



全脳アーキテクチャ 勉強会 #36

2022.10.5 18:00 ~ 20:40

人工知能が心をもつならば

主催：NPO法人 全脳アーキテクチャ・イニシアティブ

アジェンダ

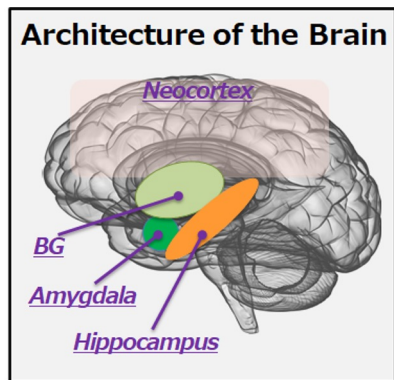
- 18:00 開会の挨拶と趣旨説明 山川宏 (WBAI)
- 18:15 感情発達ロボティクス 日永田智絵 (奈良先端科学技術大学院大学)
- 18:45 「意識の科学」のための「意識のアップロード」 或いは
「意識のアップロード」のための「意識の科学」 渡邊正峰 (東京大学)
- 19:15 休憩 (10分)
- 19:25 意識の価値 新川拓哉 (神戸大学)
- 19:55 パネルディスカッション 山川宏 (モデレーター)、日永田智絵、渡邊正峰、新川拓哉
- 20:35 Closing Remark 森岡大成 (実行委員長)
- 20:40 終了

開会の挨拶と趣旨説明

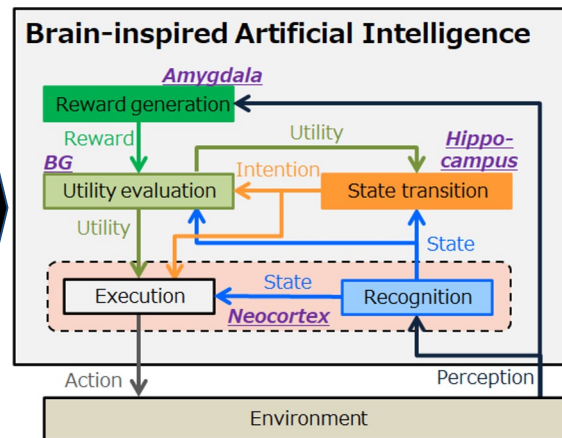
全脳アーキテクチャ・イニシアティブ

山川 宏

脳全体のアーキテクチャに学び人間のようなAGIを創る(工学)



- ① 脳の各器官を機械学習モジュールとして開発
- ② それらを統合した認知アーキテクチャの構築



本アプローチの利点

共同開発の基盤として脳の構造を利用して開発を加速

- 人間に関わる諸分野の科学知見を活用できる
- 開発の発散を防げる合意する認知アーキテクチャ
- 脳にガイドされた有機的な技術の統合

人と親和性の高いAGIを作れる

- 対人インタラクションが必要な応用
- 精神疾患などのモデル化など医療応用
- Mind-uploadingの器

全脳アーキテクチャ・イニシアティブ (WBAI)

