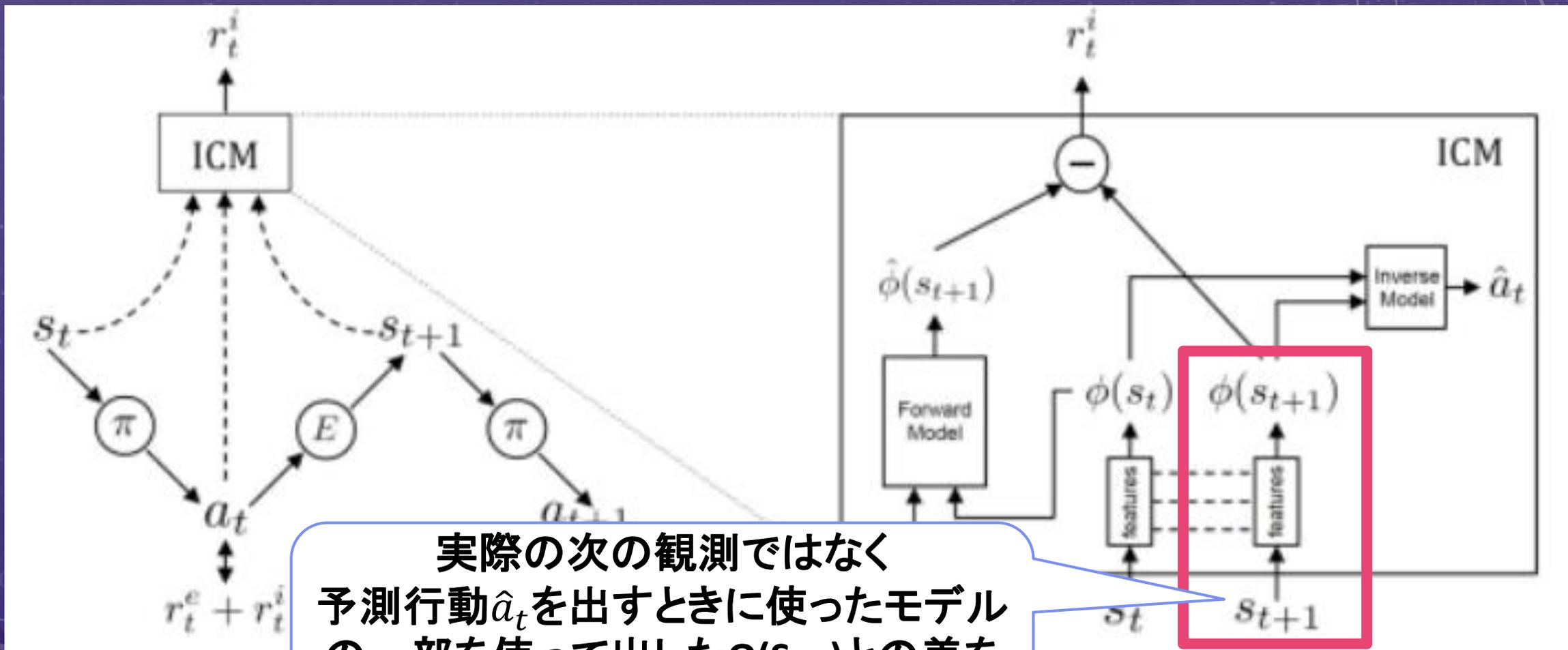
- “Curiosity-driven Exploration by Self-supervised Prediction”では
- 観測 O_t, O_{t+1} から行ったアクション a_t を予測する**逆推定モデルを先に学習させ**そのその中間層に本質的な特徴が学習されているとして観測の情報から**擾動に依存しないアクションに関する情報を抽出している。**



“Curiosity-driven Exploration by Self-supervised Prediction” モデル全体図



“Curiosity-driven Exploration by Self-supervised Prediction” モデル全体図



実際の次の観測ではなく
予測行動 \hat{a}_t を出すときに使ったモデル
の一部を使って出した $\hat{\phi}(s_{t+1})$ との差を
報酬にするよ

