

HAIとWBAのアプローチの差異と接合点から 見るインタラクションのデザイン

坂本孝丈（静岡大学）

第33回 全脳アーキテクチャ勉強会 @オンライン

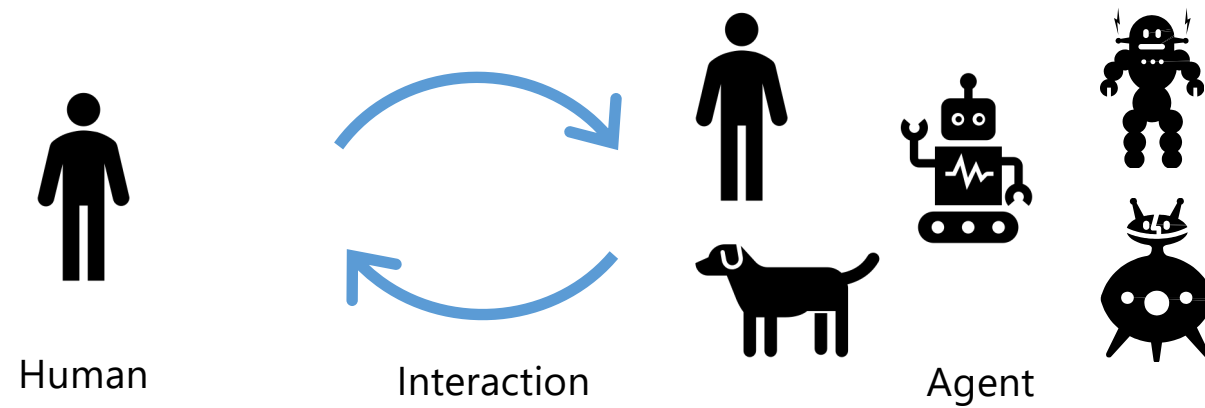
自己紹介

坂本孝丈 （情報学 博士）

静岡大学 創造科学技術大学院 特任助教

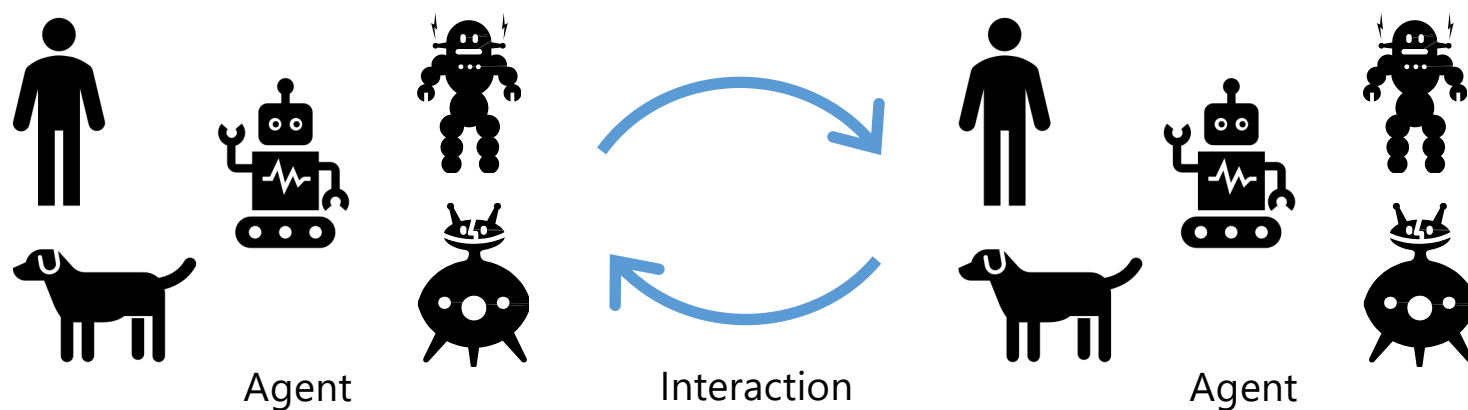
専門： Human-Agent Interaction, 認知科学

Human-Agent Interaction (HAI)



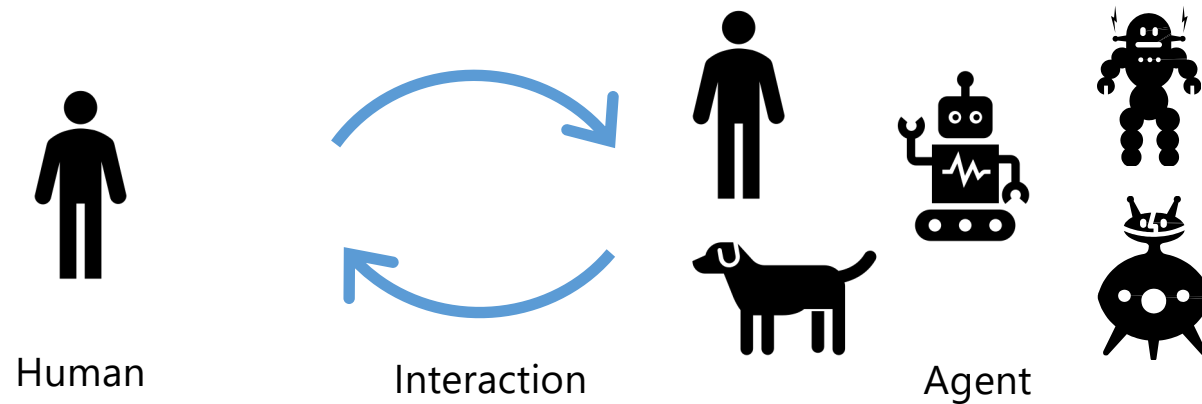
Human-Agent Interaction (HAI)

比較や他領域との接続を考えるうえでは



Human-Agent Interaction (HAI)

基本的にはここにつながるはず



HAI関連研究の流れ

- 1960年頃～ 認知革命→ 認知科学
- 1990年頃～ 4E（身体性や環境との相互作用の重要性）
- 1990年頃～ Agent：自律的に環境と相互作用する存在
- 1996年 Media equation
- 2000年頃～ 人とエージェントのインタラクションに関する議論
- 2006年～ HAIシンポジウム
- 2013年～ International conference of Human-Agent Interaction
- 2014年～ 認知的インタラクションデザイン学

HAI関連研究の流れ

1960年頃～ 認知革命→ 認知科学

1990年頃～ 4E（身体性や環境との相互作用の重要性）

「心」へのアプローチ

哲学，心理学，人工知能，言語学，人類学，神経科学の各分野
との学祭的研究 [Gardner 1987]

キー：表象，認知は計算である（コンピュータによるモデル化）

HAI関連研究の流れ

1960年頃～ 認知革命→ 認知科学

1990年頃～ 4E（身体性や環境との相互作用の重要性）

1990年頃～ Agent：自律的に環境と相互作用する存在

「心」へのアプローチ

Embodied：	身体化された	[Pfeifer & Scheier 1999]
Embedded：	（行動が生じる環境に）埋め込まれ	[Gibson 1979]
Extended：	（身体の外側に）拡張された	[Gibson 1979]
Enactive：	行動に基づいた	[Varela et al.1991]

HAI関連研究の流れ

1960年頃～ 認知革命→ 認知科学

1990年頃～ 4E（身体性や環境との相互作用の重要性）

1990年頃～ Agent：自律的に環境と相互作用する存在

1996年 Media equation

自律的に動作するシステムがAgentと呼ばれ始める
「センサーを介して環境を認識し，効果器を介して環境に作用するもの」，「自律的に環境と相互作用するもの」

※ キノコ喰いロボット（1960年代，戸田）

→ 完全エージェント [Pfeifer and Scheier 1999]

→ エージェントの設計の原理 [Pfeifer and Bongard 2006]

HAI関連研究の流れ

1960年頃～ 認知革命→ 認知科学

1990年頃～ 4E（身体性や環境との相互作用の重要性）

1990年頃～ Agent：自律的に環境と相互作用する存在

1996年 Media equation

2000年頃～ 人とエージェントのインタラクションに関する議論

自律的に動作するコンピュータに対して社会的に
振る舞う人間の傾向 [Reeves and Nass 1996]

HAI関連研究の流れ

1960年頃～ 認知革命→ 認知科学

1990年頃～ 4E（身体性や環境との相互作用の重要性）

1990年頃～ Agent：自律的に環境と相互作用する存在

1996年 Media equation

2000年頃～ 人とエージェントのインタラクションに関する議論

2006年～ HAIシンポジウム 

Human-Agent Interaction (HAI)

エージェント：「人間とインタラクションする自律システム、
または、そのような自律システムに見せかけるシステム」

HAI研究では、身体や外観などの属性を必要条件とし仮定するのではなく、それらが必要であるか否かについても議論の対象に含まれる

[山田（監修・著）2007]

HAI関連研究の流れ

インタラクション → 対象をエージェントと見なす
そのような人の認知過程を明らかにする,
あるいは, そのためのシステムの設計を行う研究領域

Cf.) Human-Robot Interaction

□ロボット技術→人がいる環境内での行動の設計

2006年～ HAIシンポジウム

2013年～ International conference of Human-Agent Interaction

2014年～ 認知的インタラクションデザイン学

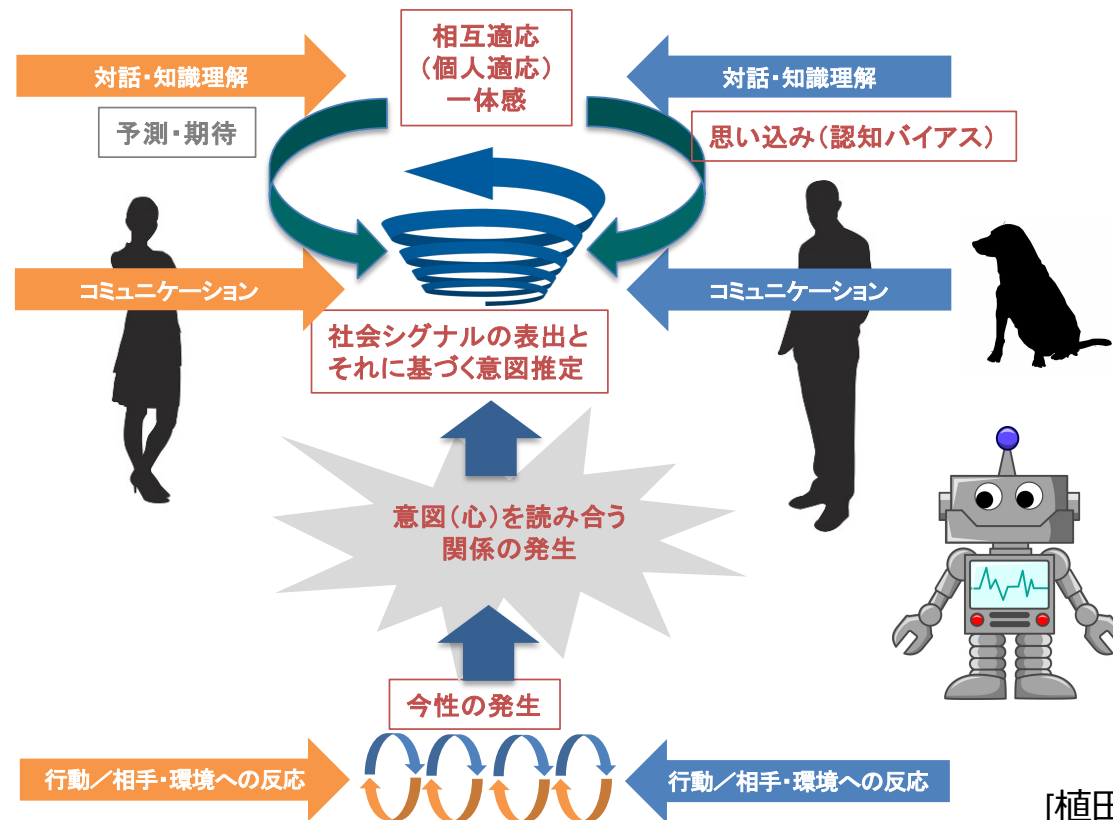
「認知的インタラクションデザイン学～意思疎通のモデル論的理解と人工物設計への応用」