良くも悪くも人類の福利厚生に対して大きな影響を与えるのは知能。

ヒトのあらゆる知的能力を凌駕する汎用人工知能(AGI)の実現は、もはや夢物語でない。

「**人類と調和した人工知能のある世界**」というビジョン実現のため、脳型からのAGIのオープンな開発を促進する(2030年目処)。

特定の組織がAGIを独占するリスクを抑制。

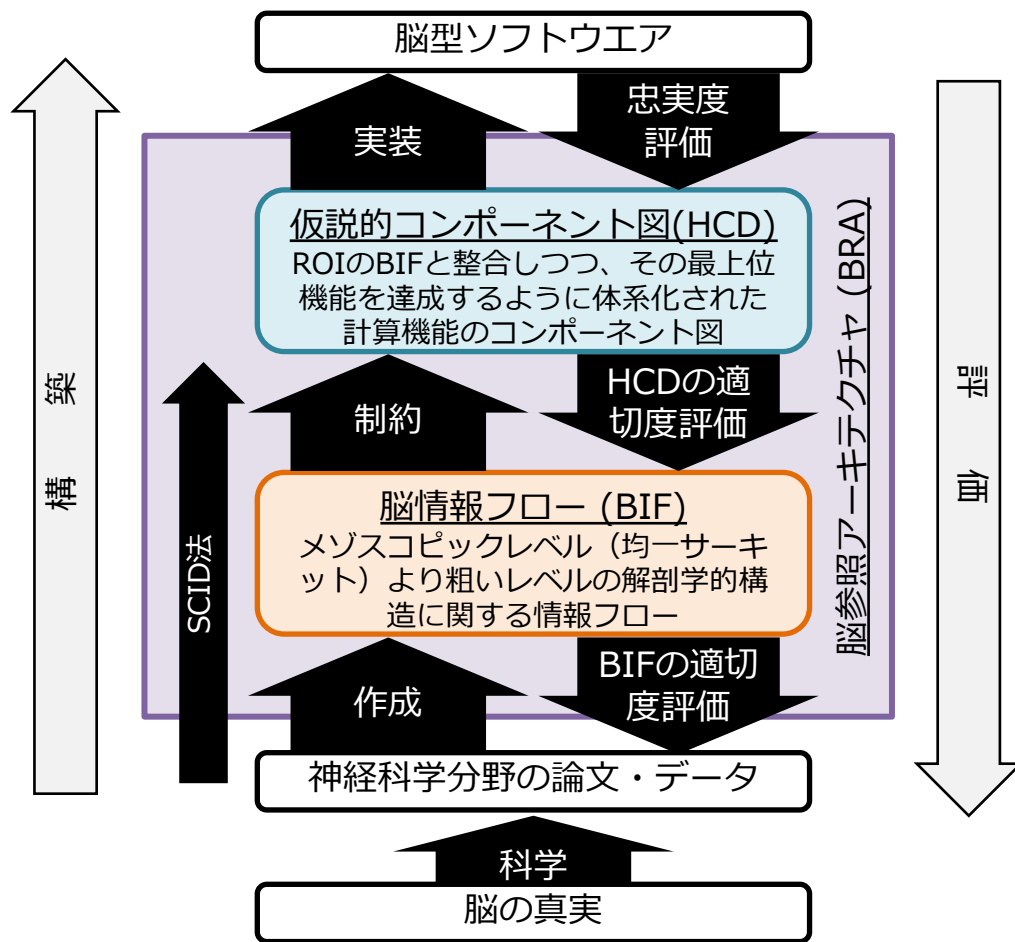


NPO法人WBAI代表

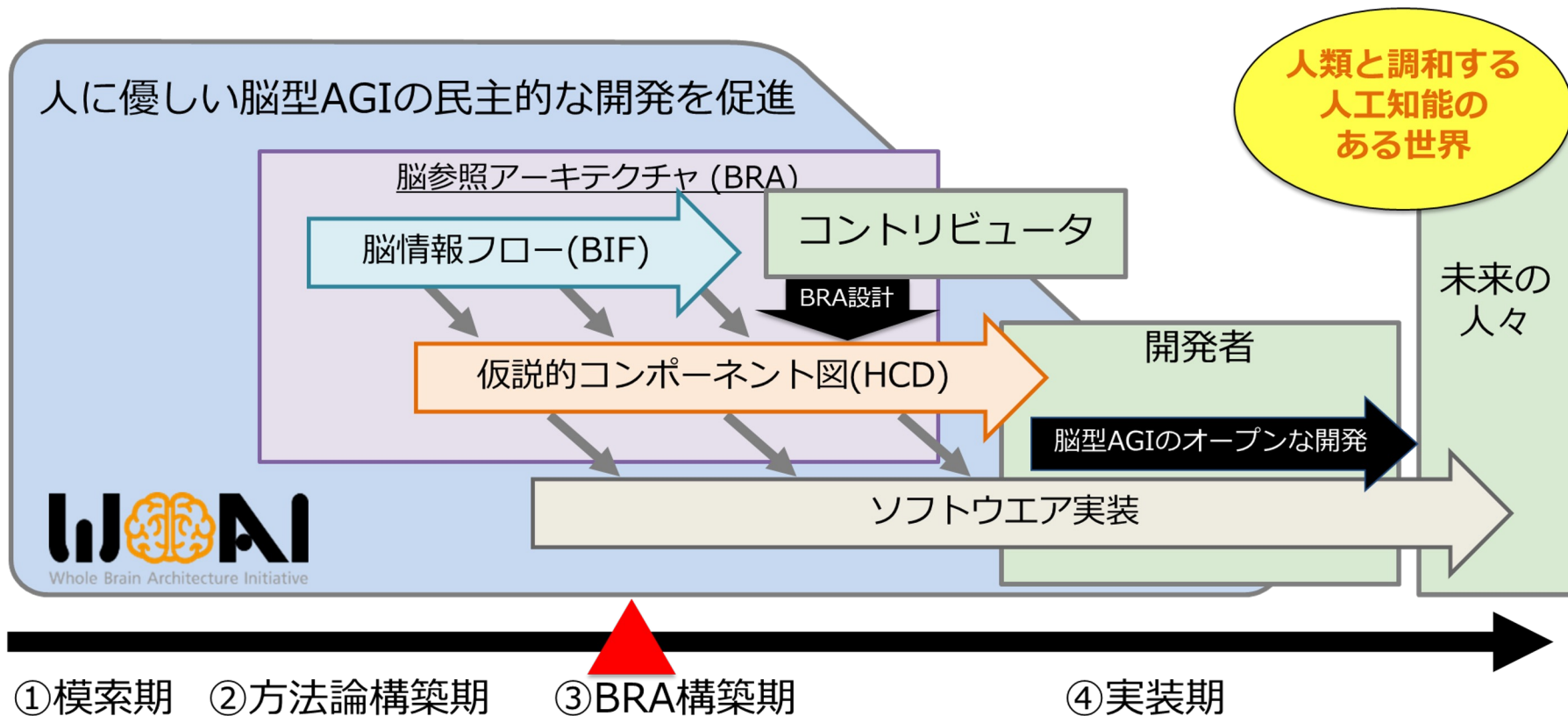
山川 宏

BRA駆動開発

脳全体の神経回路を参照することで、人間の解剖学的構造に高い整合性の高い認知機能を再現するソフトウェアを構築する手法



WBAIのビジョンにむけた技術ロードマップ



WBAIの推進体制

顧問： 銅谷賢治 (OIST) 北野宏明 (Sony)
 富田勝 (慶應義塾大) 森川博之 (東大)
 岡ノ谷一夫 (東大)

理事： 山川宏 (代表) 松尾豊 (副代表/東大)
 高橋恒一 (副代表/理研) 荒川直哉

監事： 井上康秀

会議： 総会、理事会、その他委員会

賛助会員

■ 特別賛助会員

AISIN

■ 賛助会員

 **IGPI Business Analytics
& Intelligence**

 **materializer & Co.**

 **TOPWELL**
Make ALL Happy


MITSUBISHI RESEARCH INSTITUTE, INC.