原始自律機械知能のエージェントを図式化



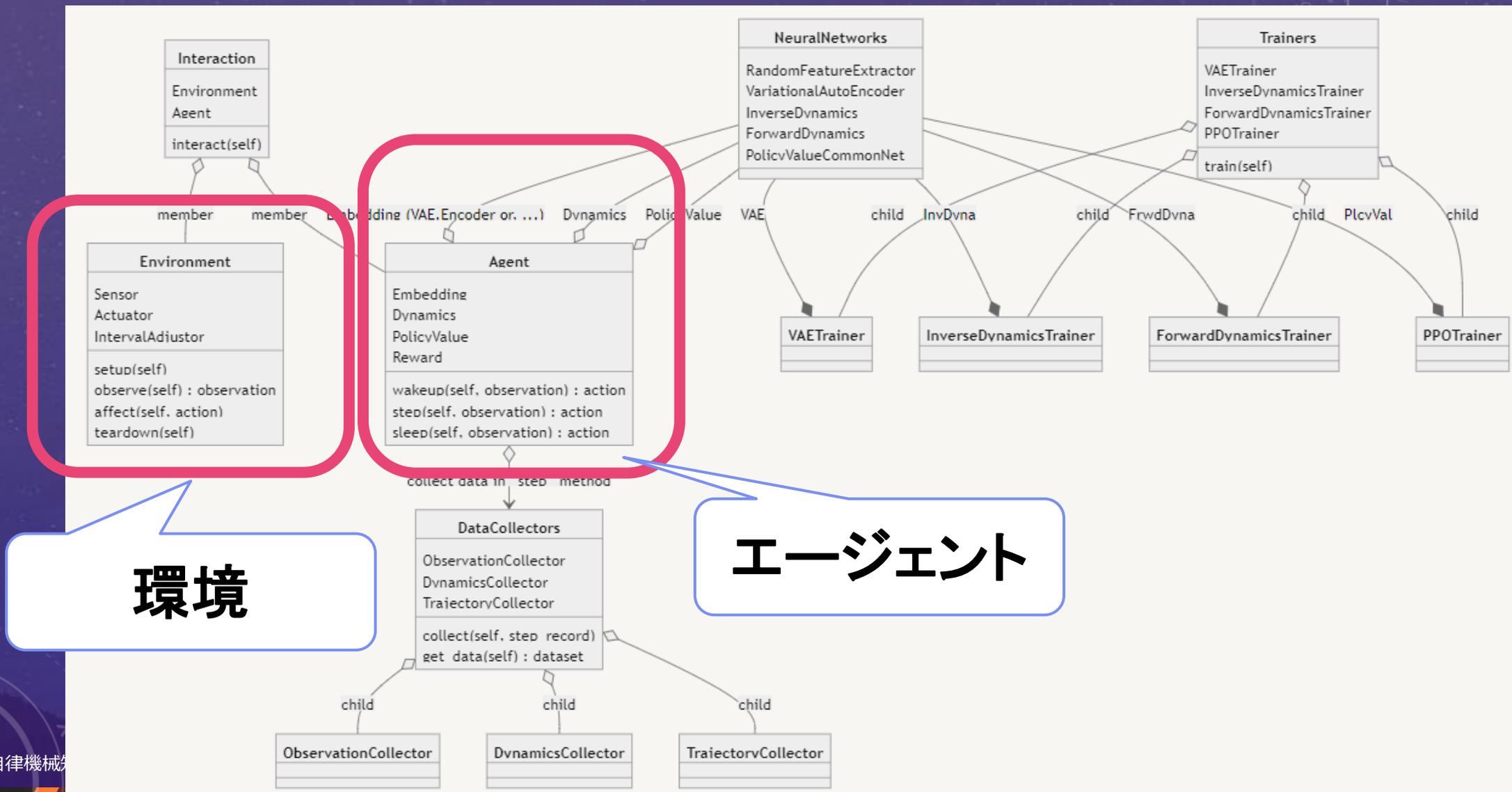
というわけで報酬は
好奇心ベースのCuriosityで
実装

原始自律機械知能のエージェントを図式化

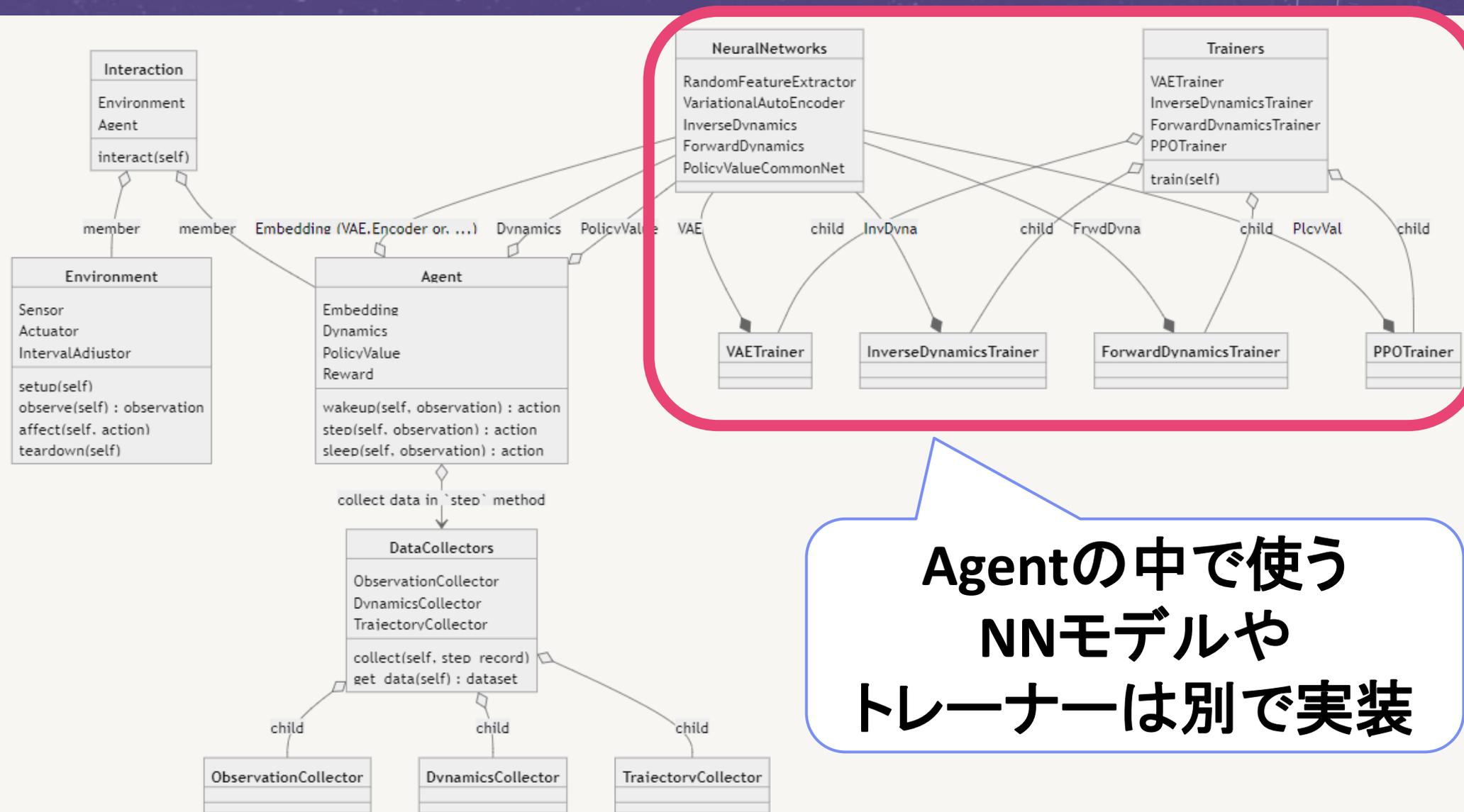


ただし、学習した後はNNのモデルの一部しか使わないためNNはそれぞれAgentから独立して実装。

