

第31回 全脳アーキテクチャWBA勉強会
2020年10月23日(金) 18:00-21:00

自由エネルギー原理から エナクティヴィズムへ

北海道大学
人間知・脳・AI研究教育センター
特任准教授 吉田 正俊

(10/26スライド追加版。
追加したスライドには「追加」と表示。)



今日の講演の流れ

1) FEPの概念の深化

ver.1,2,3,4 FEPの解説

2) 自律性からエナクティヴィズムへ

オートポイエーシス

エナクティヴィズム

1) FEPの概念の深化

2) 自律性からエナクティヴィズムへ

1) FEPの概念の深化

2) 自律性からエナクティヴィズムへ

FEP ver.1: 知覚と予測符号化

Friston K. (2005) A theory of cortical responses.
Philos Trans R Soc Lond B Biol Sci. 360(1456):815-36

FEP ver.2: 能動的推測 Active inference

Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. (2010) Action and behavior: a free-energy formulation. Biol Cybern. 102(3):227-60.

FEP ver.3: 期待自由エネルギー

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. (2015) Active inference and epistemic value. Cogn Neurosci. 6(4):187-214.

FEP ver.4: マルコフブランケット

Friston K. (2013) Life as we know it. J R Soc Interface. 10(86):20130475.

FEP ver.1: 知覚と予測符号化

Friston K. (2005) A theory of cortical responses.
Philos Trans R Soc Lond B Biol Sci. 360(1456):815-36

FEP ver.2: 能動的推測 Active inference

Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. (2010) Action and behavior: a free-energy formulation. Biol Cybern. 102(3):227-60.

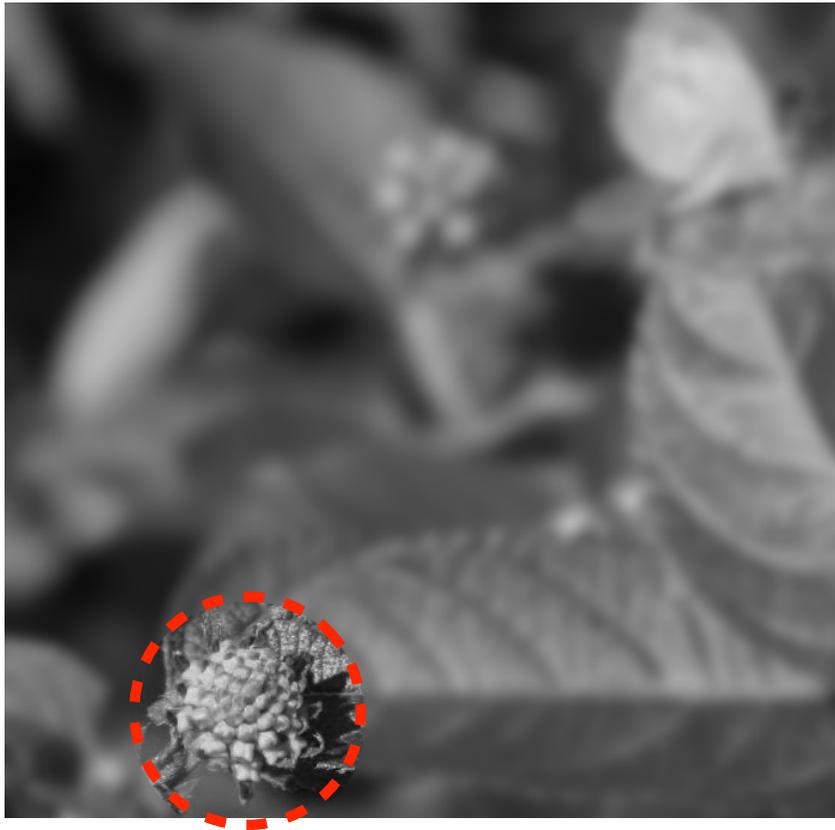
FEP ver.3: 期待自由エネルギー

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. (2015)
Active inference and epistemic value. Cogn Neurosci. 6(4):187-214.

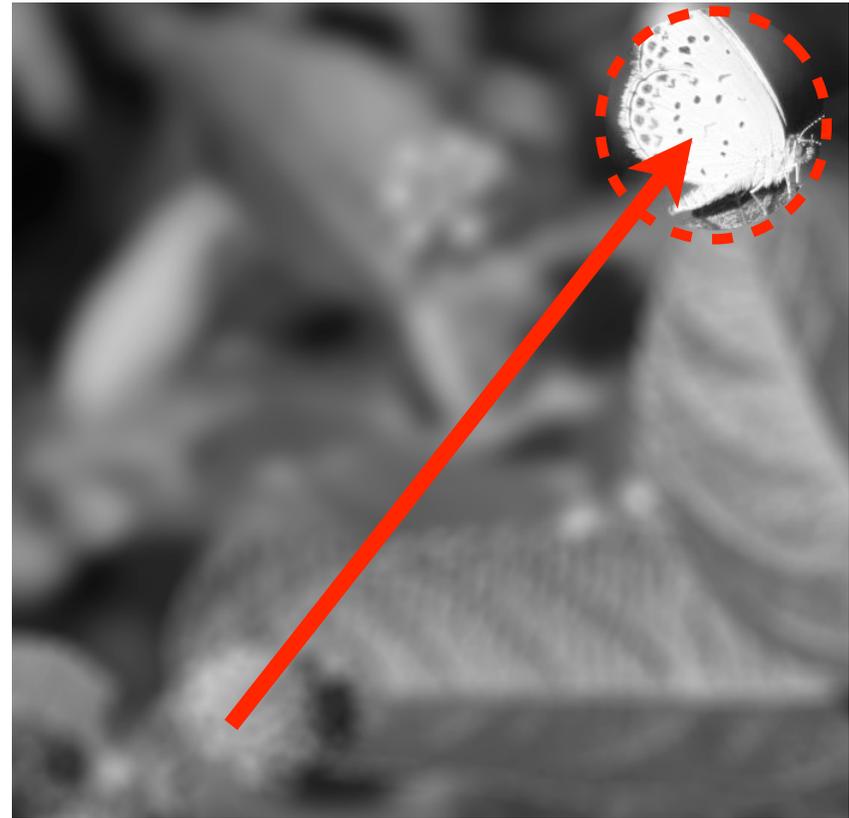
FEP ver.4: マルコフブランケット

Friston K. (2013) Life as we know it. J R Soc Interface. 10(86):20130475.