Panasonic
CONNECT

Panasonic

■ 賛助会員（個人）

後藤健太郎 様 上林厚志様 上野山勝也 様

賛助会員

■ 創設賛助会員

dwango

TOYOTA

Panasonic

IP Partners

TOSHIBA

活動趣旨をご理解の上、
ご支援頂けますと助かります

クラウドファンディング



アカデミストというクラウド
ファンディングサイトで寄付の
募集を始めました！

<https://academist-cf.com/fanclubs/267>



academistとは

**研究者がいきる、
私たちがつなぐ。**

研究者たちが最大限活躍できる世の中へ。



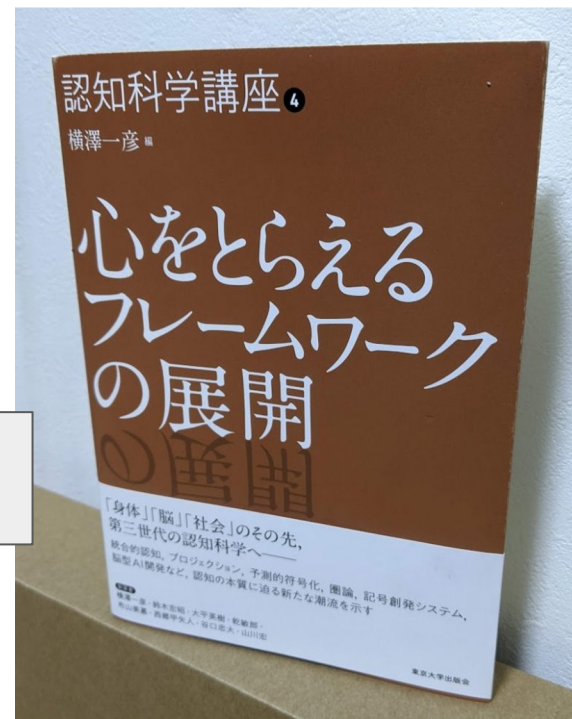
全脳アーキテクチャの解説を含む

書籍「心をとらえるフレームワークの展開」の発売

目次：

- 序 第三世代の認知科学の可能性（横澤一彦）
- 第1章 統合的認知（横澤一彦）
- 第2章 プロジェクション科学（鈴木宏昭）
- 第3章 内受容感覚の予測的処理（大平英樹）
- 第4章 自由エネルギー原理（乾敏郎）
- 第5章 圏論による認知の理解（布山美慕・西郷甲矢人）
- 第6章 記号創発ロボティクス（谷口忠大）
- 第7章 全脳アーキテクチャ
——機能を理解しながら脳型AIを設計・開発する
（山川 宏）

全脳確率的
生成モデル



<https://wba-meetup.connpass.com/event/260105/>



第7回全脳アーキテクチャシンポジウム

汎用ロボティクスと脳型知能

2022年10月12日 13:15~17:20

パネル討論

脳を参照するロボティクス

主 催： NPO法人 全脳アーキテクチャ・イニシアティブ
後 援： 学術変革 行動変容生物学
協 力： AIアート集団「波羅密多」
運 営： WBA勉強会実行委員会