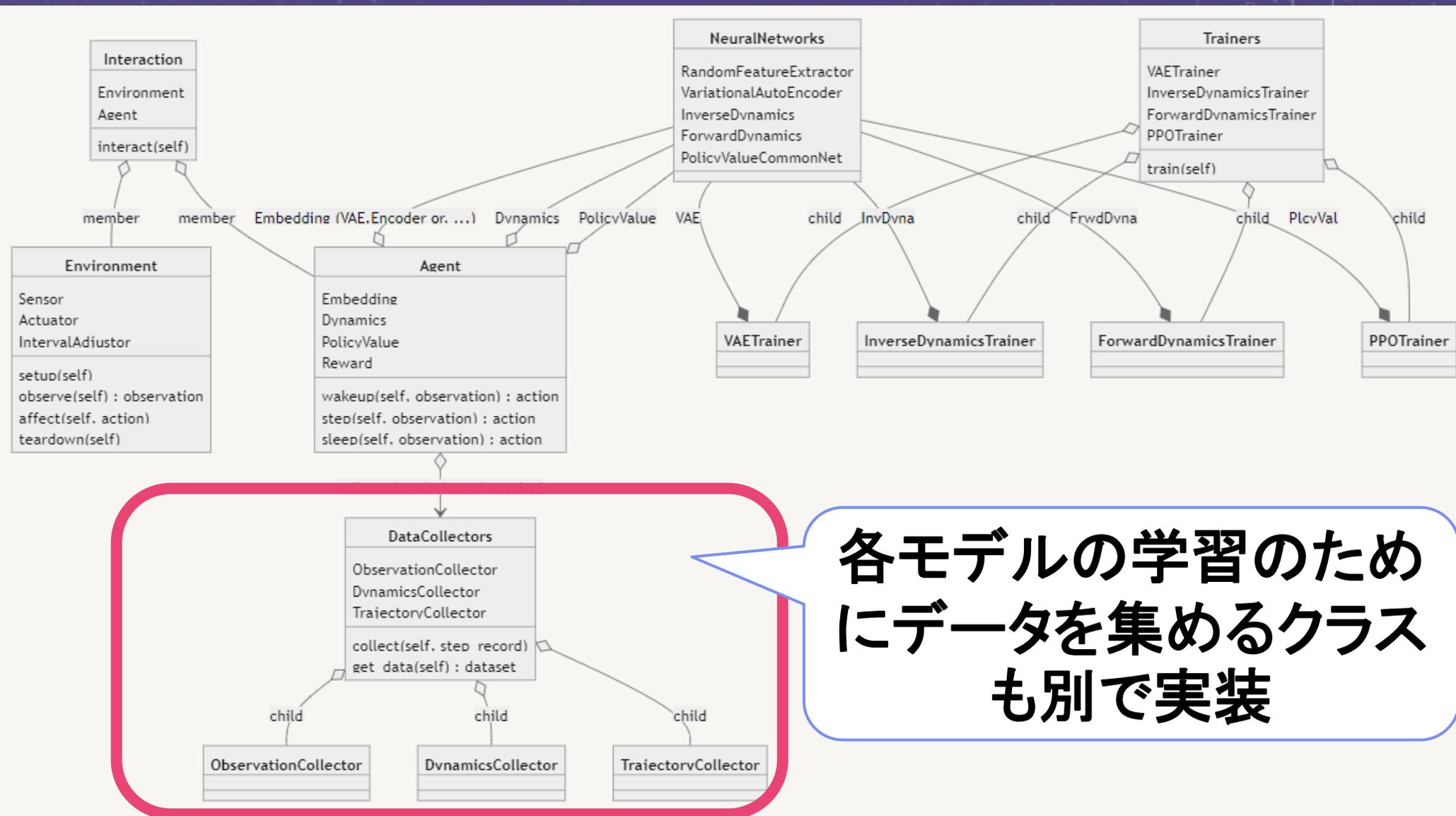
全体関係図



全体関係図



全体関係図



各モデルの学習のためにデータを集めるクラスも別で実装

使用する学習モデル一覧

- FeatureExtractor
- Inverse Dynamics
- Variational AutoEncoder (VAE)
- One Step Forward Dynamics
- PPO Policy (PolicyValueCommonNet)

現状出来ることとこれから出来るようになること

- VRChat上にとりあえず**爆誕(できた)**
- VRChatの**ワールドの探索(これから学習！)**

今後の予定

- **原始自律機械知能の学習**
- **原始自律機械知能の拡張**
- VRChat内位置情報取得ツールYAIIBAを使用し、
ワールドの探査を動きの分析
- そのためのワールド作成