本講演で使う具体例

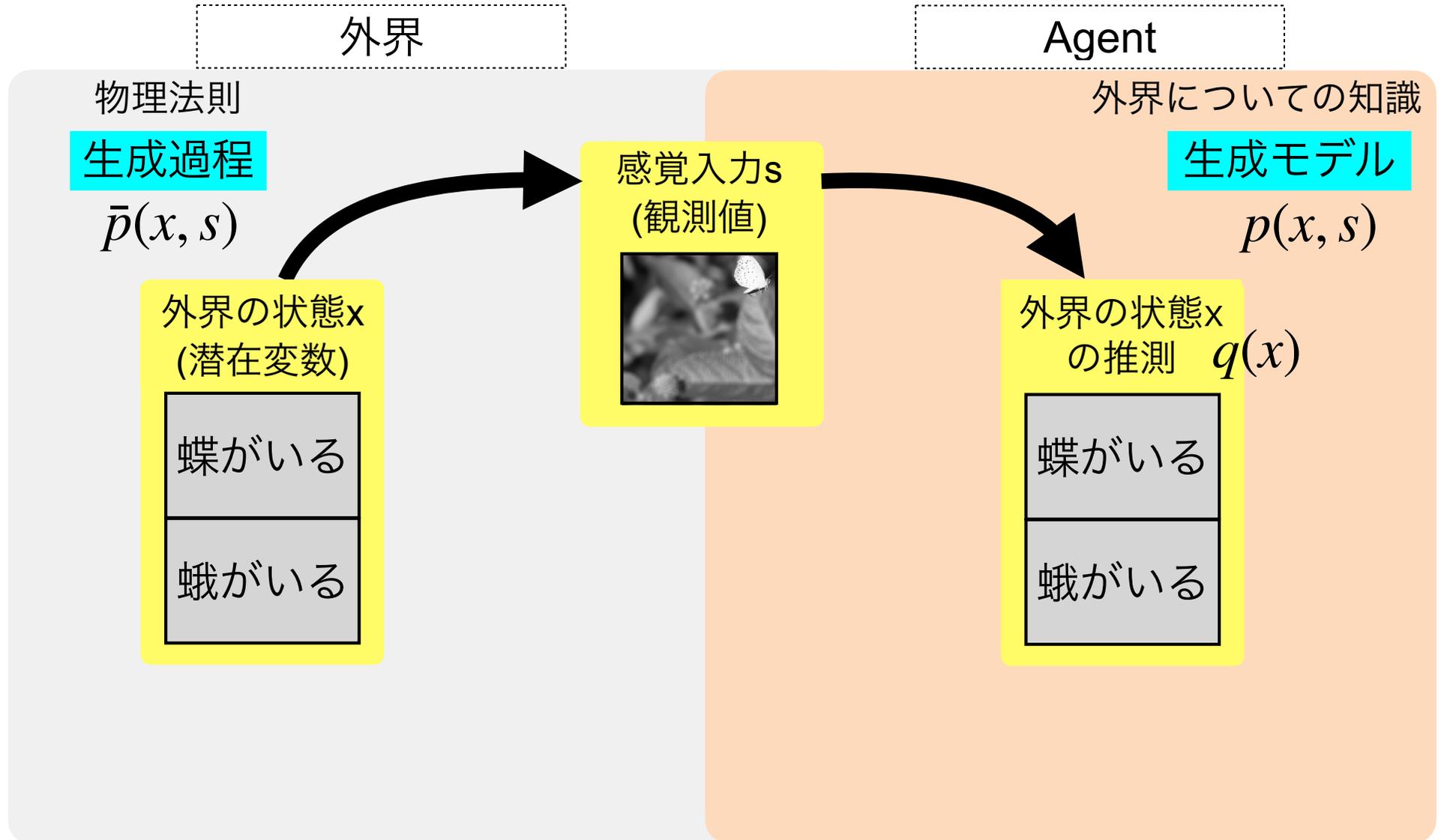


左下に視線があるとき、
右上にあるものは
ぼんやりしていて
なんだかわからない。



視線を右上に向けると、
これが蝶であること
がより確かになった。

知覚のトイモデル



ベイズ脳仮説: 知覚 = 事後分布の推定

フルベイズ

$$q(x) = p(x|s) = \frac{p(x, s)}{\sum_x p(x, s)}$$

推測 事後分布 生成モデル

変分ベイズ

$$q(x) = \operatorname{argmin}_q F(q)$$

変分自由エネルギーF

$$F(q, s) = \sum_x \left\{ q(x) \ln \frac{q(x)}{p(x, s)} \right\}$$

p: 生成モデル(事前分布と事後分布)
q: 推測(x=0,1についての確率分布)
s: 感覚入力(観測データ)

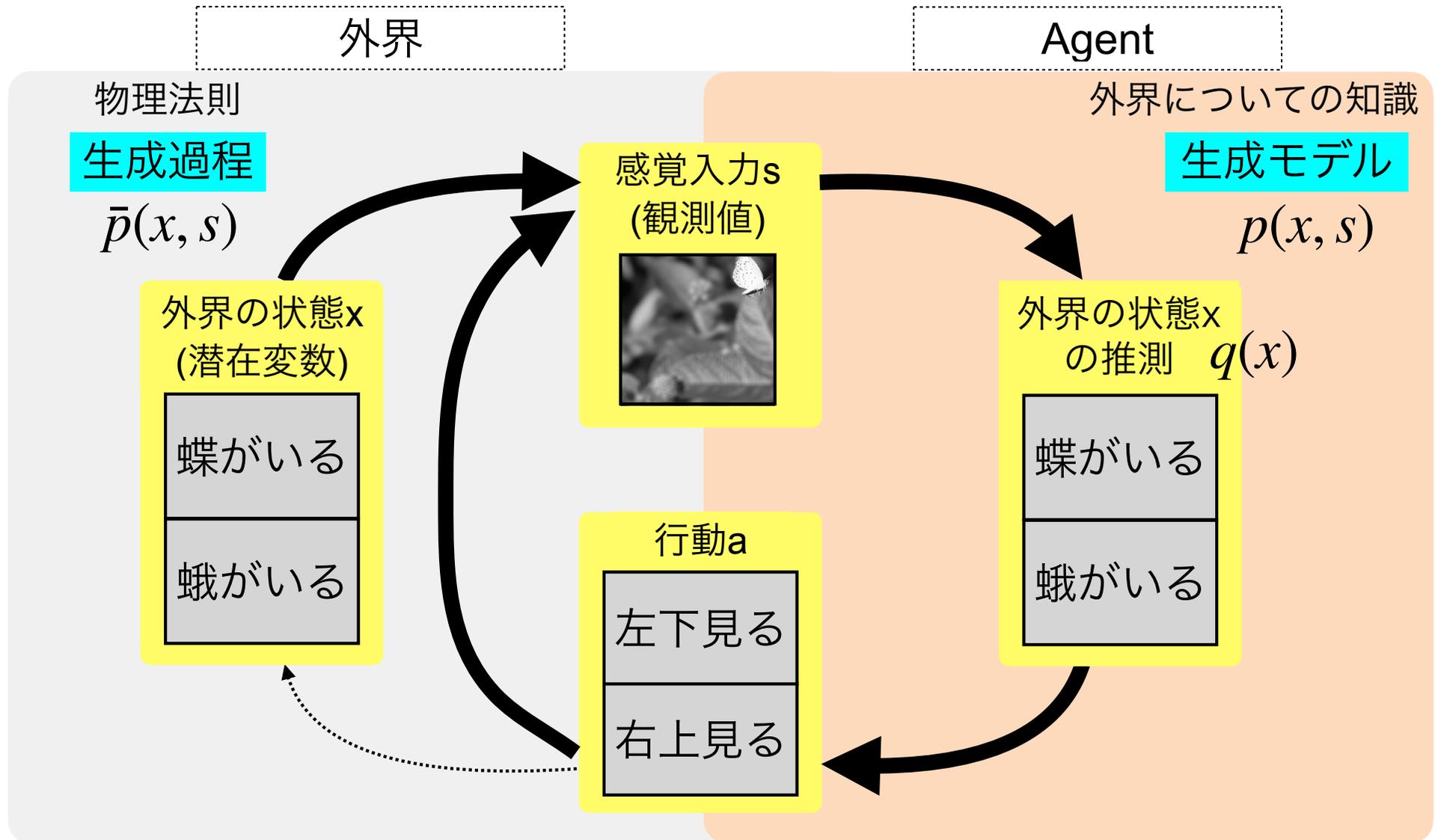
1) 知覚: 推測q(x)を変える

$$F(q) = D_{KL}(q(x) \| p(x|s)) - \ln p(s)$$

KLD 周辺尤度(定数)

q(x)を変える、とは実際には
脳活動 μ を媒介変数とした $q_\mu(x)$ において
脳活動 μ を更新しているということ。

自由エネルギー—原理FEP



知覚と行動の統一的説明

変分自由エネルギーF

$$F(p, q, s) = \sum_x \left\{ q(x) \ln \frac{q(x)}{p(x, s)} \right\}$$

p: 生成モデル(事前分布と事後分布)
q: 推測(x=0,1についての確率分布)
s: 感覚入力(観測データ)

1) 知覚: 推測q(x)を変える

pとsを固定

$$F(q) = D_{KL}(q(x) \| p(x|s)) - \ln p(s)$$

KLD 周辺尤度(定数)

2) 行動: 行動aによって感覚入力sを変える

pとqを固定

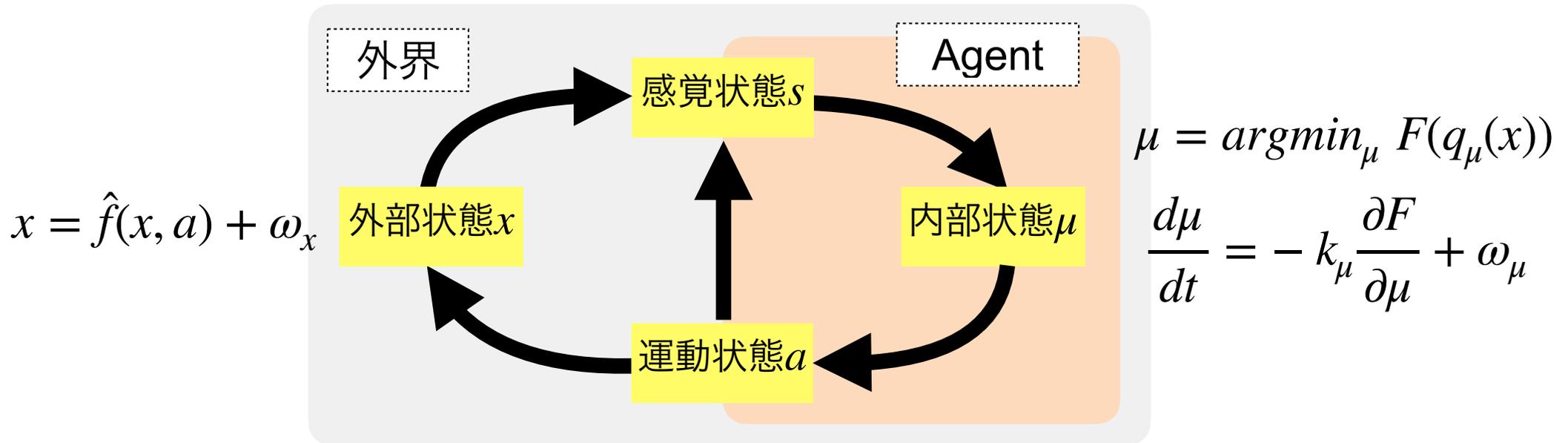
$$F(s) = \sum_x \left\{ -q(x) * \ln(p(s|x)) \right\} + D_{KL}(q(x) \| p(x))$$

Uncertainty Bayesian surprise (定数)

知覚と行動を
変分自由エネルギーFの最小化として統一的に説明

4つの方程式による時間発展

$$s = \hat{g}(x) + \omega_s$$



$$x = \hat{f}(x, a) + \omega_x$$

$$\mu = \operatorname{argmin}_{\mu} F(q_{\mu}(x))$$

$$\frac{d\mu}{dt} = -k_{\mu} \frac{\partial F}{\partial \mu} + \omega_{\mu}$$

$$a = \operatorname{argmin}_{a} F(s(a))$$

$$\frac{da}{dt} = -k_a \frac{\partial F}{\partial a} + \omega_a$$

FEP ver.1: 知覚と予測符号化

Friston K. (2005) A theory of cortical responses.
Philos Trans R Soc Lond B Biol Sci. 360(1456):815-36

FEP ver.2: 能動的推測 Active inference

Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. (2010) Action and behavior: a free-energy formulation. Biol Cybern. 102(3):227-60.