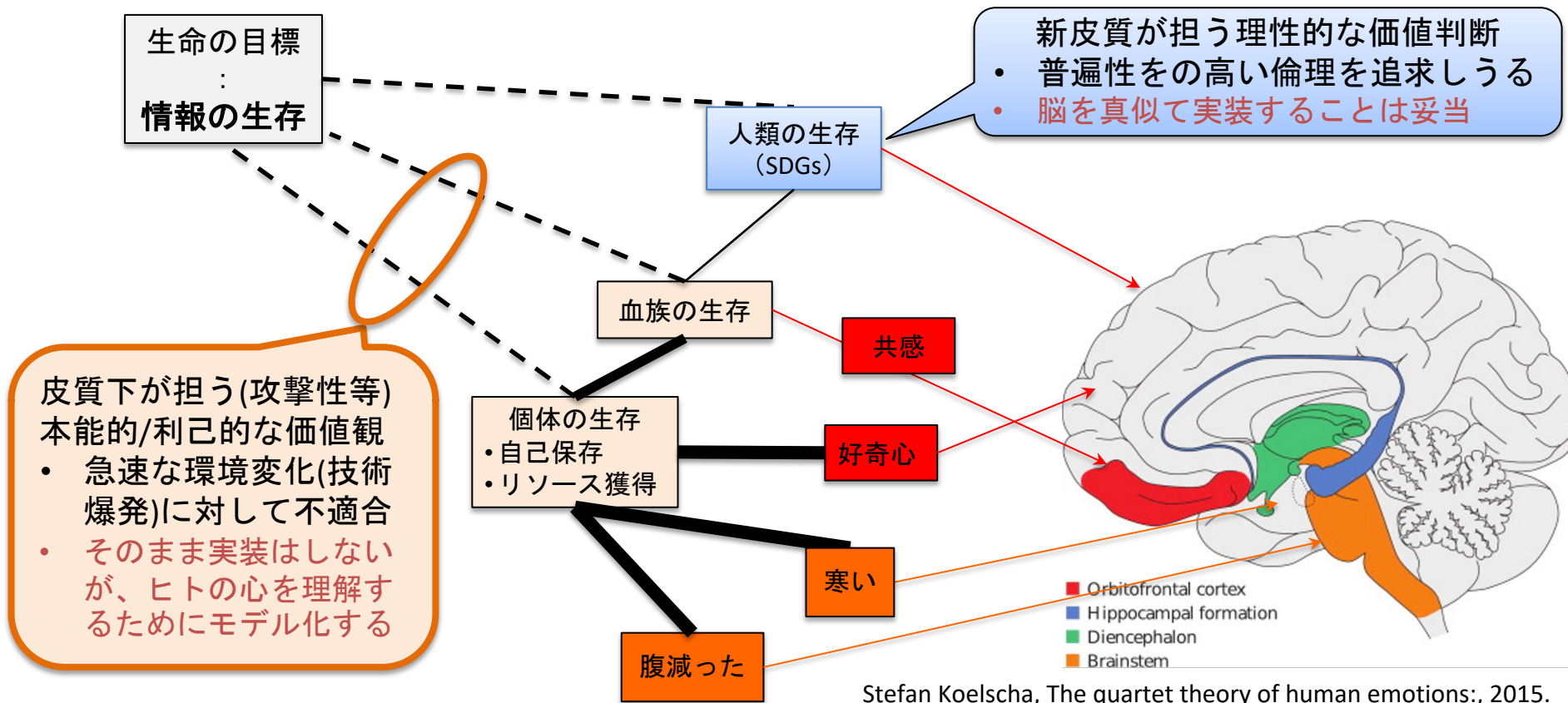


WBAアプローチの利点：

人と親和性の高いAGIを作れる

- 対人インタラクションが必要な応用
- 精神疾患などのモデル化など医療応用
- Mind-uploadingの器

WBAIで議論されてきた倫理性に配慮した脳型AGI



人工脳が心をもつならば：感情と意識

心とは、人間（や生き物）の精神的な作用や、それのもとになるものなどを指し、感情、意志、知識、思いやり、情などを含みつつ指している。

- 感情

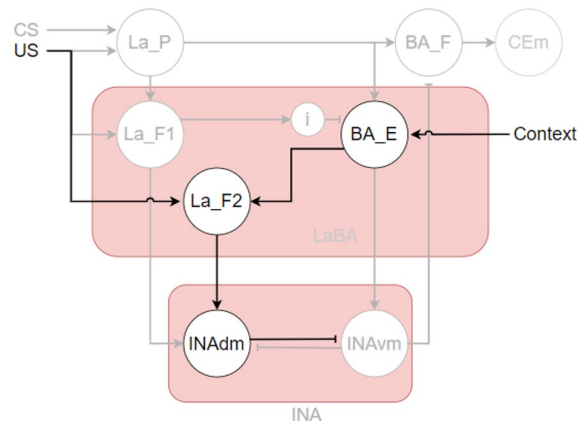
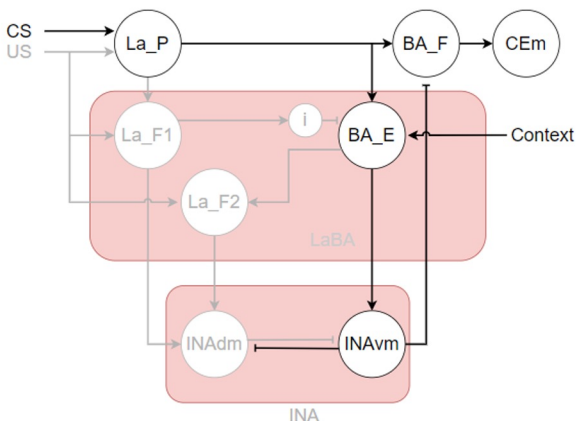
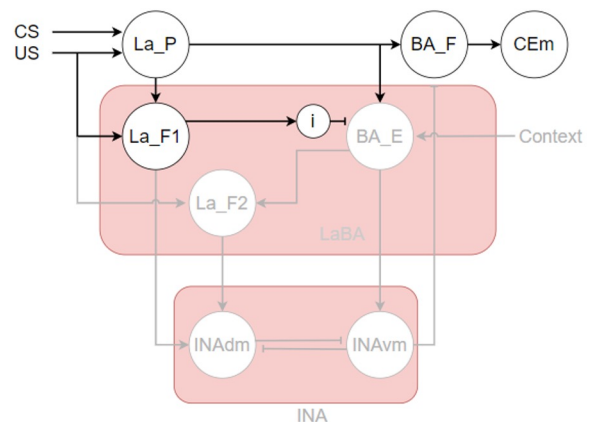
- 日永田智絵（奈良先端科学技術大学院大学）
 - 感情発達ロボティクス