平成 26 ～ 30 年度 文部科学省科学研究補助金新学 術領域研究（研究領域提案型）採択課題



「認知的インタラクションデザイン学～意思疎通のモデル論的理解と人工物設計への応用」

平成 26 ～ 30 年度 文部科学省科学研究補助金新学術領域研究（研究領域提案型）採択課題



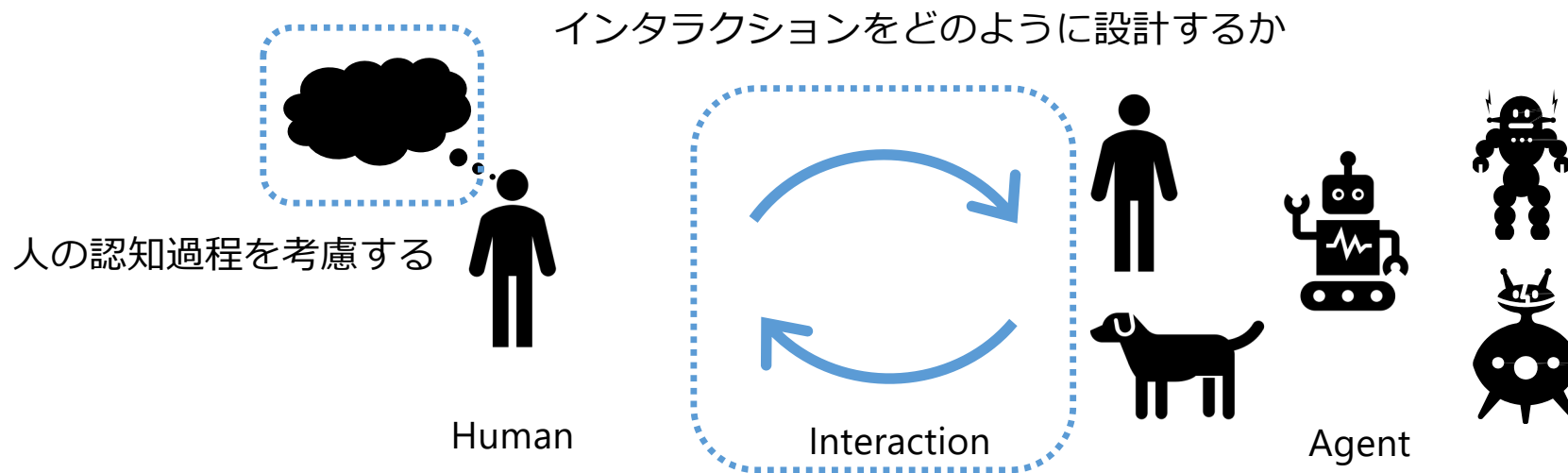
HAI関連研究の流れ

- 1960年頃～ 認知革命→ 認知科学
- 1990年頃～ 4E（身体性や環境との相互作用の重要性）
- 1990年頃～ Agent：自律的に環境と相互作用する存在
- 1996年 Media equation
- 2000年頃～ 人とエージェントのインタラクションに関する議論
- 2006年～ HAIシンポジウム
- 2013年～ International conference of Human-Agent Interaction
- 2014年～ 認知的インタラクションデザイン学

Human-Agent Interaction (HAI)

エージェント：「人間とインタラクションする自律システム、または、
そのような自律システムに見せかけるシステム」

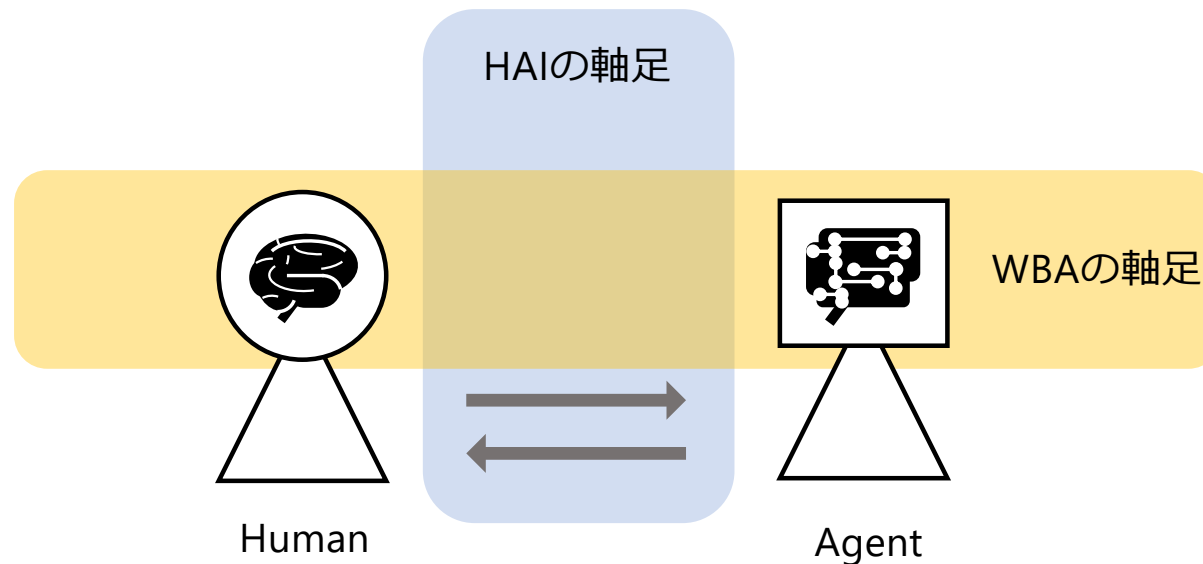
HAI研究では、身体や外観などの属性を必要条件とし仮定するのではなく、それらが必要であるか否かについても議論の対象に含まれる [山田（監修・著）2007]