FEP ver.3: 期待自由エネルギー

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. (2015) Active inference and epistemic value. Cogn Neurosci. 6(4):187-214.

FEP ver.4: マルコフブランケット

Friston K. (2013) Life as we know it. J R Soc Interface. 10(86):20130475.

行動aの扱いについての難点

行動aでサンプルされる感覚入力sが一意に決まっているとは限らない。
まだ実現していない「未来での行動の結果」を予測する必要が出てくる。=> 反実仮想

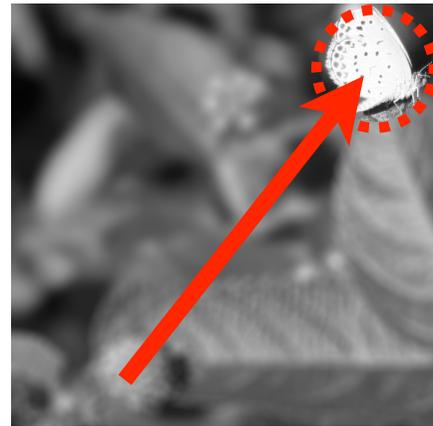
現在: 左下を見ている。
s=1を観測している。

s = 1



将来: 右上を見たときに
観測されるsは不確定

s = 2

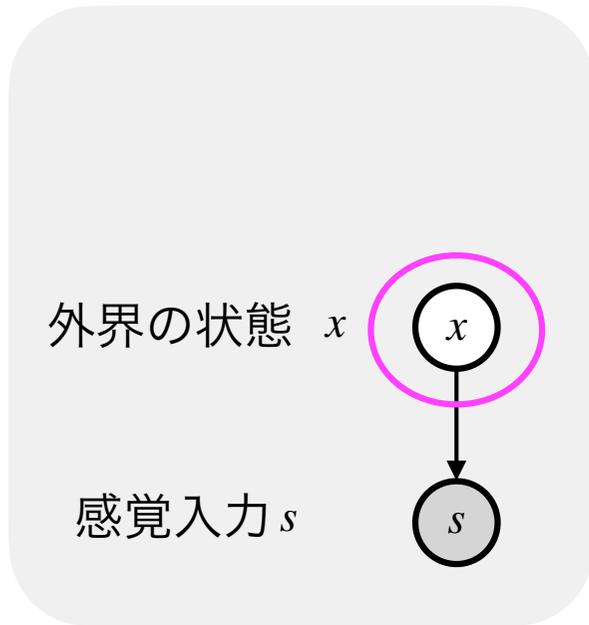


s = 3



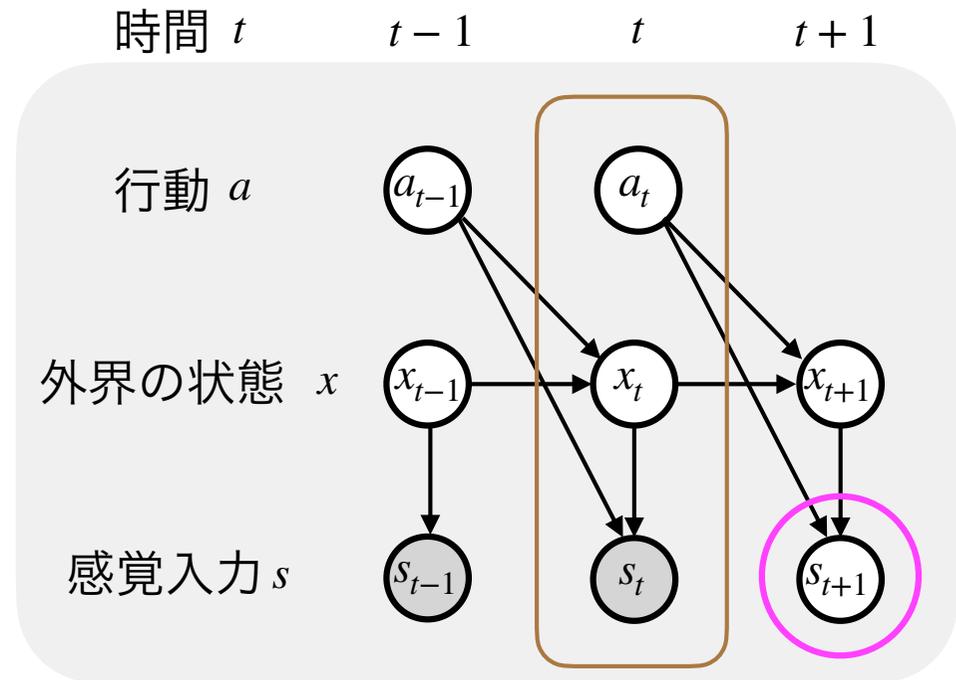
離散時間、部分観測マルコフ決定過程POMDP

知覚の生成モデル (のグラフ表現)



行動の生成モデル

(POMDPとしてのグラフ表現)



- : 観測データ
- : 隠れ状態
- : 因果的關係

まだ実現していない
「未来での行動の結果」の予測
=> 反実仮想

期待自由エネルギーG

変分自由エネルギーF 知覚: 現在の自由エネルギーが最小化する $q(x)$ に更新

$$F(q) = \mathbb{E}_{q(x)}[\ln q(x) - \ln p(x, s)] \quad x \text{の期待値を計算}$$

期待自由エネルギーG 行動: 未来の自由エネルギーの期待値が最小化する行動 a を選択

$$G(a) = \mathbb{E}_{q'_a(x,s)}[\ln q'_a(x) - \ln p_a(x, s)] \quad x \text{だけでなく } s \text{の期待値を計算}$$

$$\begin{aligned} &\simeq - \underbrace{\mathbb{E}_{q'_a(s)} [D_{KL}(q'_a(x|s) \| q'_a(x))]}_{\text{Bayesian surprise}} - \underbrace{\mathbb{E}_{q'_a(x,s)} [\ln p(s)]}_{\text{Instrumental value}} \\ &\quad \quad \quad \parallel \quad \quad \quad \parallel \\ &\quad \quad \text{相互情報量 } MI_a(x'; s') \quad \quad \quad \text{(外的な)報酬} \end{aligned}$$

[相互情報量を最大化]と[望ましい感覚入力 s (ゴール)をサンプル]
これら両方を満たす行動 a が期待自由エネルギー G を最小化する。

(=> “Explore or exploit”の実装)

Tschantz A, Seth AK, Buckley CL. (2020)
Learning action-oriented models through active
inference. PLoS Comput Biol. 16(4):e1007805.

強化学習における目的関数としての期待自由エネルギー G

(=> 「世界モデル」 「内発的動機づけ」との関連)

報酬に依存しない情報探索のモデル化

追加

実際に行動してみないと、観測される s はわからない。
しかし、あるactionでとりうる s の期待値から、
そのactionで期待される相互情報量MIが計算できる。

現在: 左下を見ている。
 $s=1$ を観測している。

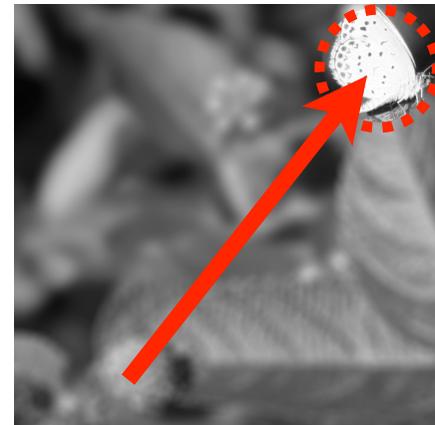
$s = 1$



現在のこの部分の
視覚サリエンス

将来: 右上を見たときに
観測される s の予測

$s = 2$



$s = 3$



a_1 (右上を見る)という行動をした
ときに期待される $MI(a_1)$

Friston K, Adams RA, Perrinet L, Breakspear M. (2012) Perceptions as hypotheses: saccades as experiments. Front Psychol. 3:151.

視覚探索を自由エネルギー最小化によって説明できた

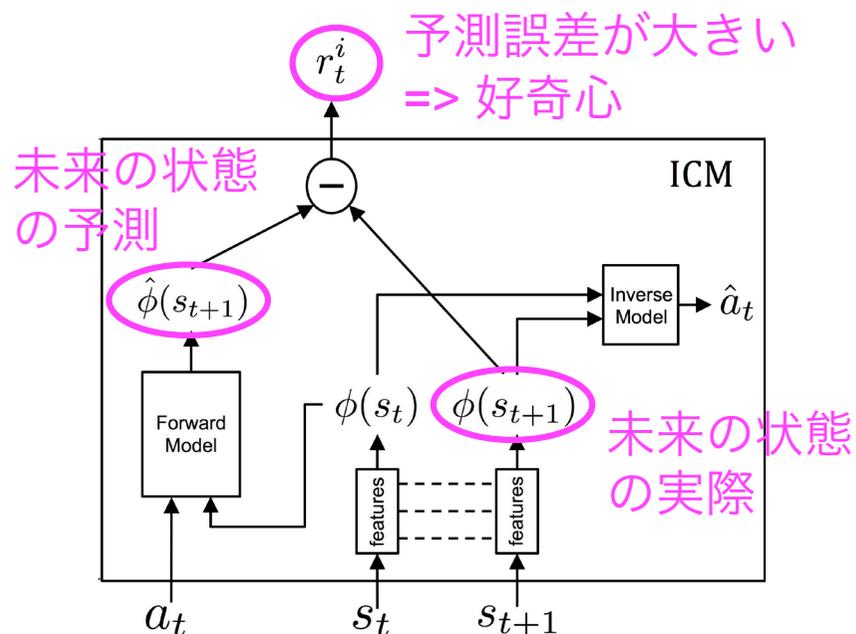
生成モデル+強化学習モデルとの類似

追加

内発的動機づけ/好奇心

強化学習では、報酬の獲得機会が少ないと価値関数がなかなか更新されない。
=> agentが環境の探索を促進する内的報酬を自ら作り出す

強化学習の基礎と深層強化学習
<https://www.slideshare.net/Shotalmai3/rldsdeepreinforcementlearning>