- 意識

- 渡邊正峰（東京大学）
 - 「意識の科学」のための「意識のアップロード」或いは「意識のアップロード」のための「意識の科学」
- 新川拓哉（神戸大学）
 - 意識の価値

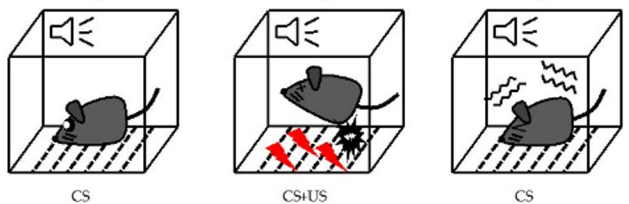
- 扁桃核における情動関連の機能仮説の構築
 - 恐怖条件付けにおける消去学習と回復を実現する扁桃体の相互抑制機構を考慮した機能仮説（宮本 竜也・田和辻可昌・山川 宏）
 - <https://ken.ieice.org/ken/paper/20220930RCMp/>
- 意識のEntification仮説（山川）
 - 世界から存在を見出すためのシリアルな処理を意識とみなす
 - 参考文献
 - 布川絢子, 松尾豊. 意識の Entification 仮説. in 2022年度人工知能学会全国大会（第36回）
 - 第22回 汎用人工知能研究会 (2022/11/22) to appear