HAIのアプローチとWBAのアプローチ

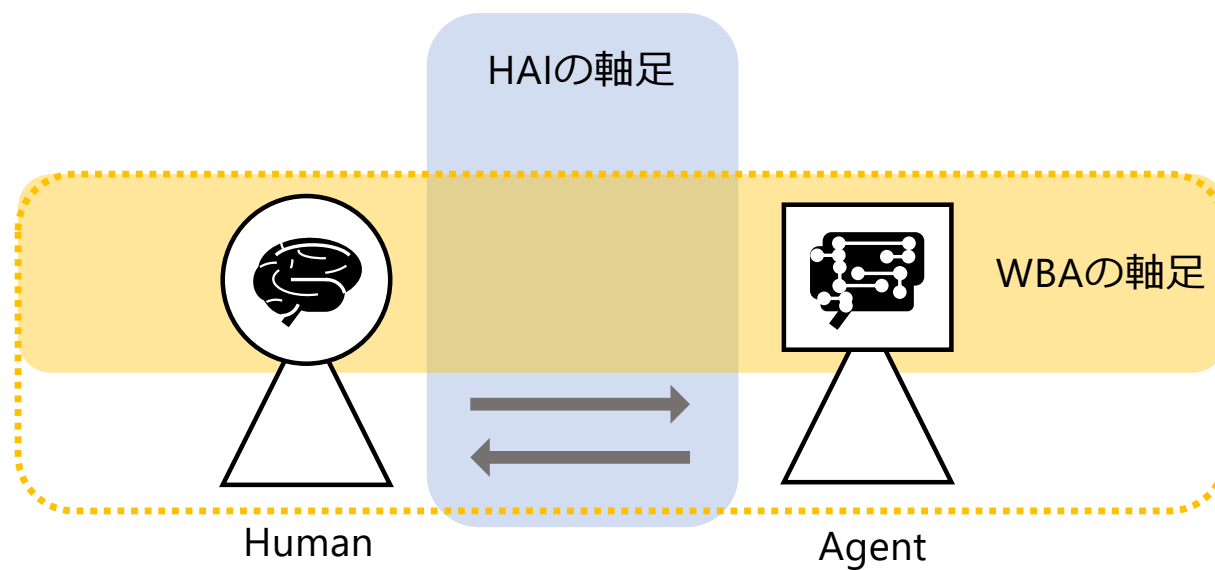
HAIのアプローチ：人とエージェントのインタラクションの設計

WBAのアプローチ：人の脳全体を参考にしたアーキテクチャの設計



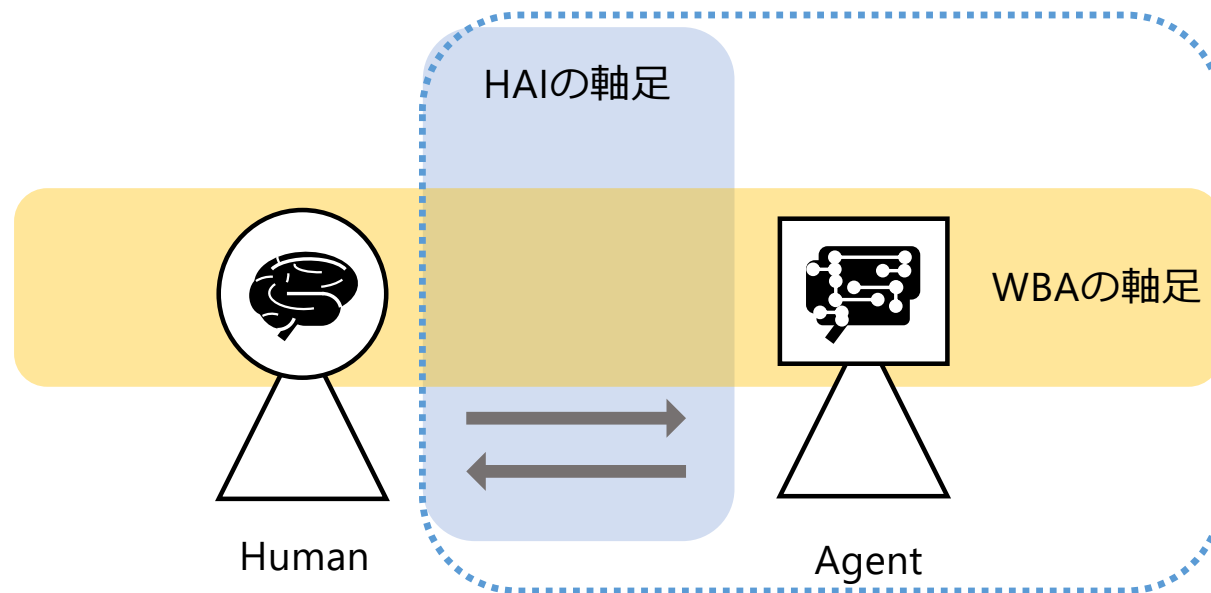
WBA→HAI

大澤さんの話（おそらく）



HAI→WBAの距離（エージェントを設計する）

Agentを設計するうえで脳を仮定する必要があるのか？

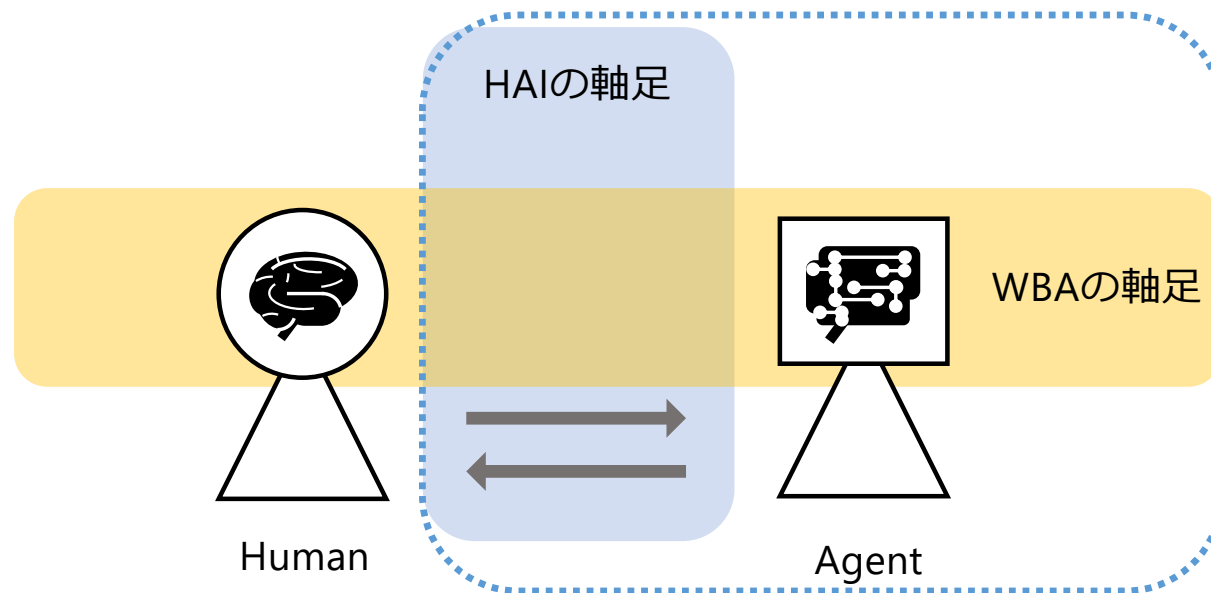


HAI→WBAの距離（エージェントを設計する）

Agentを設計するうえで脳を仮定する必要があるのか？

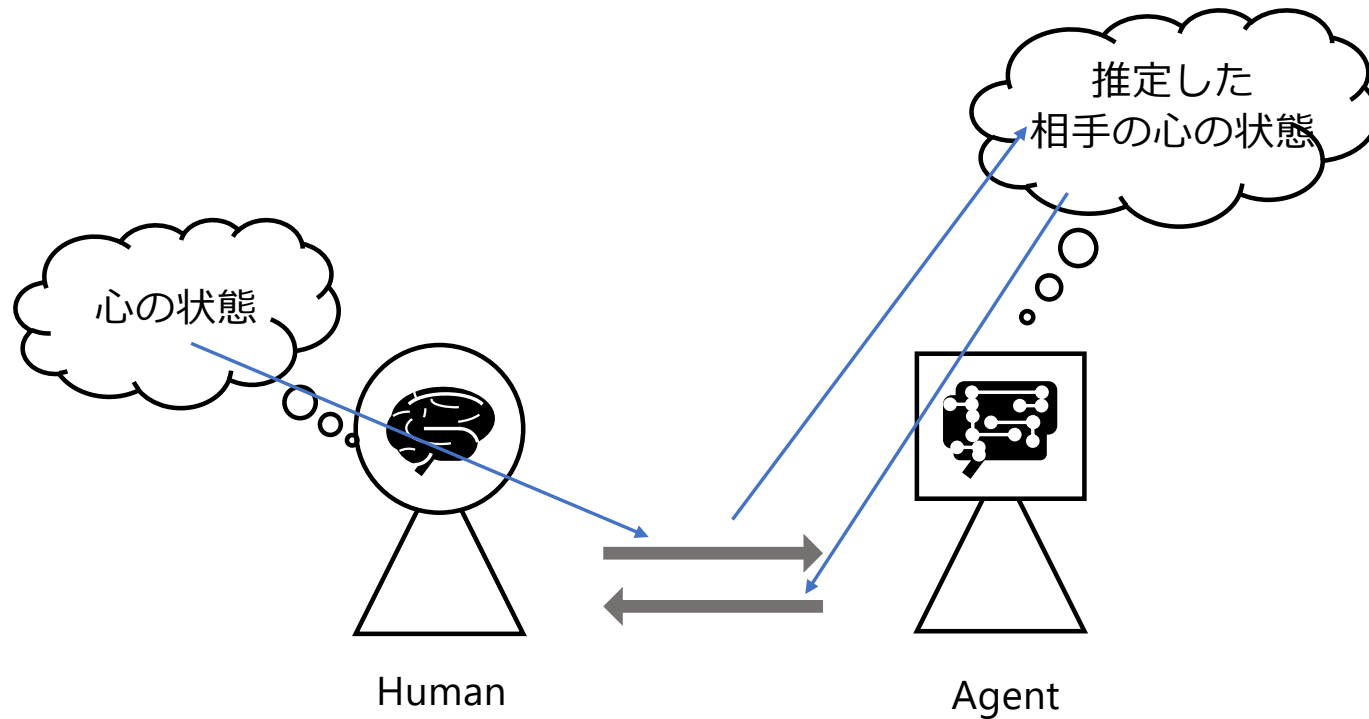
人は相手の**心**を想定しているけれど、**脳**は想定していないのでは？

HAI的には人が対象（エージェント）に**心**を見出すかどうか重要



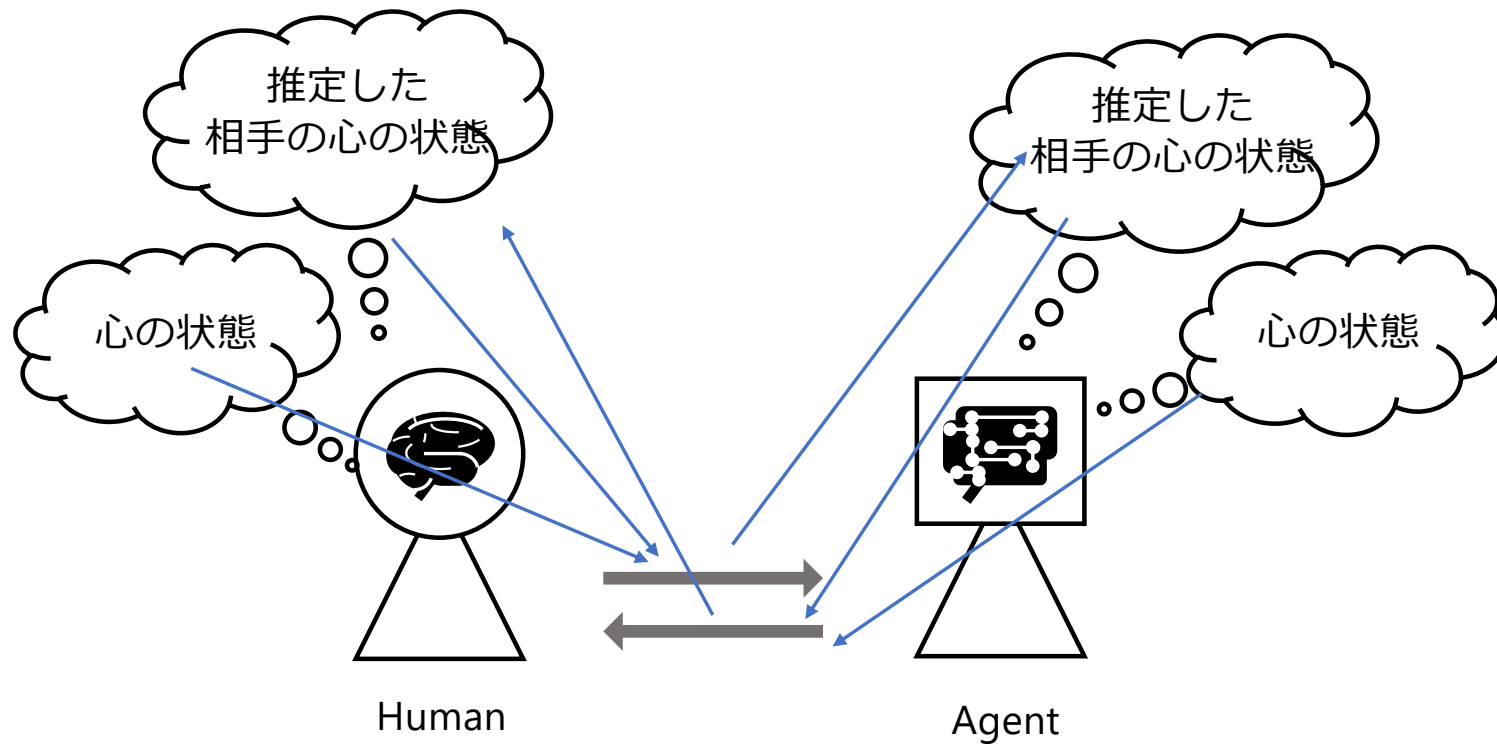
HAI→WBAの距離（エージェントを設計する）

人と（人工物）エージェントのインタラクションの問題