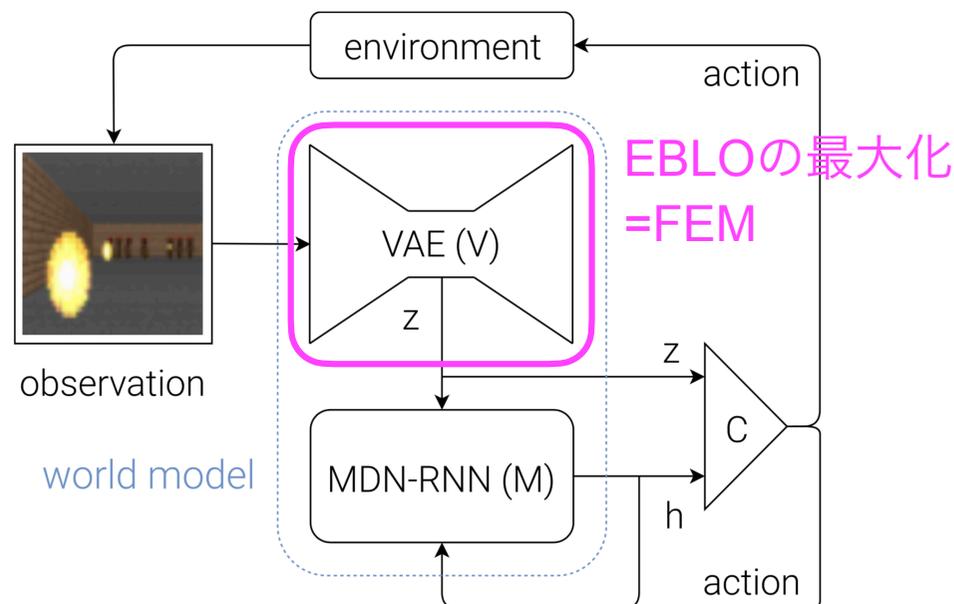
Pathak, Agrawal, Efros and Darrell. Curiosity-driven Exploration by Self-supervised Prediction. ICML 2017.

世界モデル

「世界モデルとは、環境からの限られた情報を元に、環境のモデルを学習によって内部的に構築する枠組みである。...直接には観測できない、過去/未来・反実・観測不能な状態の挙動を把握できるようになり、目的に応じた行動選択の性能を高めることができる。」

JSAI-2020: OS-18 「世界モデルと知能」

<http://www.sig-agi.org/sig-agi/events/jsai2020-os-worldmodel>



Ha, D. & Schmidhuber, J. (2018) World Models. arXiv:1803.10122

期待自由エネルギーG

変分自由エネルギーF 知覚: 現在の自由エネルギーが最小化する $q(x)$ に更新

$$F(q) = \mathbb{E}_{q(x)}[\ln q(x) - \ln p(x, s)] \quad x \text{の期待値を計算}$$

期待自由エネルギーG 行動: 未来の自由エネルギーの期待値が最小化する行動 a を選択

$$G(a) = \mathbb{E}_{q'_a(x,s)}[\ln q'_a(x) - \ln p_a(x, s)] \quad x \text{だけでなく } s \text{の期待値を計算}$$

$$\begin{aligned} &\simeq - \underbrace{\mathbb{E}_{q'_a(s)} [D_{KL}(q'_a(x|s) \| q'_a(x))]}_{\text{Bayesian surprise}} - \underbrace{\mathbb{E}_{q'_a(x,s)} [\ln p(s)]}_{\text{Instrumental value}} \\ &\quad \quad \quad \parallel \quad \quad \quad \parallel \\ &\quad \quad \text{相互情報量 } MI_a(x'; s') \quad \quad \quad \text{(外的な)報酬} \end{aligned}$$

難点

- 常に正確な知覚 $q(x) = p(x|s)$ を仮定
- $\ln p(s)$ を s の事前分布(=ゴール)と読み替える強引な解釈

そもそもこれって「原理」か？

FEP ver.1: 知覚と予測符号化

Friston K. (2005) A theory of cortical responses.
Philos Trans R Soc Lond B Biol Sci. 360(1456):815-36

FEP ver.2: 能動的推測 Active inference

Friston KJ, Daunizeau J, Kilner J, Kiebel SJ. (2010) Action and behavior: a free-energy formulation. Biol Cybern. 102(3):227-60.

FEP ver.3: 期待自由エネルギー

Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. (2015) Active inference and epistemic value. Cogn Neurosci. 6(4):187-214.

FEP ver.4: マルコフブランケット

Friston K. (2013) Life as we know it. J R Soc Interface. 10(86):20130475.

自由エネルギー原理 (FEP)とはなにか

「いかなる自己組織化されたシステムでも、
環境内で平衡状態でありつづけるためには、
そのシステムの(情動的)自由エネルギーを
最小化しなくてはならない」

「適応的なシステムが無秩序へ向かう自然的な傾向
に抗して持続的に存在しつづけるために必要な条件」

Friston, K. (2010). The free-energy principle: a unified brain theory?
Nature Reviews Neuroscience, 11(2), 127–138.

(なお、「平衡」状態は「定常」状態のほうが正しい)

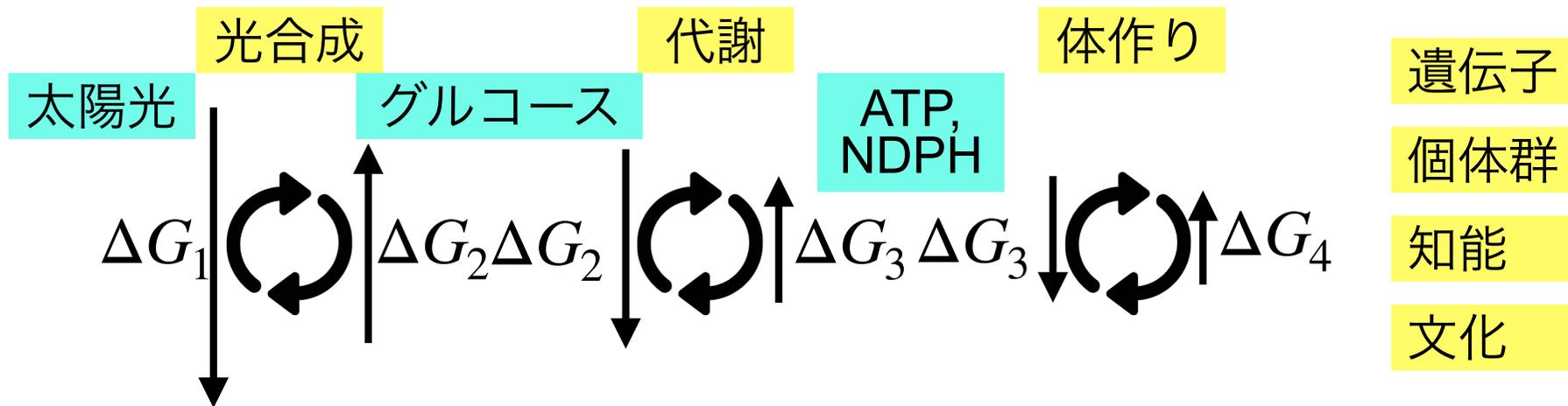
- 「なくてはならない」 => 方向づけ、変分原理
- シュレディンガーの「生命とは何か」への答え

シュレディンガーの「生命とは何か」

- 「生命を持っているものは、崩壊して平衡状態になることを免れている」
 - 「生物体は「負のエントロピー」を食べて生きている」
- (p.147注) 「エントロピーの代わりに自由エネルギーについて論ずべきであった」

『生命とは何か 物理的にみた生細胞』
岡小天・鎮目恭夫共訳、岩波書店〈岩波文庫〉、2008年5月

生物における(化学的な)自由エネルギーの連鎖



参考: 佐藤直樹 「エントロピーから読み解く生物学」 p.184

自由エネルギー原理もこの連鎖に連なることを想定している。

ベイズ脳の実現方法としての自由エネルギー最小化

外界の隠れ値 x の事後分布を推測

ベイズ脳仮説
(hypothesis)

感覚入力 s とその予測
とのあいだの誤差を最小化

予測符号化理論
(theory)

自由エネルギー
原理(principle)

事後分布の
近似的推定法としての
変分ベイズ

自由エネルギー「原理」というよりは「アルゴリズム」

First principleとしての自由エネルギー原理

自由エネルギー最小化の結果、あたかもagentが外界を推測しているかのようにふるまう

ベイズ脳仮説
(hypothesis)

感覚入力sとその予測とのあいだの誤差を最小化

予測符号化理論
(theory)

自由エネルギー
原理(principle)

- 生物が適応的に存在しつづけるためにはagentがどのように変化するかを方向づけ
- 生物にとっての「変分原理」

生物としての安定性

「良い制御器」定理 The good regulator theorem:

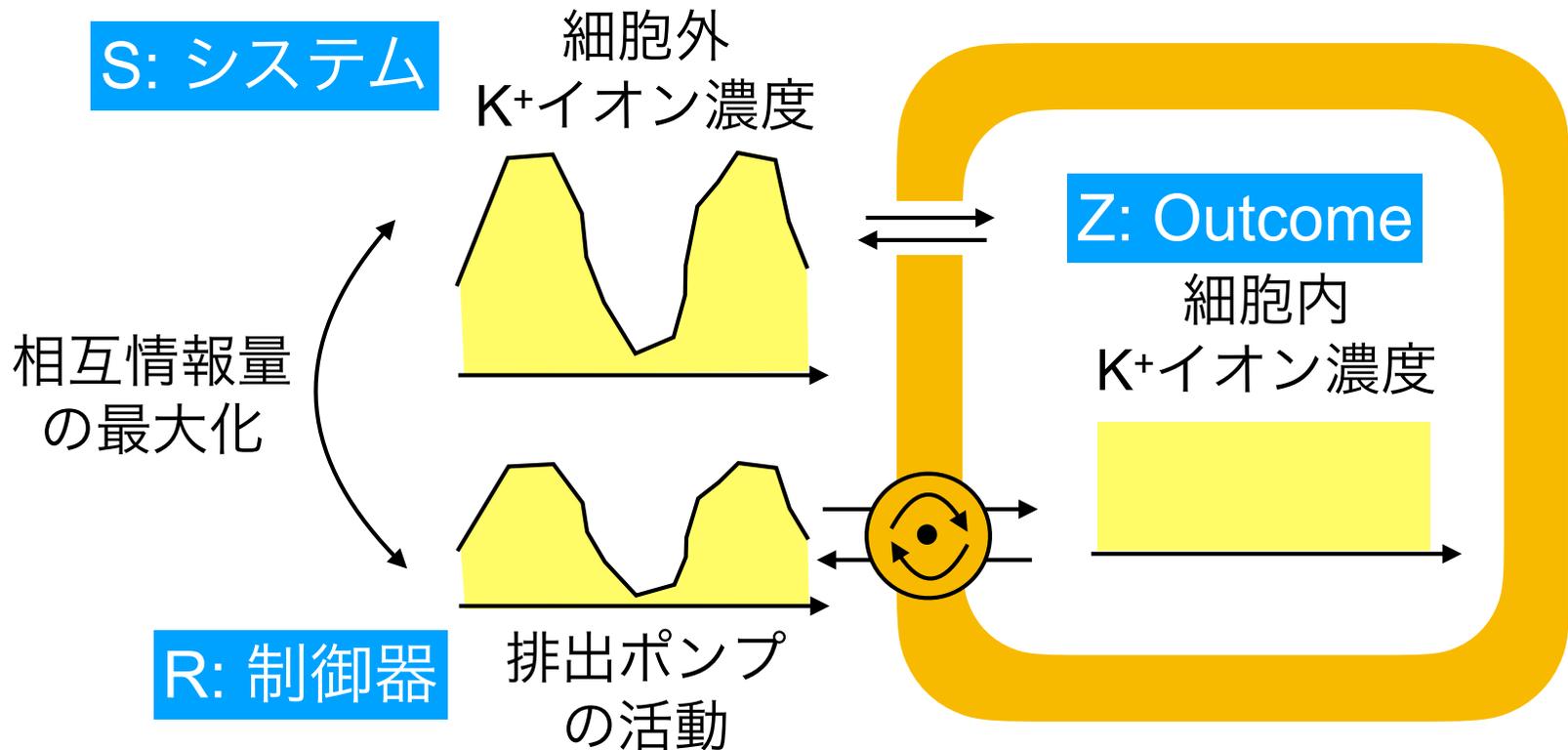
“Every good regulator of a system must be a model of that system.”

R. C. Conant and W. R. Ashby, "Every good regulator of a system must be a model of that system", Int. J. Systems Sci., 1970, vol 1, No 2, pp. 89–97

ある制御器があるシステムを最大限安定的に制御できたとき、
その制御器の状態はそのシステムの状態と同型isomorphicになる

=> 相互情報量の最大化

例: 細胞内 K^+ イオン濃度の制御



生物としての安定性

「良い制御器」定理 The good regulator theorem:
“Every good regulator of a system must be a model of that system.”

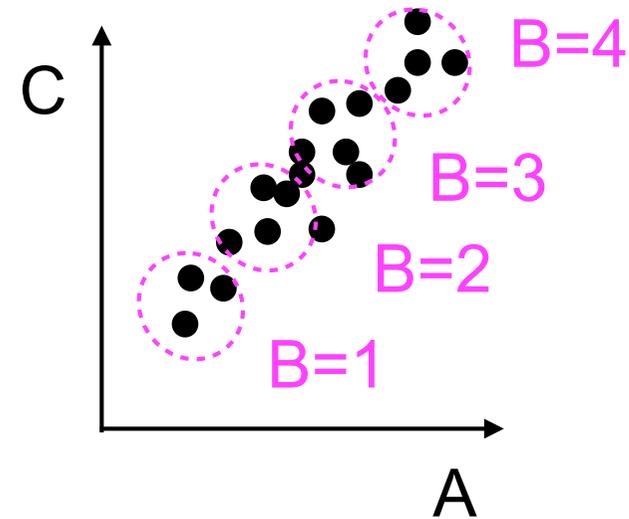
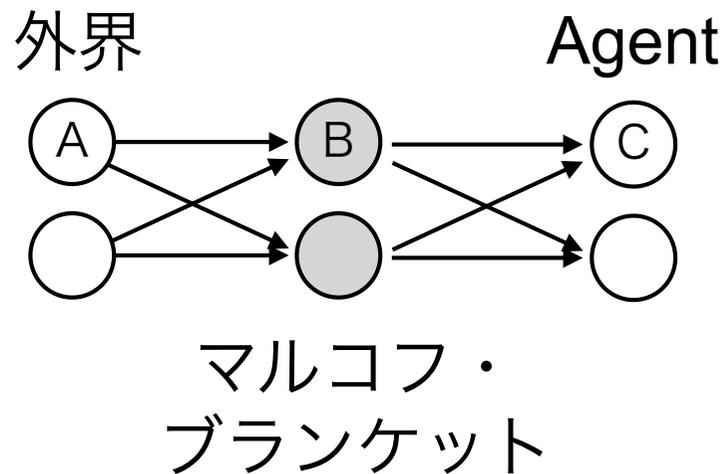
R. C. Conant and W. R. Ashby, "Every good regulator of a system must be a model of that system", Int. J. Systems Sci., 1970, vol 1, No 2, pp. 89–97

ある制御器があるシステムを最大限安定的に制御できたとき、
その制御器の状態はそのシステムの状態と同型isomorphicになる

=> 相互情報量の最大化

- 逆は真ではない。内部と外部の区別がなければ、それは死。
 - 例) 外気と石は温度の点でisomorphicだが、石が外気をモデルしているとは言わない。
 - (細胞膜などによって)外界との区別を保ちながら、それでも内部状態が外部状態と >0 の相互情報量を持つことが生存には必要。
- => 自律性autonomy (「開きながら閉じる」)

マルコフ・ブランケット: ベイズ的自律性



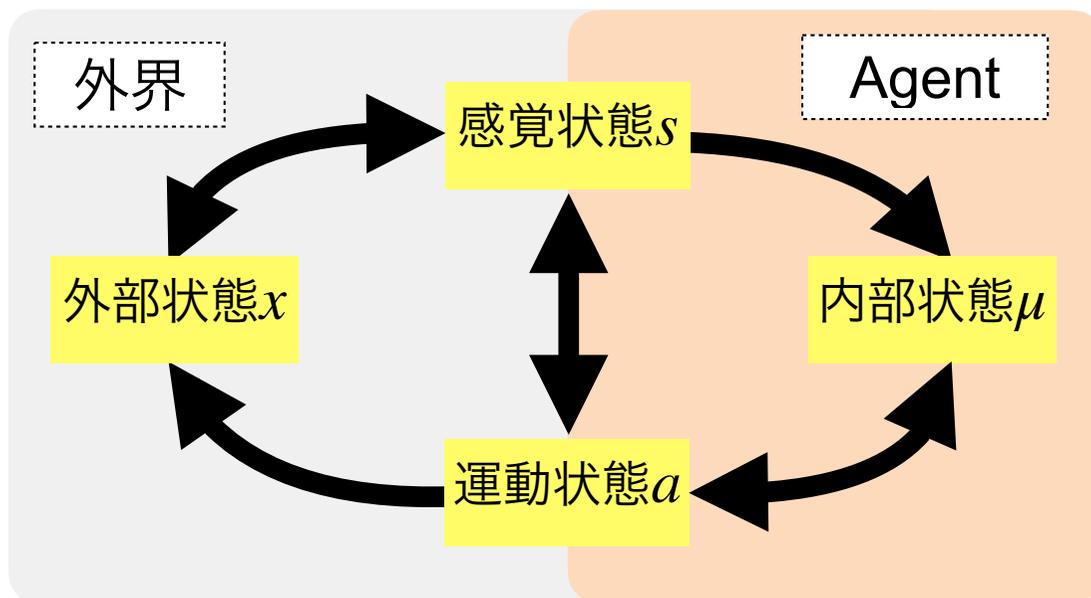
CとAは高い相関だが、Bの値ごとに見ると無相関 = 条件付き独立

$$MI(A; C) > 0 \quad MI(A; C | B) = 0$$

このようなBをマルコフ・ブランケットと呼ぶ。

FEP ver.4: Markov blanket

外界の中で安定して内部状態を保つagentを
ベイジアン因果グラフとマルコフ・ブランケットによって表現

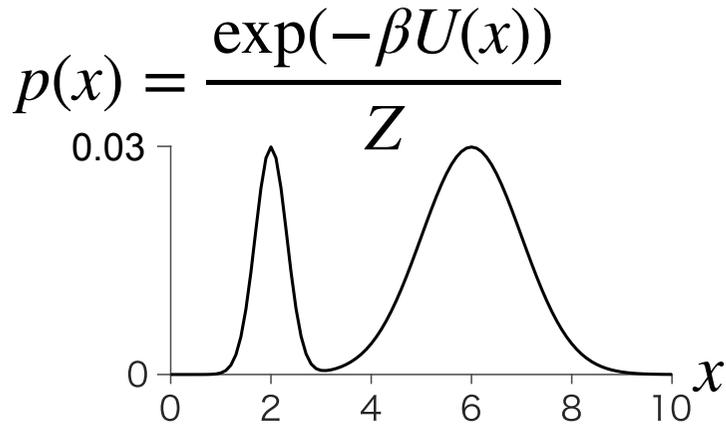


マルコフブランケットによって自律性を持ち、
平衡状態から遠く離れて安定している状態(非平衡定常状態NESS)
においては、変分自由エネルギーが最小化されているとき、
あたかも内部状態が外部状態を推測しているようにふるまう、
というのがver. 4 FEP。