自律機械知能の可能性

- VRChatで移動以外の動作の実装や学習。(物をつかむなど)
- ほかのアバターや声に対しての反応するようになるなどの拡張。

↓ (将来的に出来るかは不明ですが)

- ゲームワールドの対戦相手やギミック作動の操作役
- 中の人不在時の留守番役など

自律機械知能の可能性

- VRChatで移動以外の操作
- ほかのアバターや声に対

なんてのができると良いな
と思いました！（個人の感想です）
（おど）
緊張。

↓（将来的に出来るかは不明ですが）

- ゲームワールドの対戦相手やギミック作動の操作役
- 中の人不在時の留守番役など

自律機械知能の可能性

- VRChatで移動以外の
- ほかのアバターや声に対

なんてのができると良いな
と思いました！（個人の感想です）
（かど）
張。

↓（将来的に出来るかは不明ですが）

- ゲームワーク
- 中の人イ

動の操作役

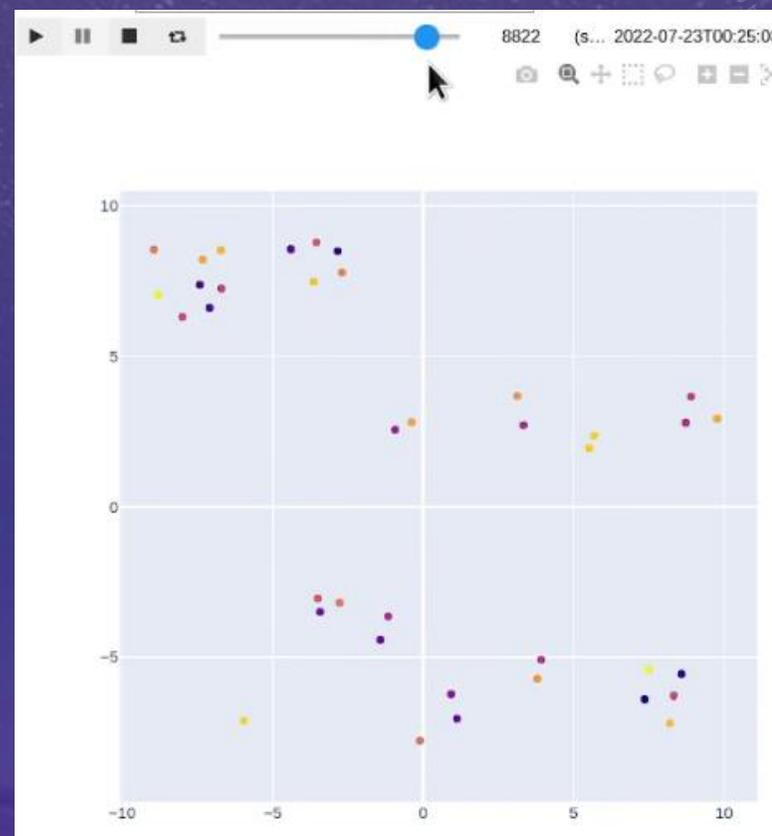
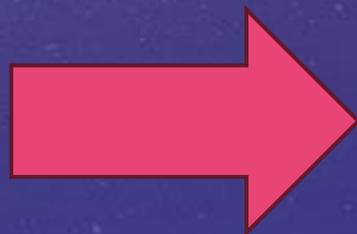
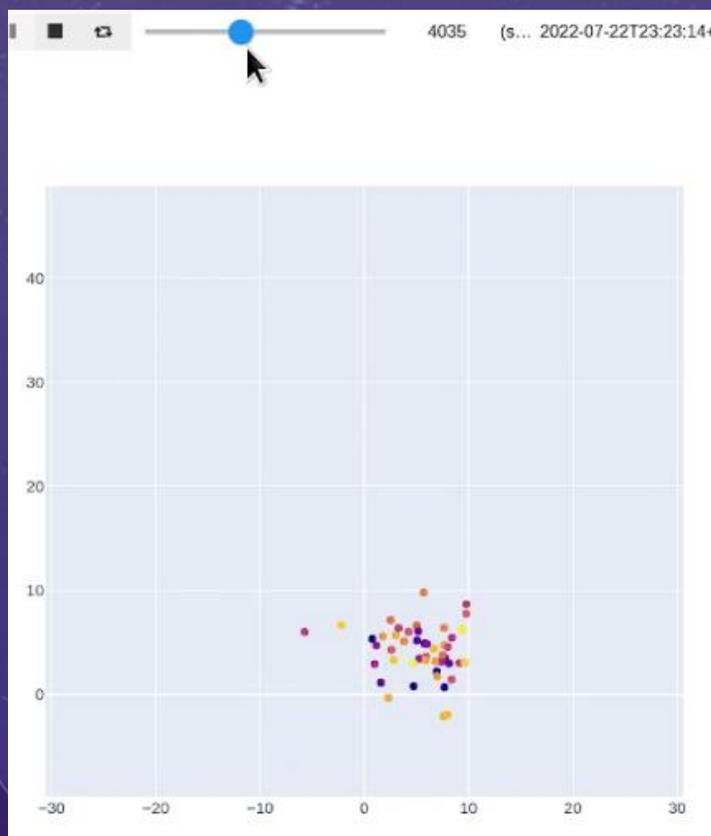
可能性は無限大！！！！