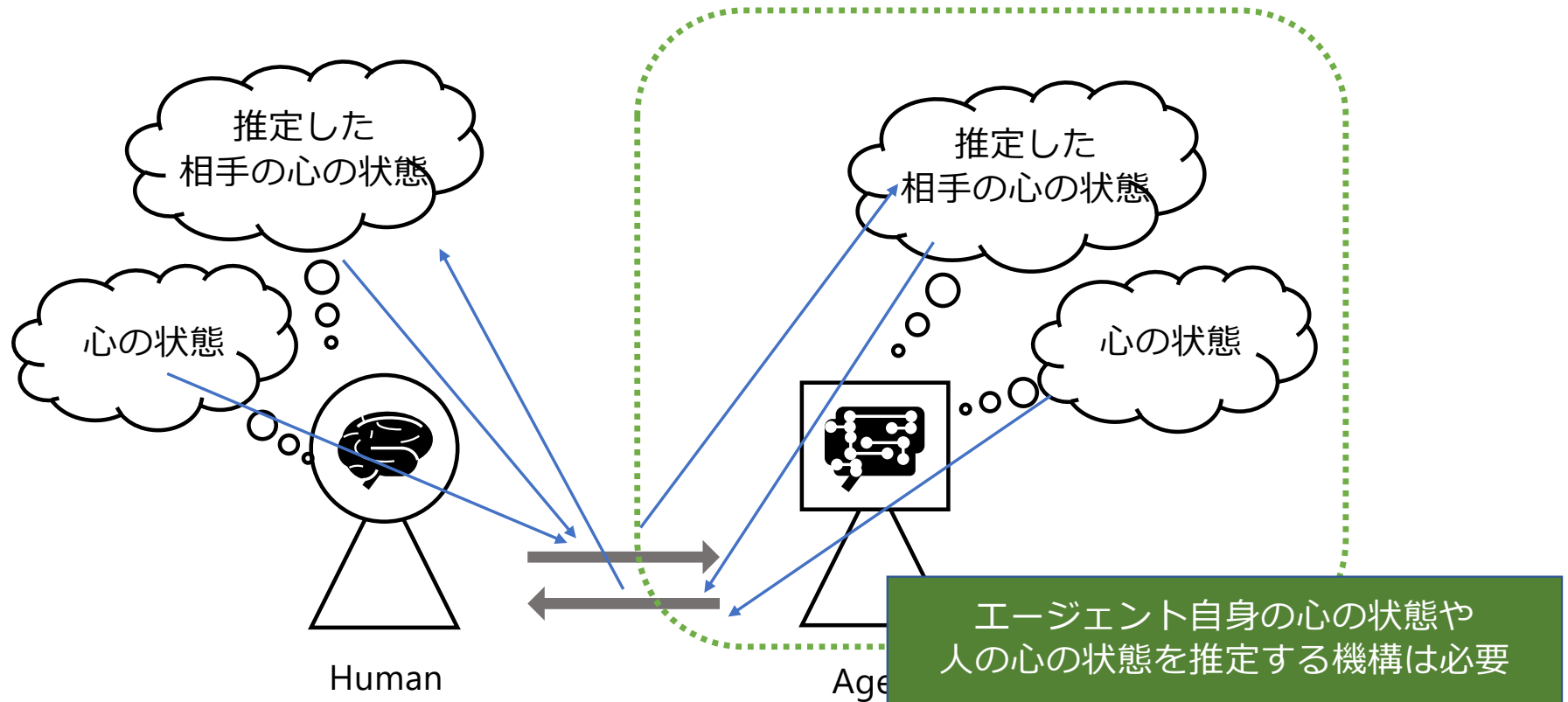
HAI→WBAの距離（エージェントを設計する）

人と（人工物）エージェントのインタラクションの問題



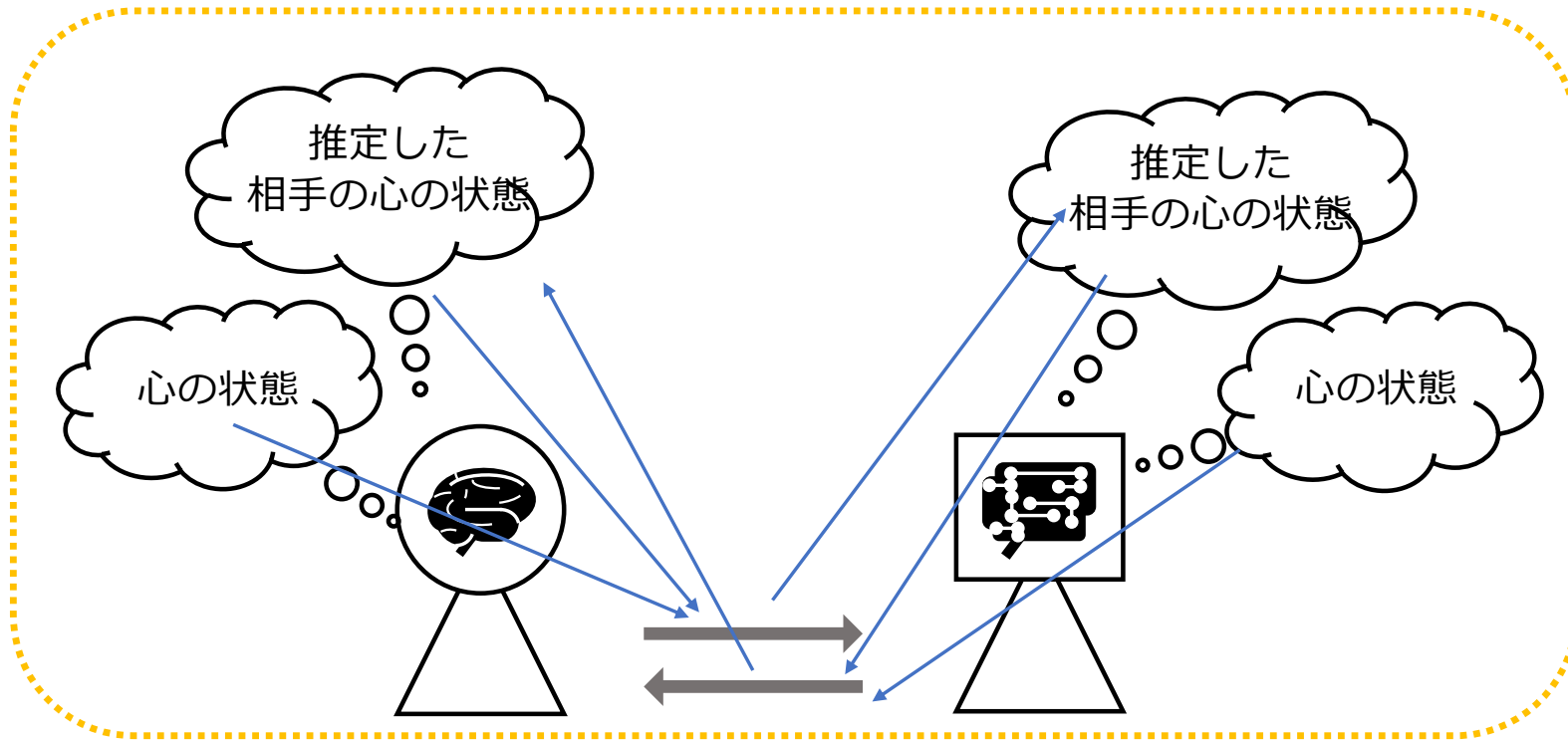
HAI→WBAの距離（エージェントを設計する）

人と（人工物）エージェントのインタラクションの問題



HAI→WBAの距離（エージェントを設計する）

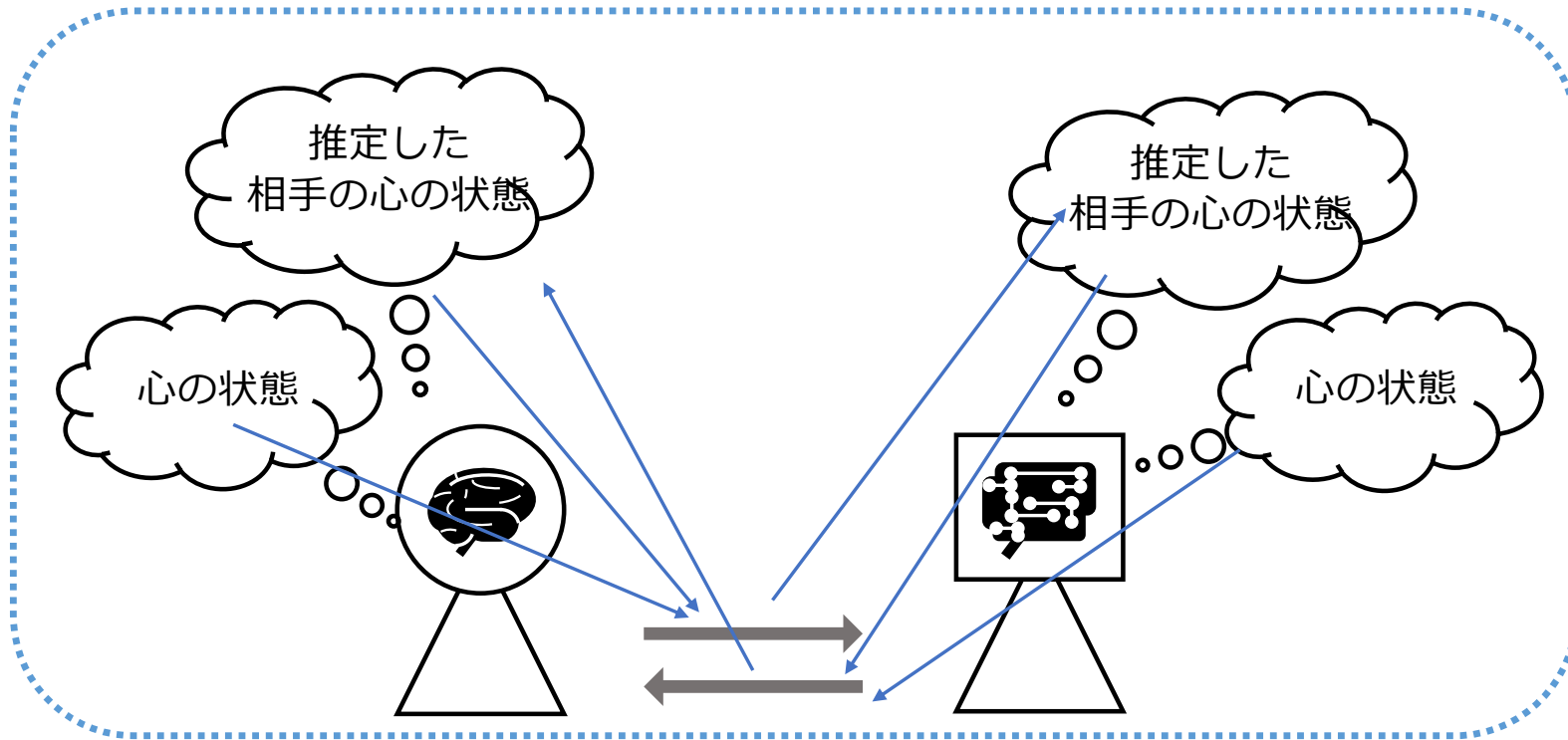
人と（人工物）エージェントのインタラクションの問題



エージェントが人の脳のアーキテクチャを持つ ⇒ 人同士のインタラクションと同質？

HAI→WBAの距離（エージェントを設計する）

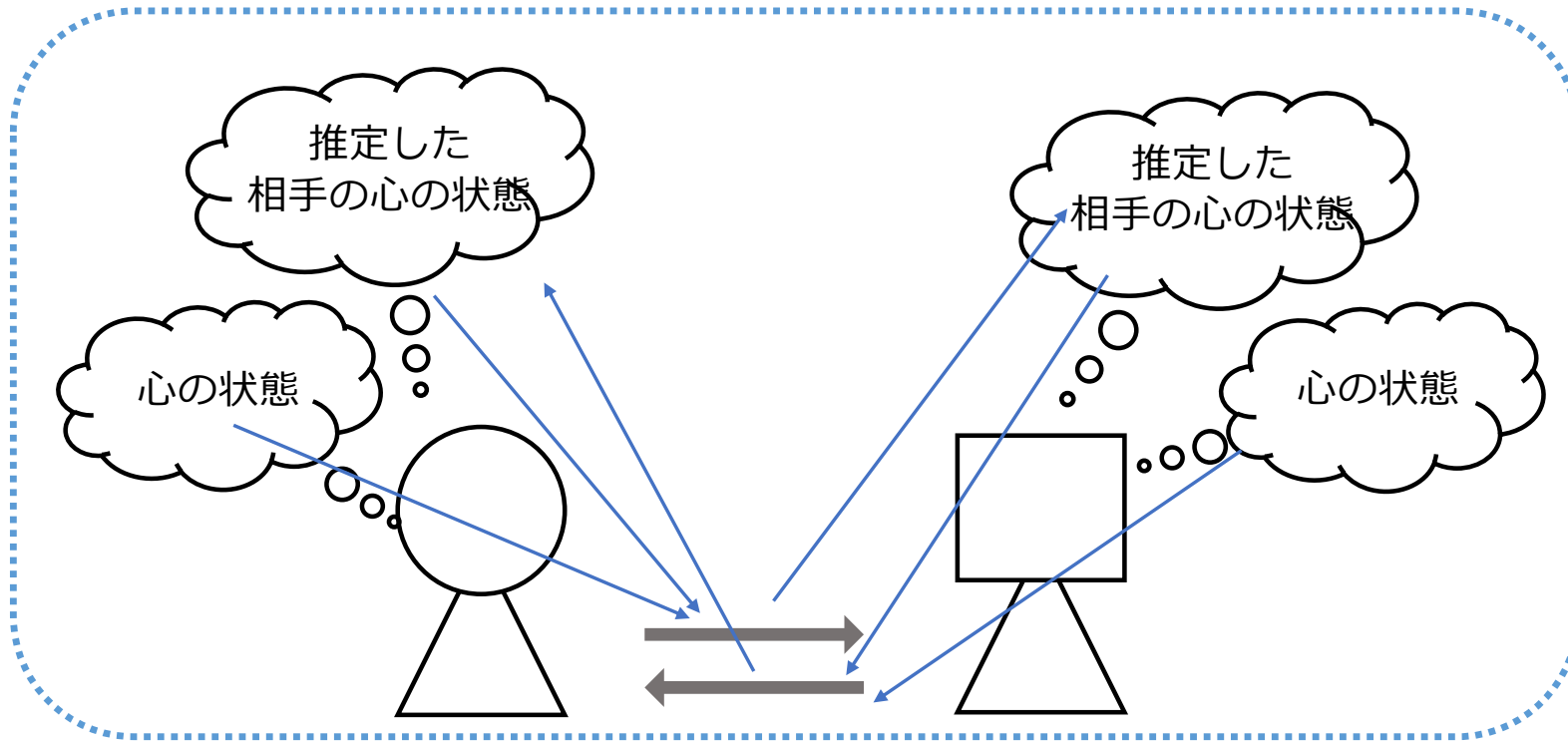
人と（人工物）エージェントのインタラクションの問題



人同士と同質なインタラクションを再現するために必要な機構
（仮定する心の状態やエージェントが持つ機能）は何か？

HAI→WBAの距離（エージェントを設計する）

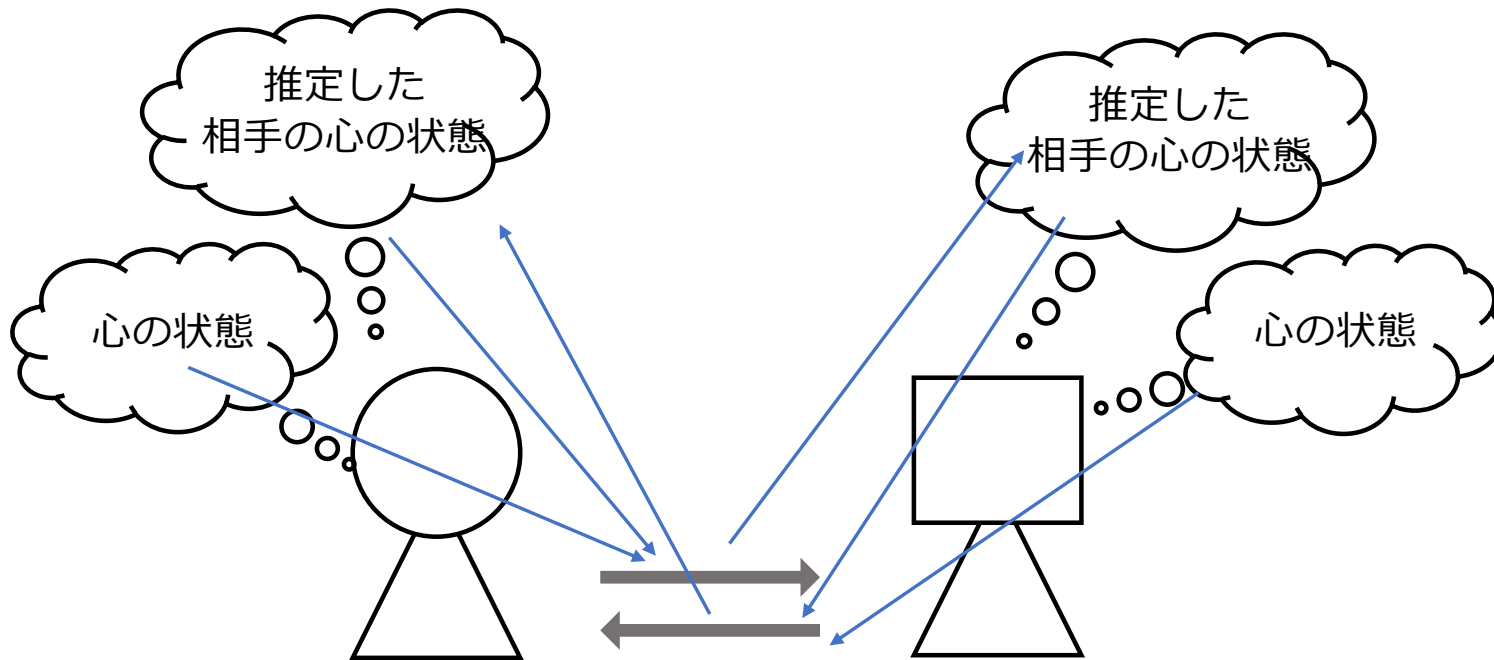