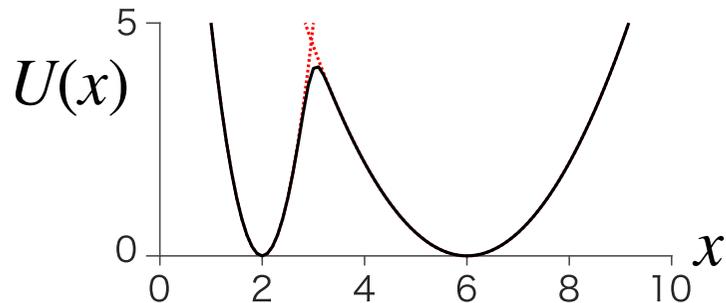
非平衡定常状態NESSの扱い

平衡 (equilibrium)

確率密度



エネルギー地形



$$F = -\frac{1}{\beta} \ln Z$$

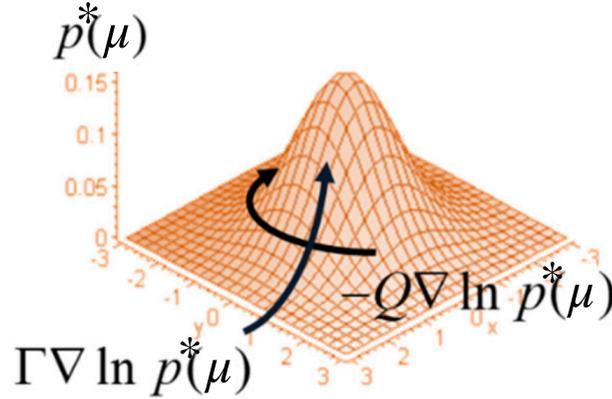
非平衡定常状態 (NESS: non-equilibrium steady state)