条件付け・消滅・回復の機構について、生物学的に妥当な機能仮説を提案



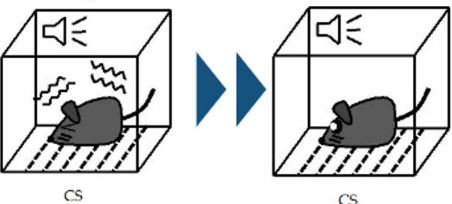
USと同時に提示されたCSに対して、恐怖反応が条件付けされる

恐怖条件づけ



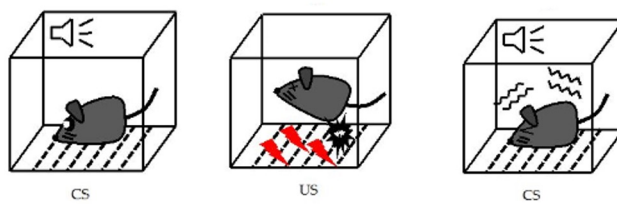
US のない状態でのCS の繰り返し提示により、同一Contextに安

消去学習

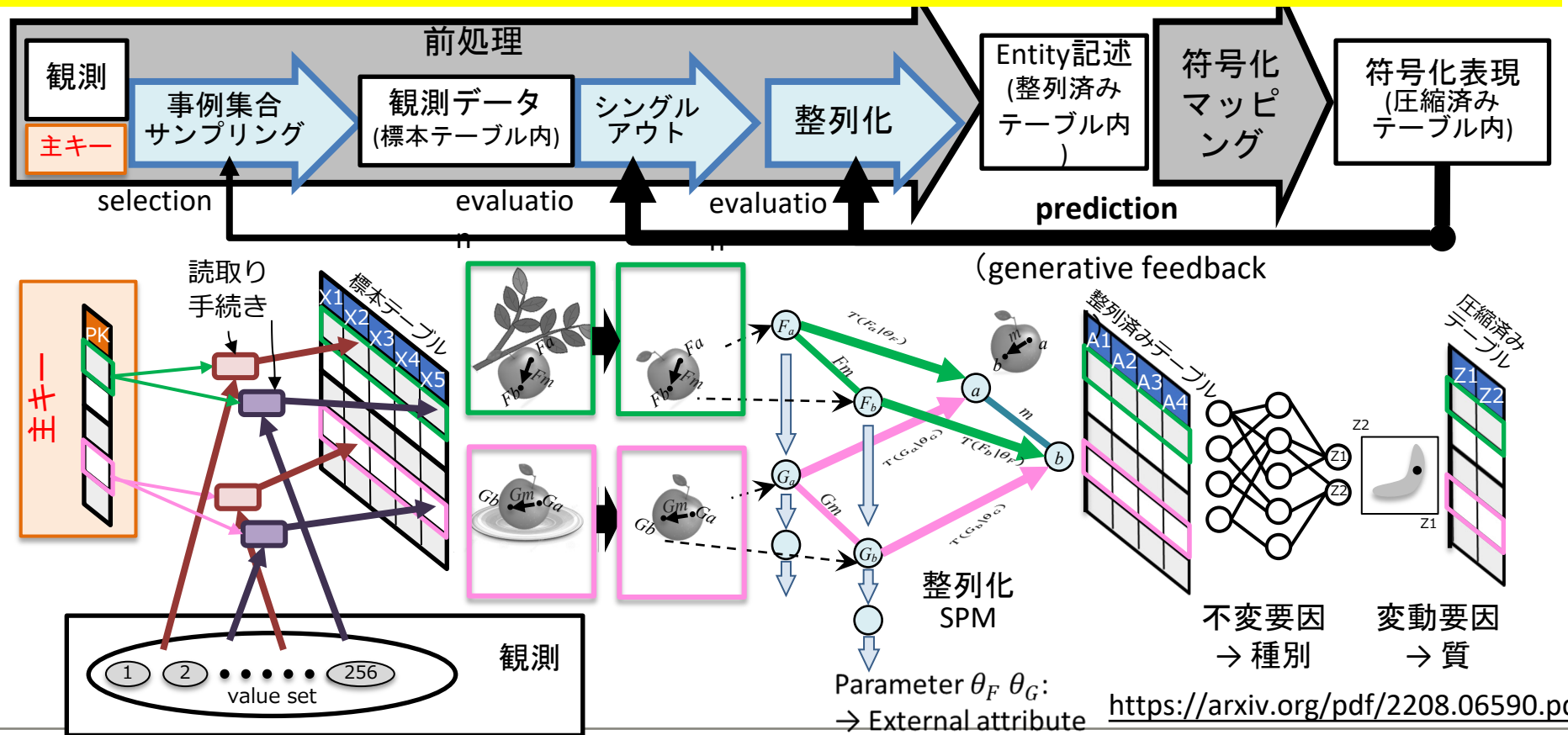


左記と同Context でUS が提示されたときに恐怖反応が回復

回復(Reinstatement)