人と（人工物）エージェントのインタラクションの問題



人同士と同質なインタラクションを再現するために必要な機構
（仮定する心の状態やエージェントが持つ機能）は何か？

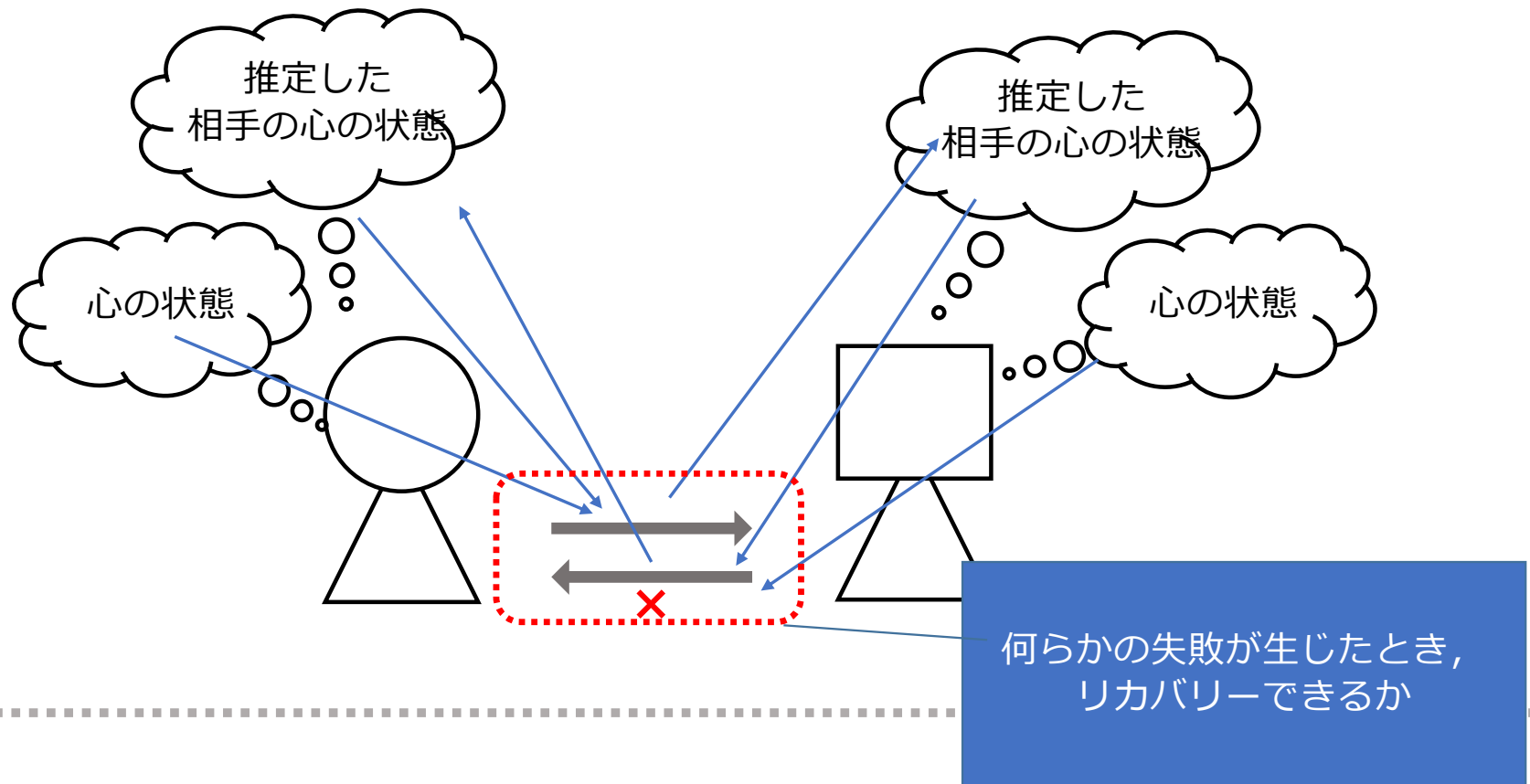
HAI→WBAの距離（エージェントを設計する）

インタラクションの問題における汎用性 → 長期的インタラクション



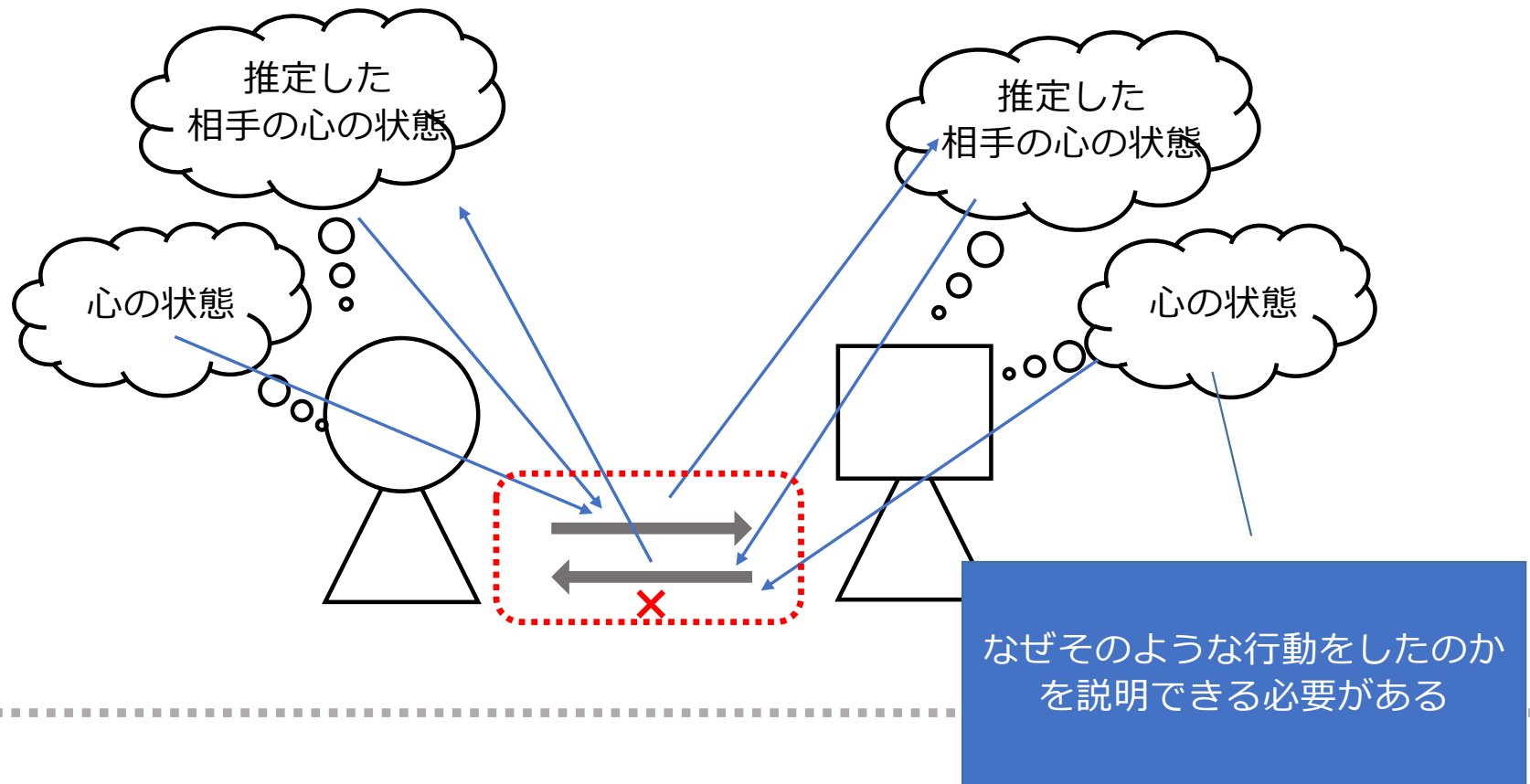
HAI→WBAの距離（エージェントを設計する）

インタラクションの問題における汎用性 → 長期的インタラクション



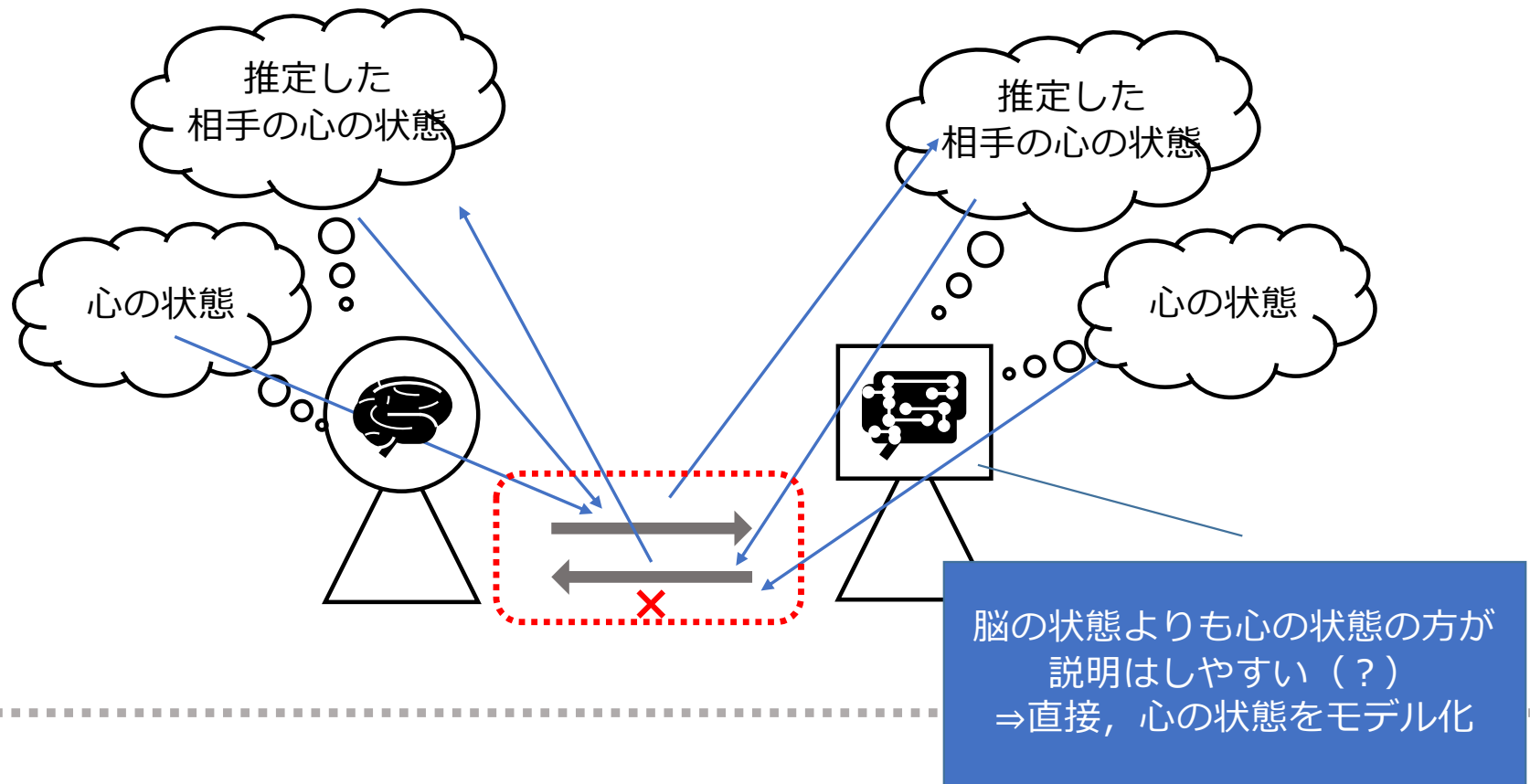
HAI→WBAの距離（エージェントを設計する）

インタラクションの問題における汎用性 → 長期的インタラクション



HAI→WBAの距離（エージェントを設計する）

インタラクションの問題における汎用性 → 長期的インタラクション



HAI→WBAの距離（エージェントの設計）

HAIの視点：人が心を見出す対象はある範囲に収まっている（はず）

→ 人が相手を他者と見做す最小限のアーキテクチャはありそう