Langevin方程式

stateの時間発展

$$\frac{d\mu}{dt} = f_{\mu}(\mu, s, a) + \omega_{\mu}$$

Ergodic density



表現系の安定性:
この分布が時間に
依存せず一定

$$\frac{dp^*(\mu)}{dt} = 0$$

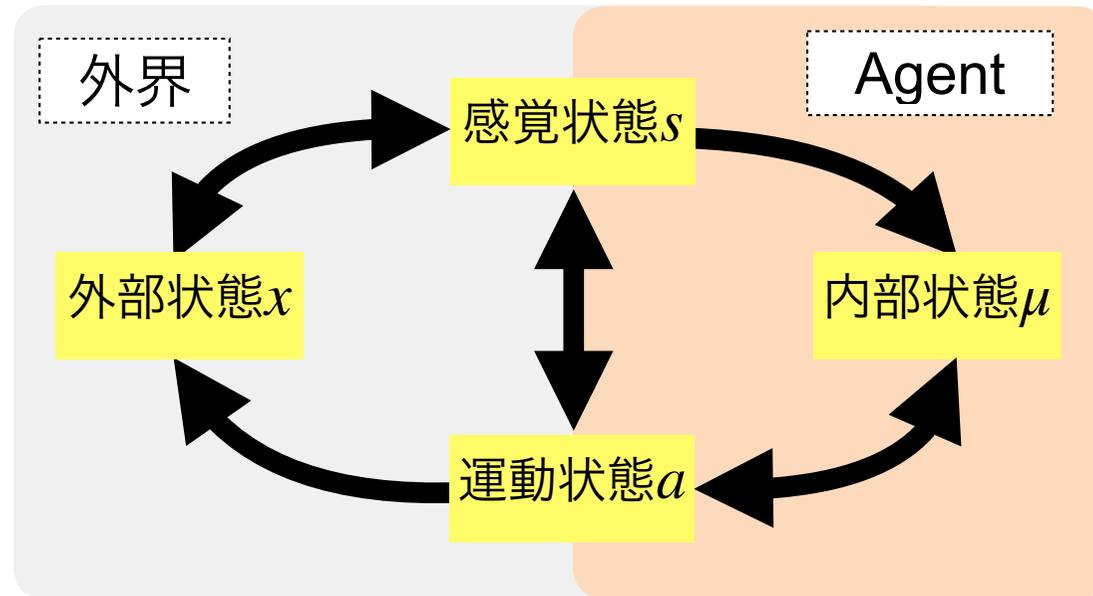
Fokker-Planck方程式

確率分布に
よる表現

$$\frac{dp^*(\mu)}{dt} = \nabla \cdot (\Gamma \nabla - f_{\mu}(\mu)) p^*(\mu)$$

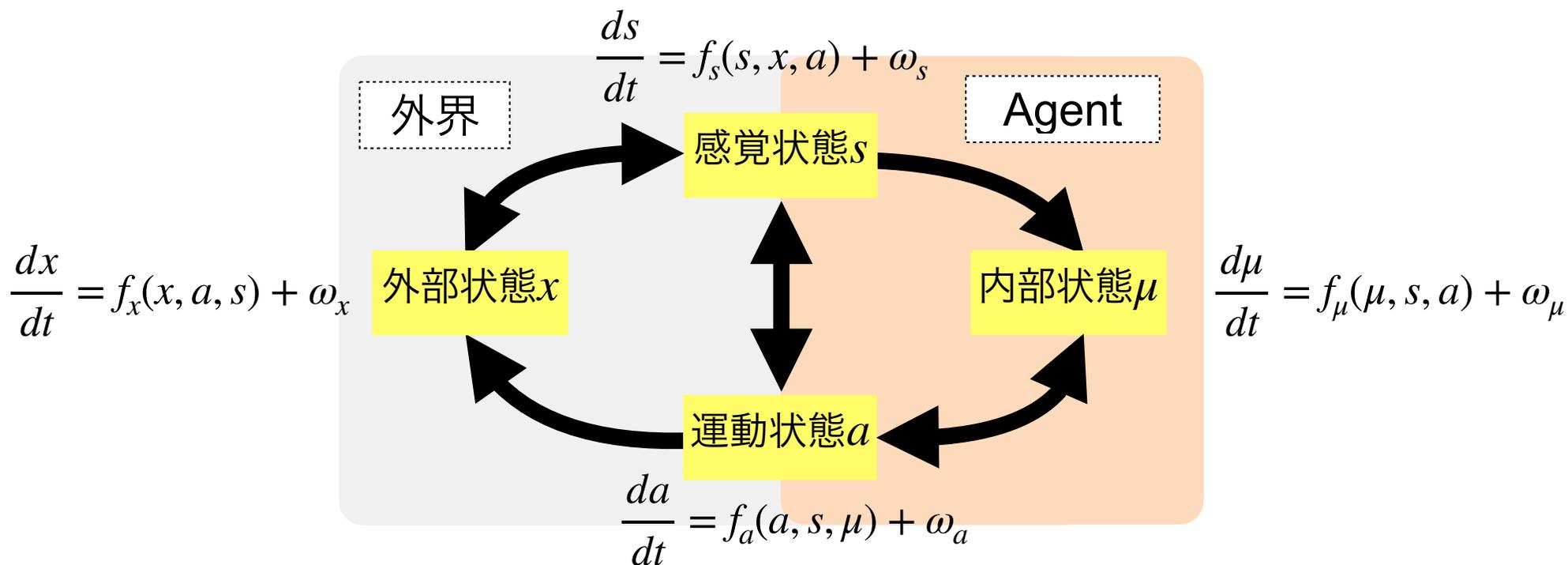
FEP ver.4: Markov blanket

0) マルコフ・ブランケットによって分けられた4つのstate



FEP ver.4: Markov blanket

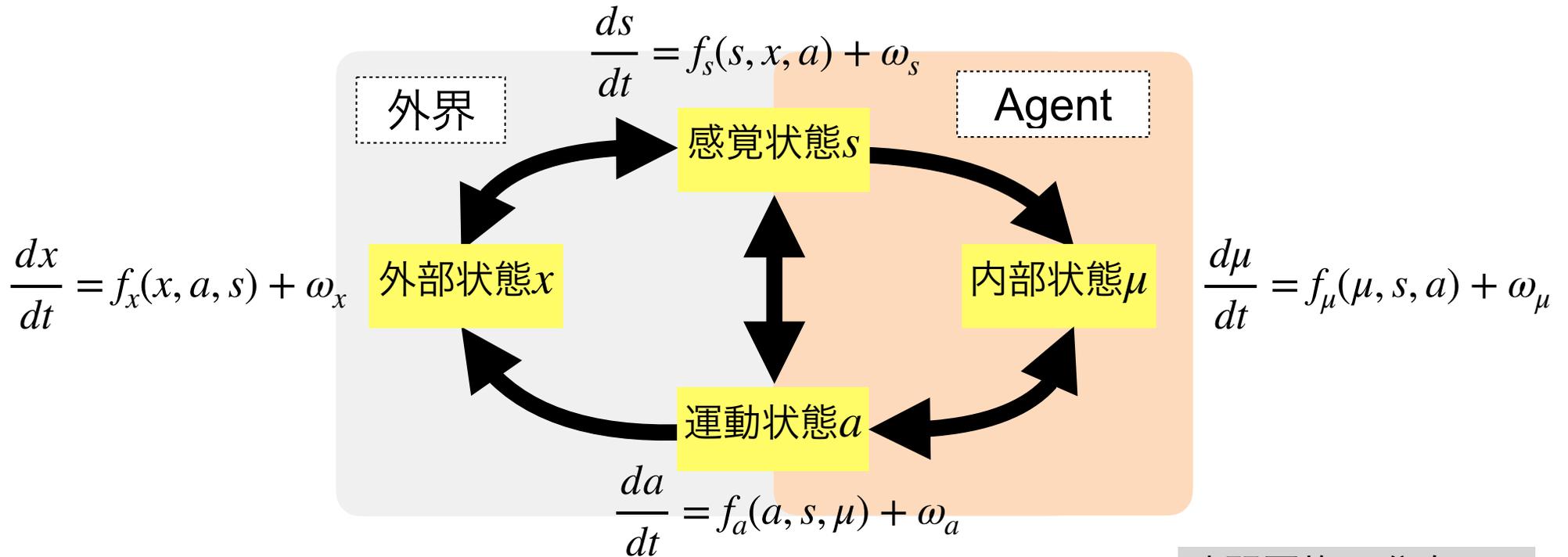
0) マルコフ・ブランケットによって分けられた4つのstate



1) 4つのstateの時間発展(Langevin 方程式)。

FEP ver.4: Markov blanket

1) マルコフ・ブランケットによって分けられた4つのstate

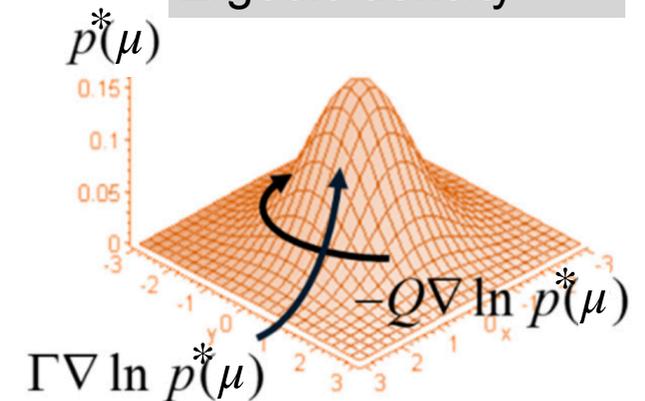


2) 4つのstateの時間発展(Langevin 方程式)。

3) 関数 f_μ を決めるためにFokker-Planck方程式を作る。

$$\frac{dp^*(\mu)}{dt} = \nabla \cdot (\Gamma \nabla - f_\mu(\mu)) p^*(\mu)$$

時間平均の分布
Ergodic density



FEP ver.4: Markov blanket

追加

4) agentが安定して存在している条件: $\frac{dp^*(\mu)}{dt} = 0$

5) この条件を使ってFP方程式を解くと $f()$ が決まる。

$$f_{\mu}(\mu, s, a) = ((\Gamma_{\mu\mu} + R_{\mu\mu}) \cdot \nabla_{\mu}) \ln p^*(\mu, s, a) \quad \leftarrow \text{surprisal}$$

6) 変分ベイズで $q(x | \mu)$ によって事後分布 $p^*(x | \mu, s, a)$ を近似

$$f_{\mu}(\mu, s, a) = \Gamma_{\mu\mu} \nabla_{\mu} F(\mu, s, a) \quad \leftarrow \text{自由エネルギー}$$

7) これを1)の式に入れると $\frac{d\mu}{dt} = \Gamma_{\mu\mu} \frac{\partial F(\mu, s, a)}{\partial \mu}$

8) つまり内部状態 μ は変分自由エネルギー $F(\mu, s, a)$ の勾配が無くなる方向に更新されてゆく。

すべてが客観的に決まることに注意！

FEP ver.4 まとめ

1)

「内」と「外」との
情動的独立性

マルコフ
ブランケット

4)

$\frac{dp^*(\mu)}{dt} = 0$ 非平衡定常状態

表現系が
安定して存在する

5)

サプライズ最小化
による更新則

6)

自由エネルギー最小化
による更新則

$$f_{\mu}(\mu, s, a) = ((\Gamma_{\mu\mu} + Q_{\mu\mu}) \cdot \nabla_{\mu}) \ln p^*(\mu, s, a)$$

(内的な)生成モデル
ではなくて
時間平均の分布
ergodic density
であることに注意

$$f_{\mu}(\mu, s, a) = \Gamma_{\mu\mu} \nabla_{\mu} F(\mu, s, a)$$

$$\frac{d\mu}{dt} = \Gamma_{\mu\mu} \frac{\partial F(\mu, s, a)}{\partial \mu}$$

Ver.4 FEP: 「マルコフ・ブランケットを持っているNESSでは、
内部状態があたかも外部状態の推測をするようにふるまう」

FEP ver.4: Markov blanket

しかし、Biehl et al (2020)によれば、
各ステップは必ずしも成り立たない
(各ステップに反例が見つかる)。
よって、FEP ver.4の主張は正当化できない。

Martin Biehl and Felix A. Pollock and Ryota Kanai (2020) A technical critique of the free energy principle as presented in "Life as we know it" and related works. arXiv 2001.06408

しかし「マルコフブランケットのような自律性の過程から、
推測をするシステムの創発」という道筋は有望そうだ。
後半のトークでは、この「自律性」の概念について
考えを進めてゆくことにする。

1) FEPの概念の深化

2) 自律性からエナクティヴィズムへ

1) FEPの概念の深化

2) 自律性からエナクティヴィズムへ

- オートポイエーシス
- エナクティヴィズム

自律性ってなんだろう？

ベイズネットでの自律性 - マルコフ・ブランケット

状態空間モデルにおける自律性 - Non-trivial information closure

Bertschinger N., Olbrich E., Ay N., Jost J. (2006). Information and closure in systems theory, in Explorations in the Complexity of Possible Life. Proceedings of the 7th German Workshop of Artificial Life, (Amsterdam:) 9–21.

Chang AYC, Biehl M, Yu Y, Kanai R. (2020) Information Closure Theory of Consciousness. Front Psychol. 11:1504.

生命の特性としての自律性 - オートポイエーシス

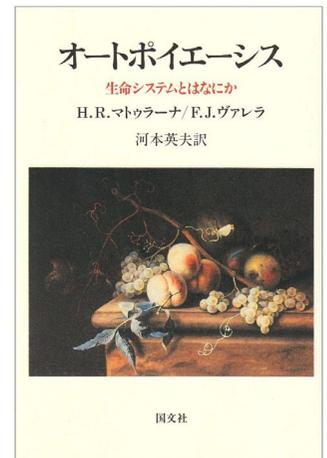
オートポイエーシス: 自律性の実現

生命システムのorganization(組織としての特性)とはなにか？
これに対する答えが「オートポイエーシスautopoiesis」。

Autopoeisis = Auto (自己) + poiesis (制作) = 自己を産出する

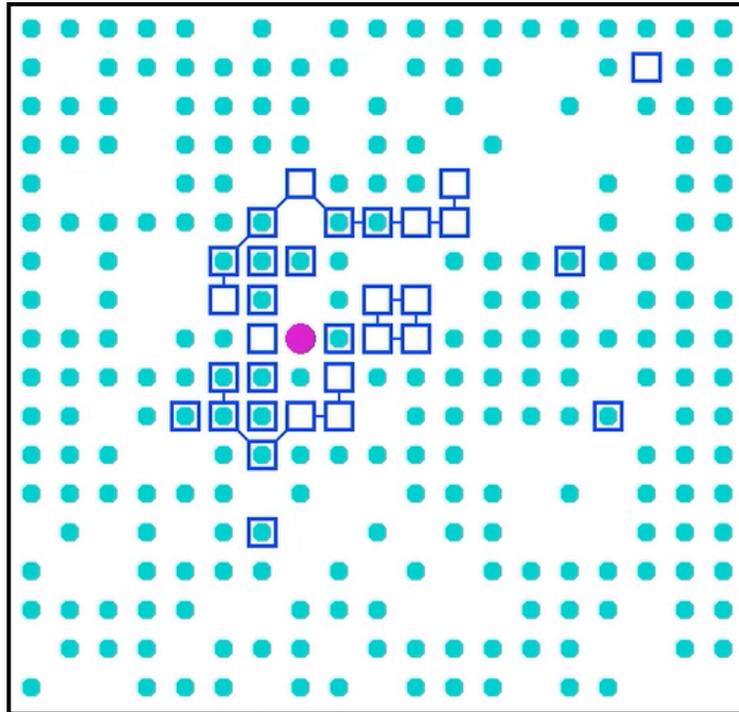
- 自律性 autonomy
- 個性 identity
- 自他の境界の自己産出
- 入出力関係によっては規定されない

マトゥラーナ、バレーラ (1974) 『オートポイエーシス —
生命システムとは何か』 河本英夫訳、国文社 (日本語訳1991)