分析用のワールドについて

設計案

YAIBAについて

- ワールドに存在するアバターの位置情報を取得するツール



ワールド1

- シンプルな正方形のワールド
- 四方の壁はそれぞれ別々の色 (模様)
- 1辺概ね10m~20m程度のある程度の広さを持つ
- 壁の1面に動画が再生できるスクリーンを設置



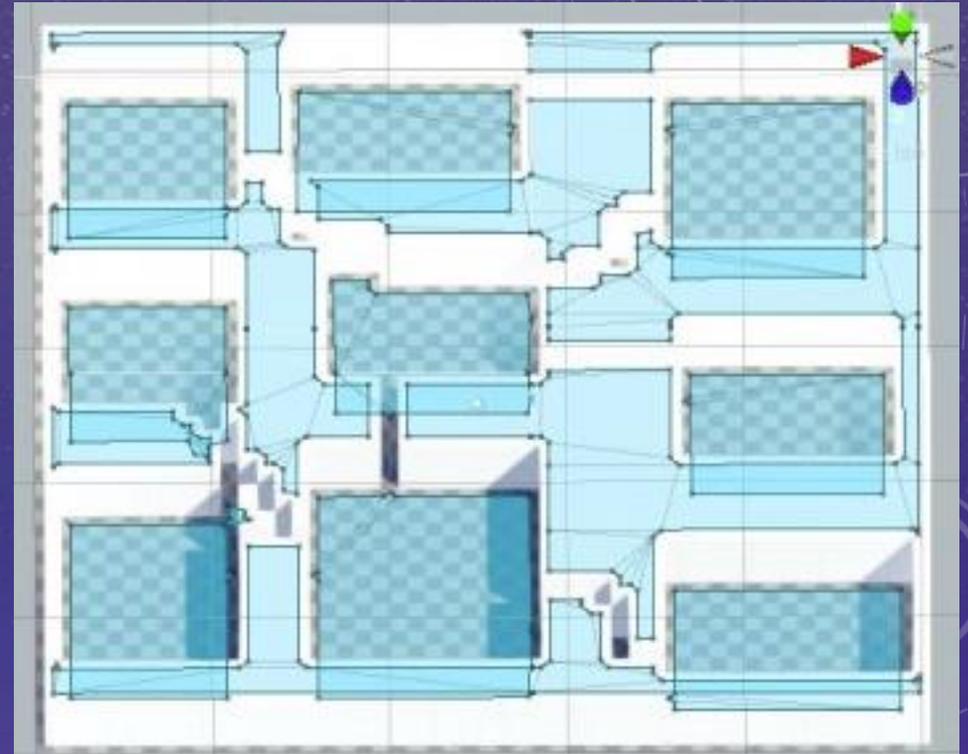
ワールド2 第1案

- 自律機械知能が探査を行うことができるワールドとして迷路のワールドとする。
- 迷路サイズはサイズ別に数パターン作成する。
- 5m四方、10m四方、20m四方など
- 途中、動画再生用のスクリーンも用意する。