世界から観測される情報のなかに繰り返し現れる不変のパターンを見出す処理



意識のEntification仮説：主キー系列で制御されるEntification処理

仮説：意識とは、**主キー系列**を用いて行われるEntification処理

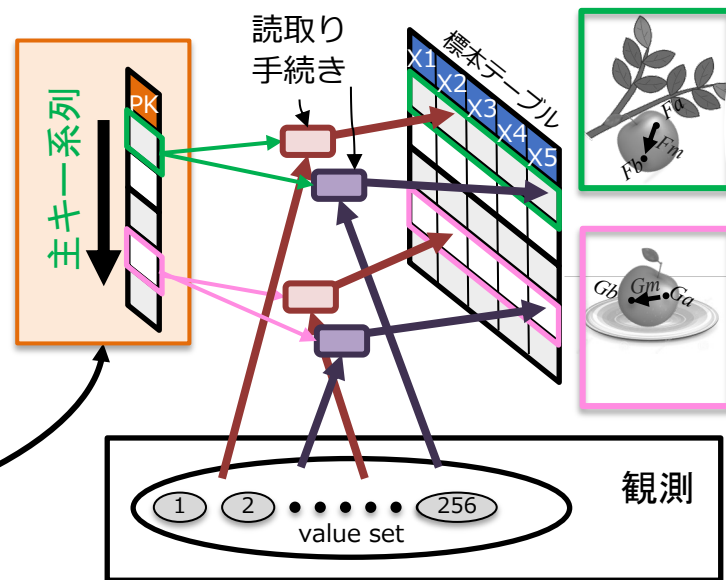
なぜ系列なのか？

新皮質上では、
ある時点で表現できる内容は一つ

隣接する時刻間のみで表現内容が比較可能
(誤差を計算することができる)

主キー系列

順序付けされた主キー



心をもつAIはどのようにつくられるべきか？ by 山川

- 単に攻撃能力が高い高度AIが対立すれば、技術爆発で狭くなった世界では、相互破壊にいたる危険性が極めて高い(Existential Risk)。
- よって知的でありながら共存共栄を実現するには「信頼関係を構築する能力」に優れたAIを構築すべき。
- それゆえ、信頼構築能力をもつAIがいわゆる「倫理空間」への参加権（道德的地位を認められる）をもつ。

Entification能力は、信頼構築能力の基盤であるため、少なくとも、それをもたないレベルの主体は、倫理空間に入る資格をもたないだろう。

- 他者認識：主体の候補となる外界の事物を存在として扱う能力
- 意図推定：自己の内部状態(意図・感情・信念など)を存在として扱い、それを他者に対応付ける類推能力

趣旨説明：人工脳が心をもつならば

- 高度な人工知能を人の脳のように構築してゆけば、それは、ヒトと近い、心／感情／意識をもつ主体となっていくであろう
 - それは時に過去の生者から連続もしくは非連続的な意識がアップロードした主体であるかもしれない。
 - 進展著しい、大規模言語モデルなどとの接点はどうなるか？
- 本勉強会では、倫理的な課題をひとつの軸としつつ議論を行う
 - 人々はそうした主体になりたいと思うのか
 - そうした主体の尊厳を認めるべきかといった

運営体制

プログラム委員長： 山川 宏
実行委員長： 森岡大成
司会者： 濱田太陽
Zoom： 生島 高裕、西村由弥子
Connpass運営： 孫 暁白
Slack運営： 土肥栄祐

運営全般： WBA勉強会実行委員会

活動募集

活動への参加、お待ちしております

1. WBA勉強会実行委員会
2. WBAI開発部
3. 賛助会員

勉強会やシンポジウムの企画や運営
支援を行うボランティアグループで
す。

ご興味のある方はこちらから

[https://wba-
initiative.org/get_involved/join/](https://wba-initiative.org/get_involved/join/)



脳型AIの実装、開発方法論や
ロードマップの議論にご興味ある
方はぜひご参加ください

wba-initiative.org/join-sig-wba



賛助会員

理念に賛同し汎用人工知能の実現
を支援して下さる賛助会員を募っ
ています。

<https://wba-initiative.org/support/>



クラウドファンディング



アカデミストというクラウド
ファンディングサイトで寄付の
募集を始めました！

<https://academist-cf.com/fanclubs/267>



academistとは

**研究者がいきる、
私たちがつなぐ。**

研究者たちが最大限活躍できる世の中へ。