単純な構造（要素やエージェントが持つ機能）からモデルを拡張していくことで、
生じ得るインタラクションの違いを明らかにしていく

→ ボトムアップ

→ 特定の機能があれば良いのであって、（人の）脳を仮定する必要はない

一方で、WBAでは人の脳のアーキテクチャに注目している

→ トップダウン（?）

（※ ただしHAI研究が必ずしもボトムアップのアプローチというわけではない...むしろ少数?）

HAI→WBAの距離（エージェントの設計）

HAIの視点：人が心を見出す対象はある範囲に収まっている（はず）

→ 人が相手を他者と見做す最小限のアーキテクチャはありそう

単純な構造（要素やエージェントが持つ機能）からモデルを拡張していくことで、
生じ得るインタラクションの違いを明らかにしていく

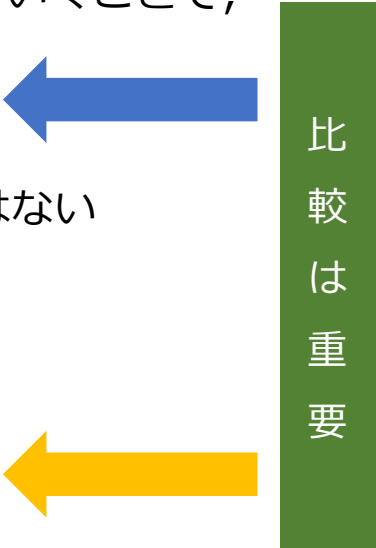
→ ボトムアップ

→ 特定の機能があれば良いのであって、（人の）脳を仮定する必要はない

一方で、WBAでは人の脳のアーキテクチャに注目している

→ トップダウン（?）

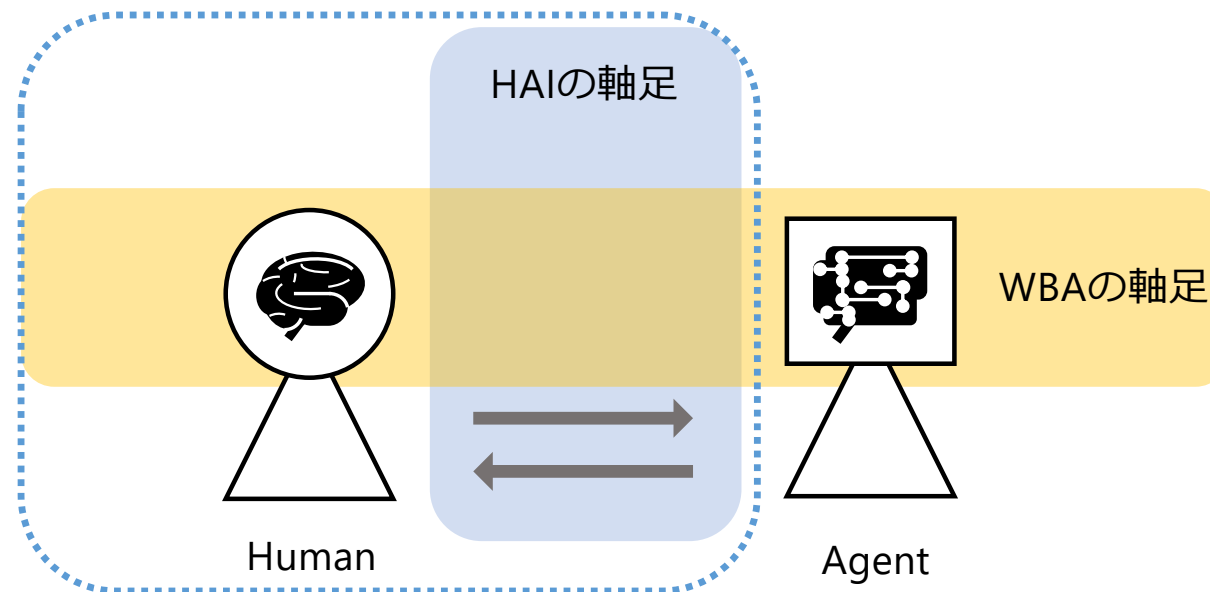
（※ ただしHAI研究が必ずしもボトムアップのアプローチというわけではない...むしろ少数?）