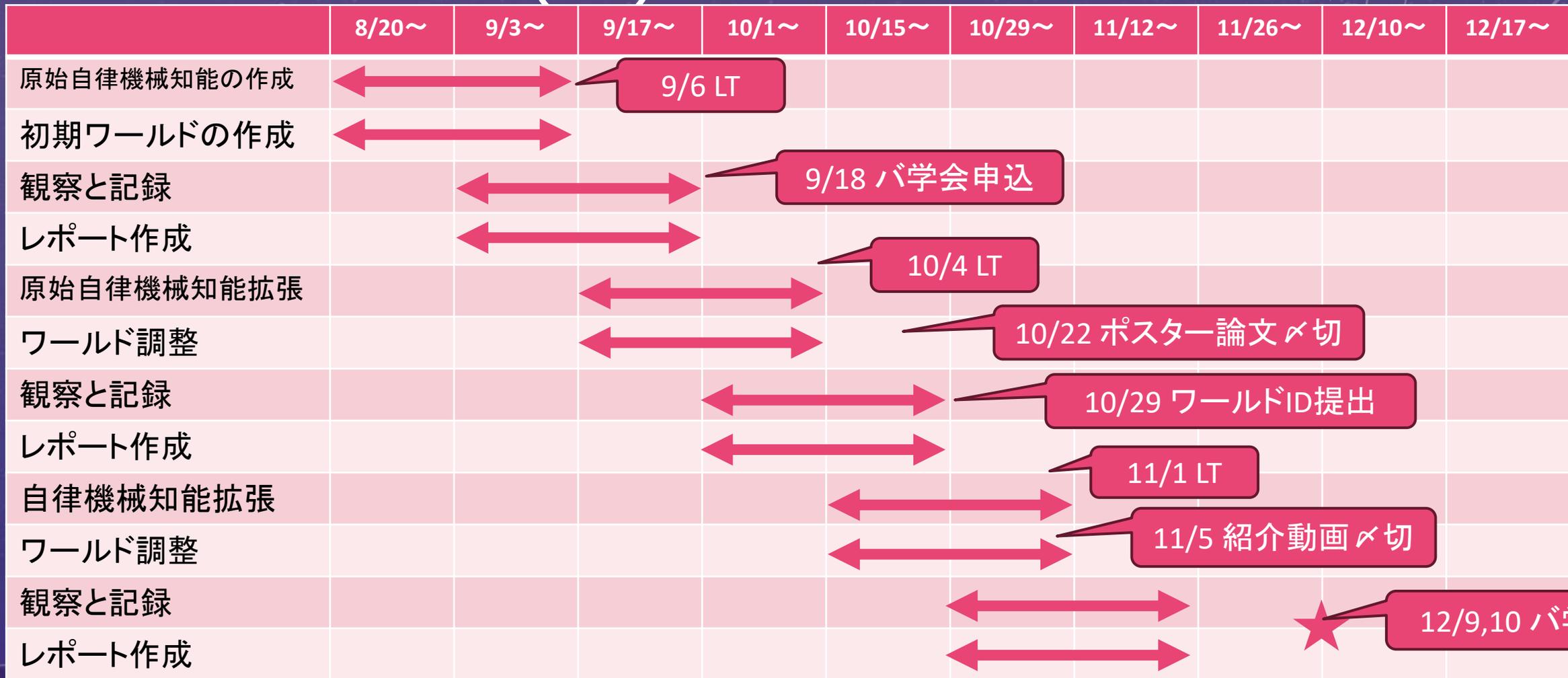


ワールド2 第2案

- 不思議なダンジョン風のワールド
- 迷路と同様にサイズ別に複数パターン作成する。
- 2部屋、4部屋、8部屋など
- 途中、動画再生用のスクリーン、家具などの移動の邪魔となるオブジェクトなども用意する。



全体スケジュール(仮)



ML集会自律機械知能プロジェクト 募集

- ML集会自律機械知能プロジェクトでは以下の点のアイデアや意見など**積極的に募集**しています。
- 自律機械知能の**拡張案やコンセプト・アイデア**
- 自律機械知能の**ふるまいの分析案**
- 自律機械知能の**検証ワールド案**

その他気づいた点など是非ご意見いただければと思います。



ありがとうございます

By ML集会

げそん<GesonAnko>

ocha_krg

myxy

Earl Klutz (クルツ)