比較
は
重
要

HAI→WBAの距離（人を理解する）

人の認知プロセスを理解するという視点



HAI→WBAの距離（人を理解する）

人の認知プロセスを理解するという視点

HAI：インタラクションにおける人の行動や心の状態の変化を予測する

WBA：人の脳内で起きている情報の流れを予測する

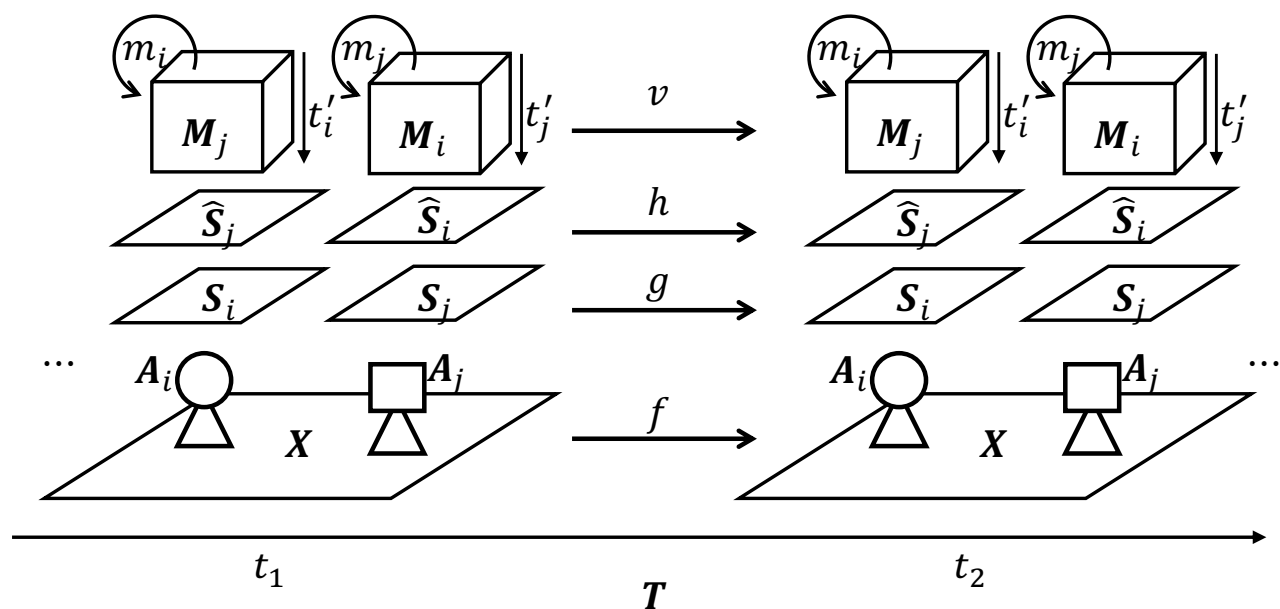
人の脳内で起きている情報の流れを予測することで、

インタラクションにおける人の行動や心の状態の変化を予測できるか？

→ 人の認知プロセスも含めてインタラクションの構造としてモデル化してしまえば
（人の）脳を仮定する必要はない？

HAI→WBAの距離（人を理解する）

認知的インタラクションのフレームワーク



$$\mathcal{S}_3 = \langle A, X, S, \hat{S}, \mathcal{M}, T, f, g, h, v \rangle$$

HAI→WBAの距離（人を理解する）

認知的インタラクションのフレームワーク

A : エージェントの特性

X : 環境

S : 内部状態

$\hat{S}$: 相手の内部状態の推定値

$\mathcal{M}$: 他者モデルを表すサブシステム

T : 時間間隔

f : 行動を決定する関数

g : 内部状態の変化を決定する関数

h : 相手の内部状態を推定する関数

v : $\mathcal{M}$ の状態空間を更新する関数

$$\mathcal{M} = \langle A', M, T', m \rangle$$

A' : エージェントが想定する相手エージェント

M : エージェントが想定する状態空間

T' : 推論の時間間隔

m : 推論のプロセス $m: A' \times M \times T' \rightarrow M$

$$f: A \times X \times S \times \hat{S} \times M \times T \rightarrow X$$

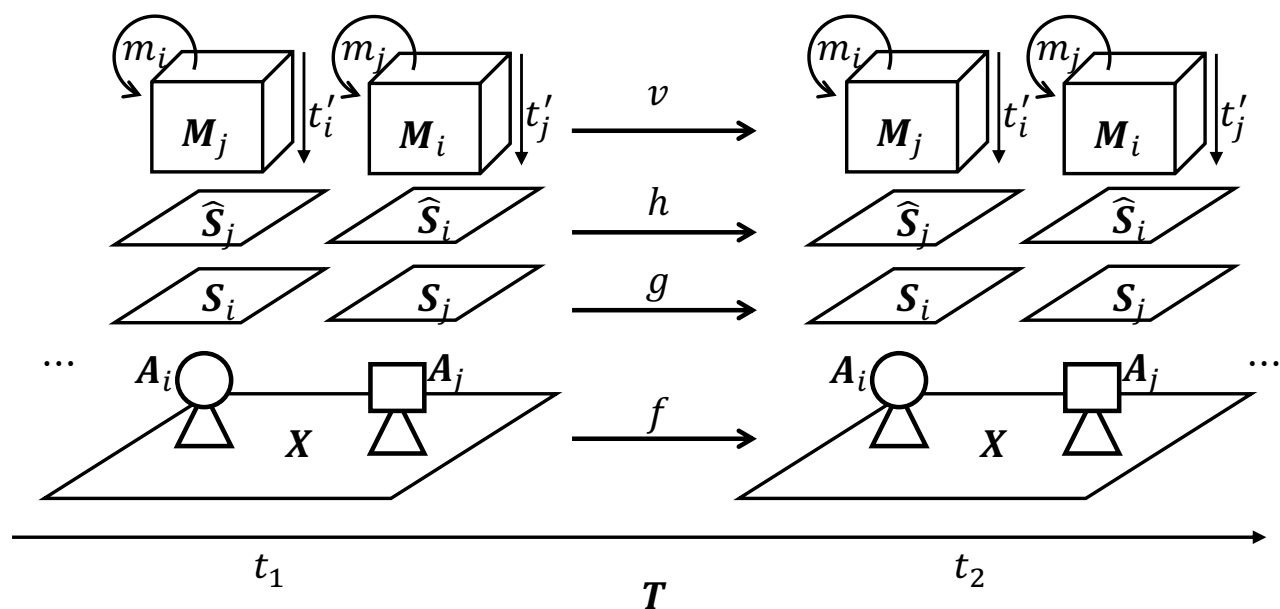
$$g: A \times X \times S \times \hat{S} \times M \times T \rightarrow S$$

$$h: A \times X \times S \times \hat{S} \times M \times T \rightarrow \hat{S}$$

$$v: A \times X \times S \times \hat{S} \times M \times T \rightarrow M$$

HAI→WBAの距離（人を理解する）

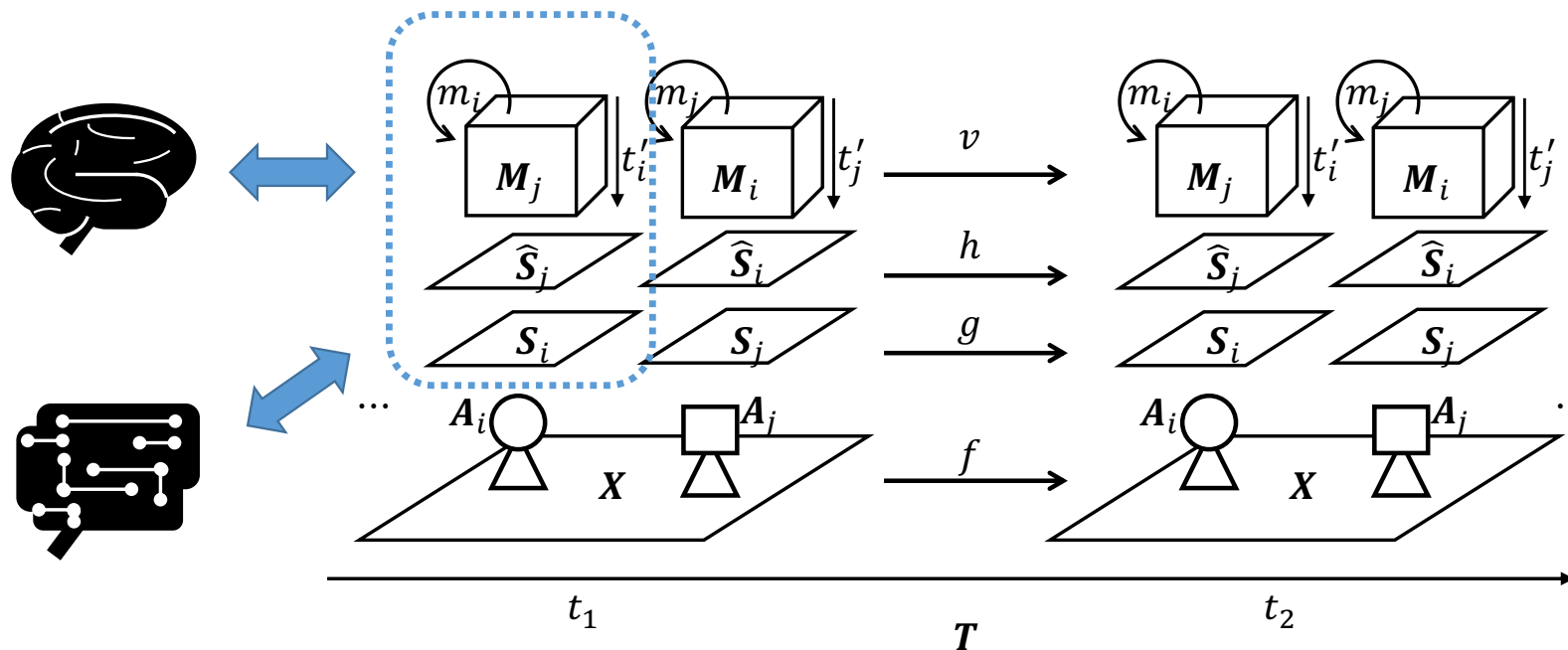
認知的インタラクションのフレームワーク



$$\mathcal{S}_3 = \langle A, X, S, \hat{S}, \mathcal{M}, T, f, g, h, v \rangle$$

HAI→WBAの距離（人を理解する）

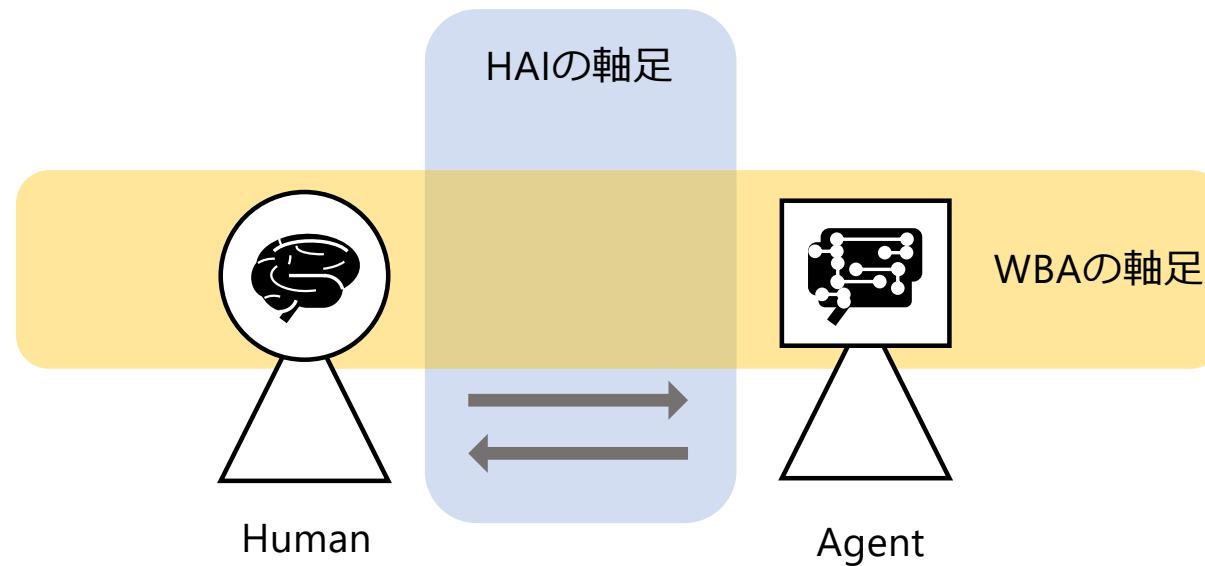
認知的インタラクションのフレームワーク



認知的インタラクションのフレームワークを考えるうえで人の脳との対応関係は検討しておく必要がある
⇒ WBAの成果物との比較も重要

HAI→WBAの接点を探る

人の認知プロセスを理解するという視点



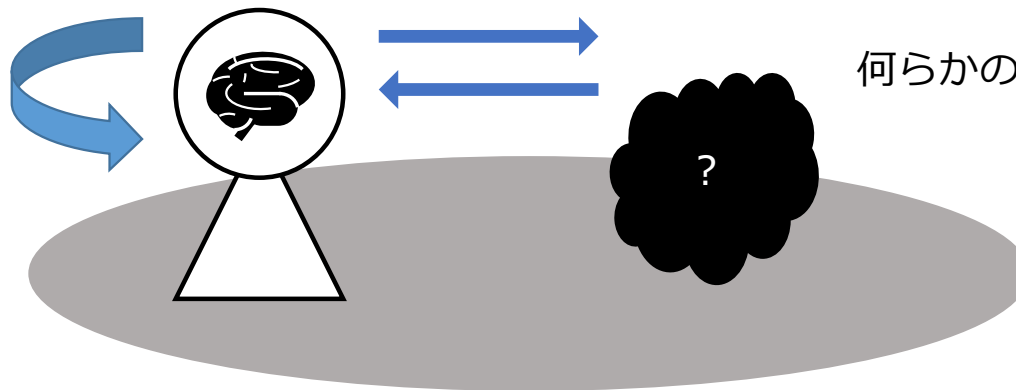
HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

人が対人的なコミュニケーションをできるようになった
(進化的な) プロセスを考える

→ 何らかの環境において適応的だったため

脳 (認知的な機能)
が進化



何らかの環境で
何らかの相手とインタラクション

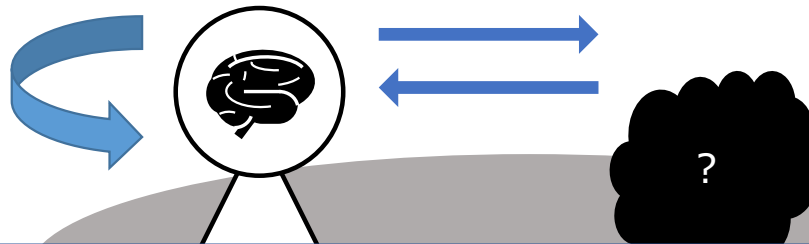
HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

人が対人的なコミュニケーションをできるようになった
(進化的な) プロセスを考える

→ 何らかの環境において適応的だったため

脳（認知的な機能）
が進化



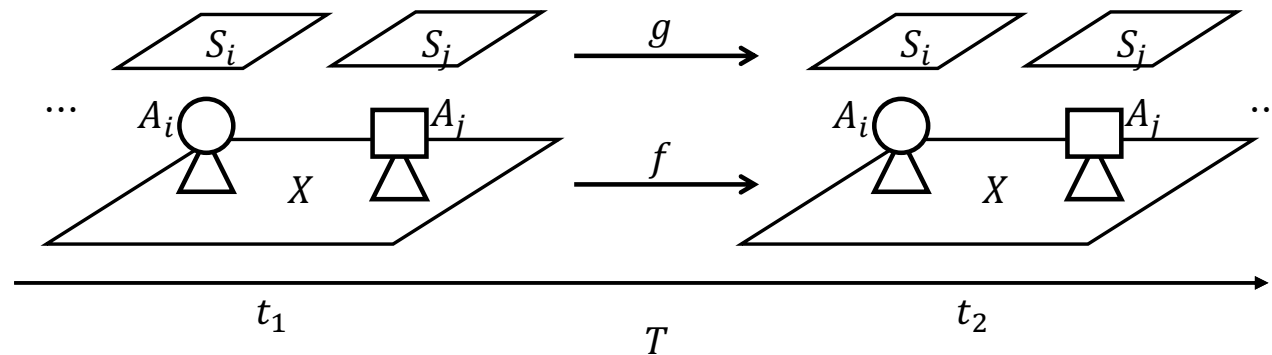
何らかの環境で
何らかの相手とインタラクション

人が現在の脳の構造（認知的な機能）を獲得した理由を探る問題
「どのような環境，どのような相手とのインタラクションにおいて，人の脳が適応的であったか」

HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

ボトムアップに

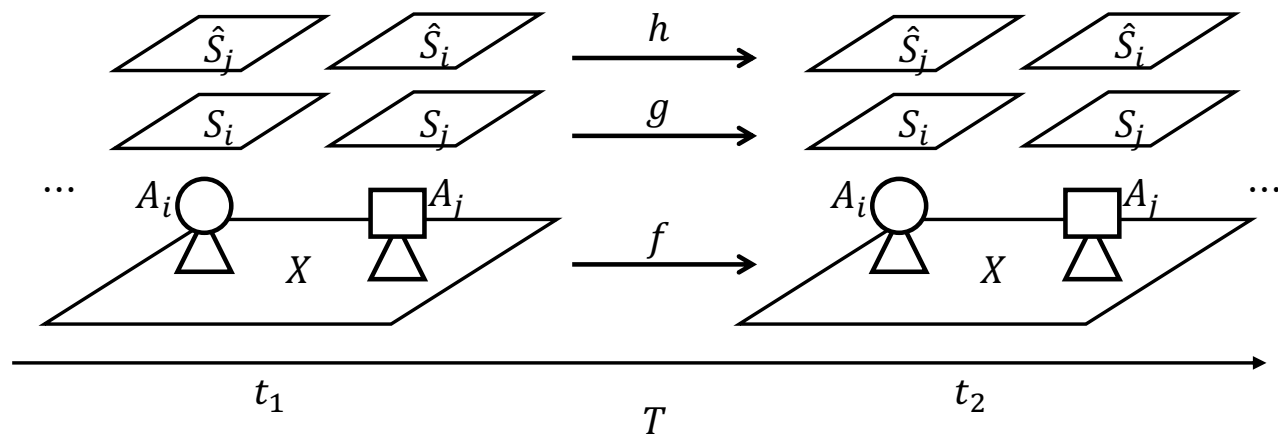


$$\mathcal{S}_1 = \langle A, X, S, T, f, g \rangle$$

HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

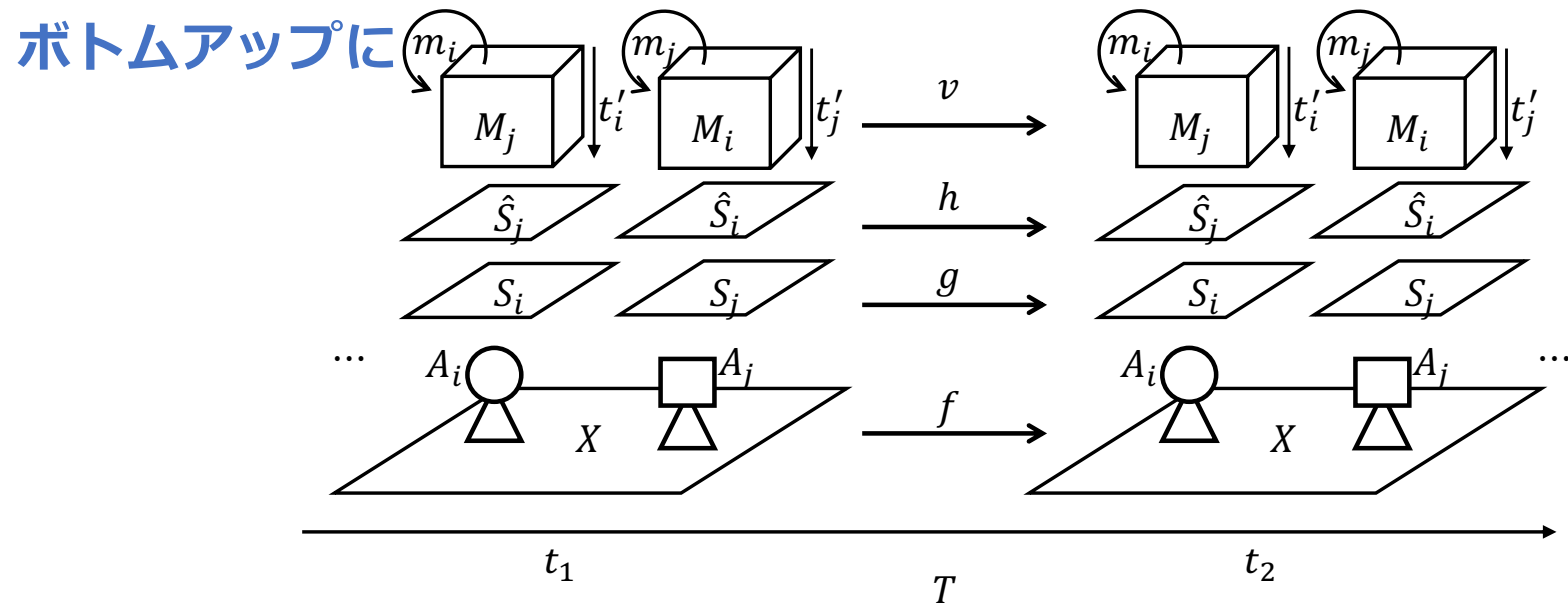
ボトムアップに



$$S_2 = \langle A, X, S, \hat{S}, T, f, g, h \rangle$$

HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

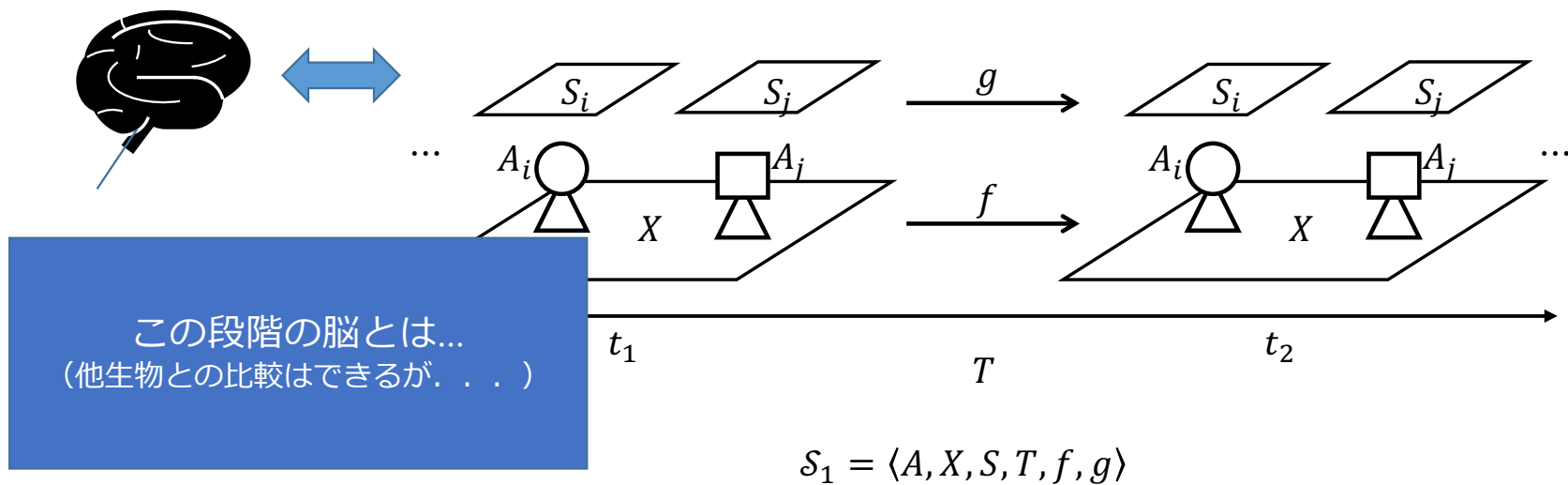


$$\mathcal{S}_3 = \langle A, X, S, \hat{S}, \mathcal{M}, T, f, g, h, v \rangle$$

HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

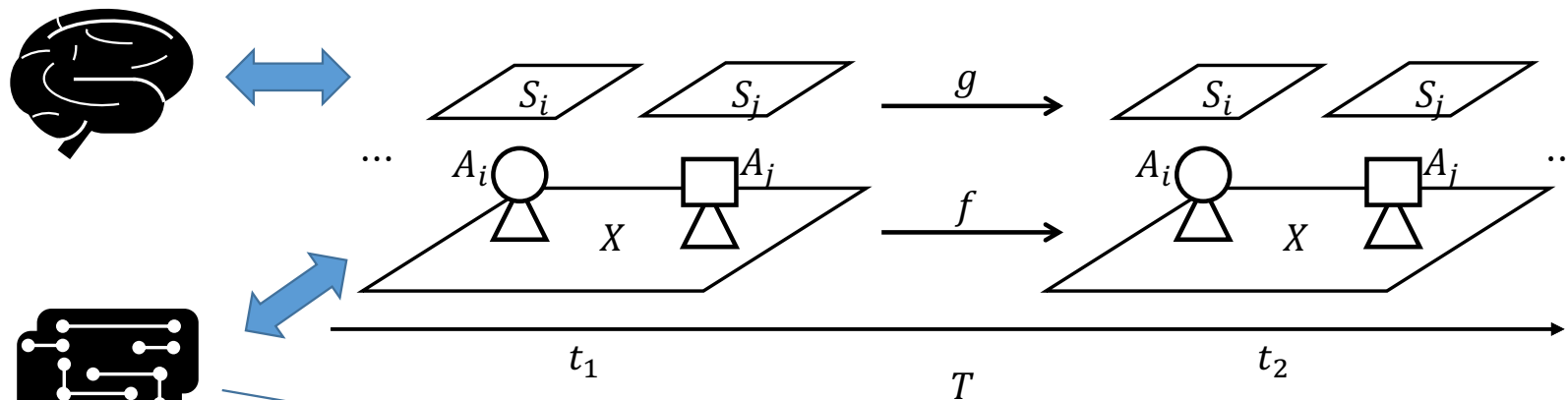
ボトムアップに



HAI→WBAの接点を探る

人の認知プロセスを理解するという視点

ボトムアップに



アーキテクチャなので機能を間引くことができるはず
→ (進化的な) 獲得プロセスを追える

まとめ

HAI研究の視点から考えた場合. . .

人工物の設計という意味では脳を仮定する必要はないかもしれない

人の認知プロセスを明らかにするという点では,

HAIの視点とWBAの視点の両方が必要である