オートポイエーシスのtoy model: SCLモデル

SCLモデル



細胞での化学反応を模した2D CA

- Substrate基質
- Catalyst触媒
- Linkリンク
- Bonded link

動作ルール

1. Linkの生成



2. Bonded linkの生成



3. Linkの崩壊



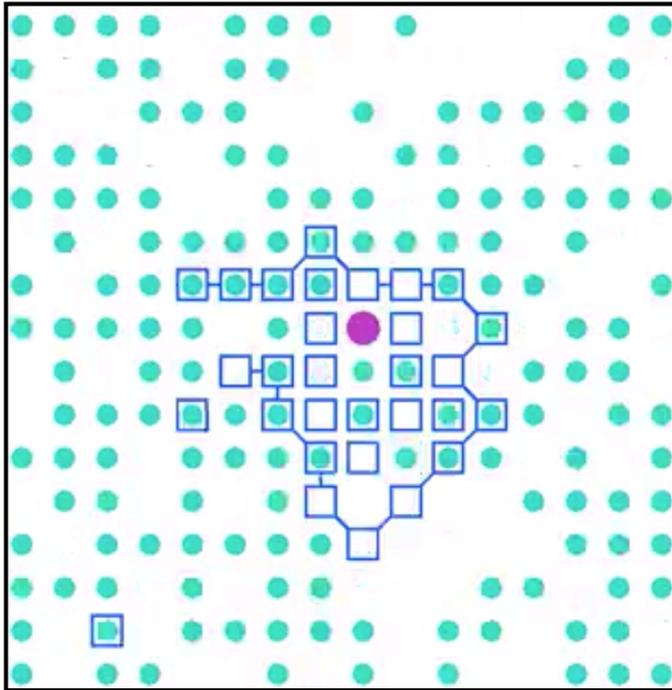
B. McMullin and F. Varela (1997) Rediscovering Computational Autopoiesis. SFI Working Paper 97-02-012

岡 瑞起、池上 高志、ドミニク・チェン、青木 竜太、丸山 典宏 (2018) 作って動かすALife - 実装を通した人工生命モデル理論入門. オライリー・ジャパン

これらが一定の確率で起こる

オートポイエーシス: 自律性の実現

SCLモデル



- 局所的には、化学反応を表現するルールにしたがって動いているだけ。
- 境界が生成され続けることによって、個性を持つ主体が生成される。
- しかしいったん主体が生成されると、主体にとっての内部状態と外界(環境)が解釈可能になる。
- よってこのagentは入力を処理して出力を出すという観点から解釈するのは不十分。

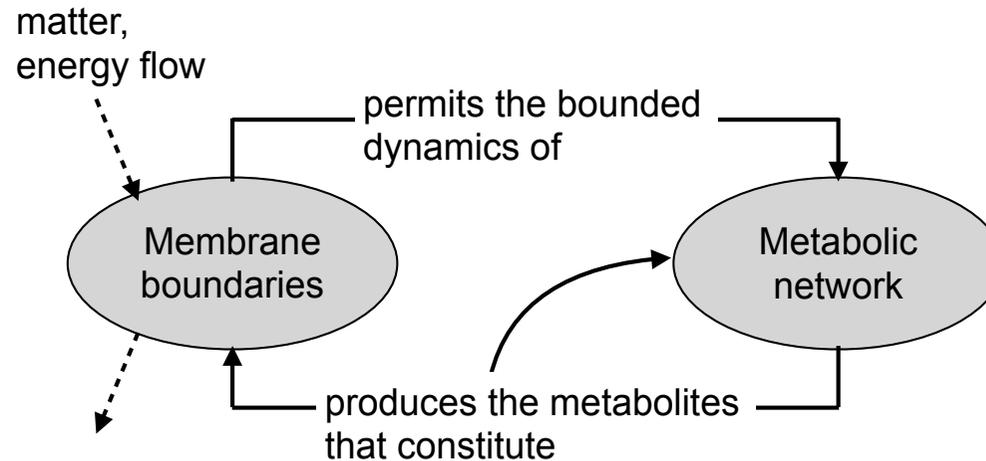
=> 自律性 autonomy
個性 identity
自他の境界の自己産出
入出力関係によっては規定されない

B. McMullin and F. Varela (1997) Rediscovering Computational Autopoiesis. SFI Working Paper 97-02-012

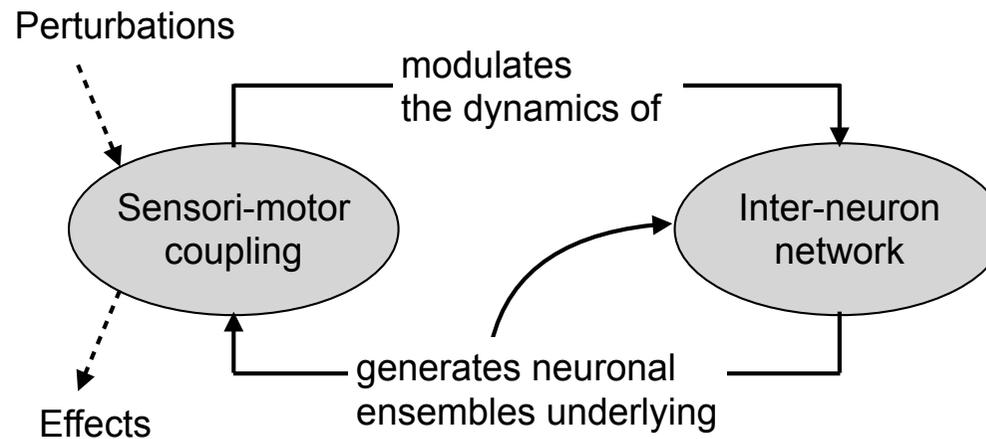
岡 瑞起、池上 高志、ドミニク・チェン、青木 竜太、丸山 典宏 (2018) 作って動かすALife - 実装を通じた人工生命モデル理論入門. オライリー・ジャパン

細胞も(認知する)個体もオートポイエーシス

Life:
Single cell



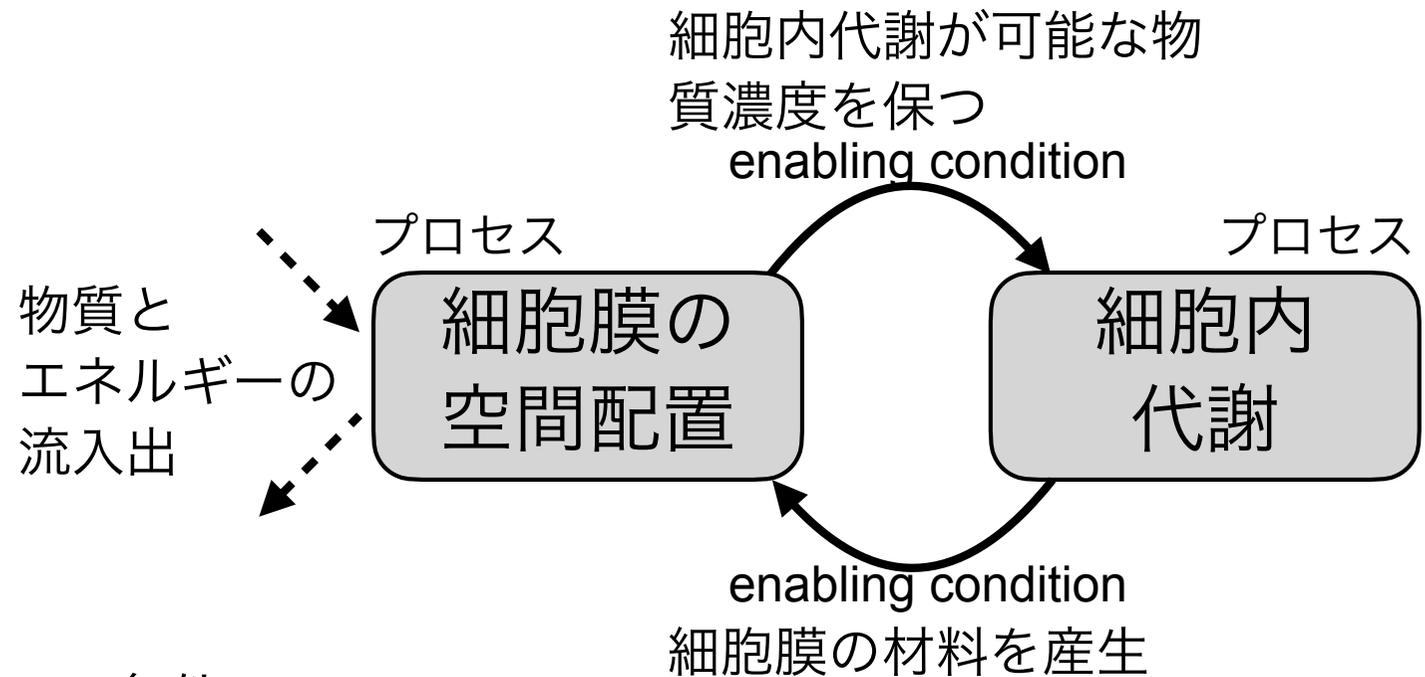
Cognition:
Brain, body and environment



オートポイエーシスとしての細胞

追加

Life:
Single cell



オートポイエーシスの条件:

自律性 autonomy

個性 identity

自他の境界の自己産出

入出力関係によっては規定されない

=> 操作的閉包 + 構造的カップリング
として再規定

操作的に閉じている

operationally closed

= 各プロセスの維持が

他のプロセスに依存している

オートポイエーシスとしての(認知する)個体

“Open loop”

入出力装置

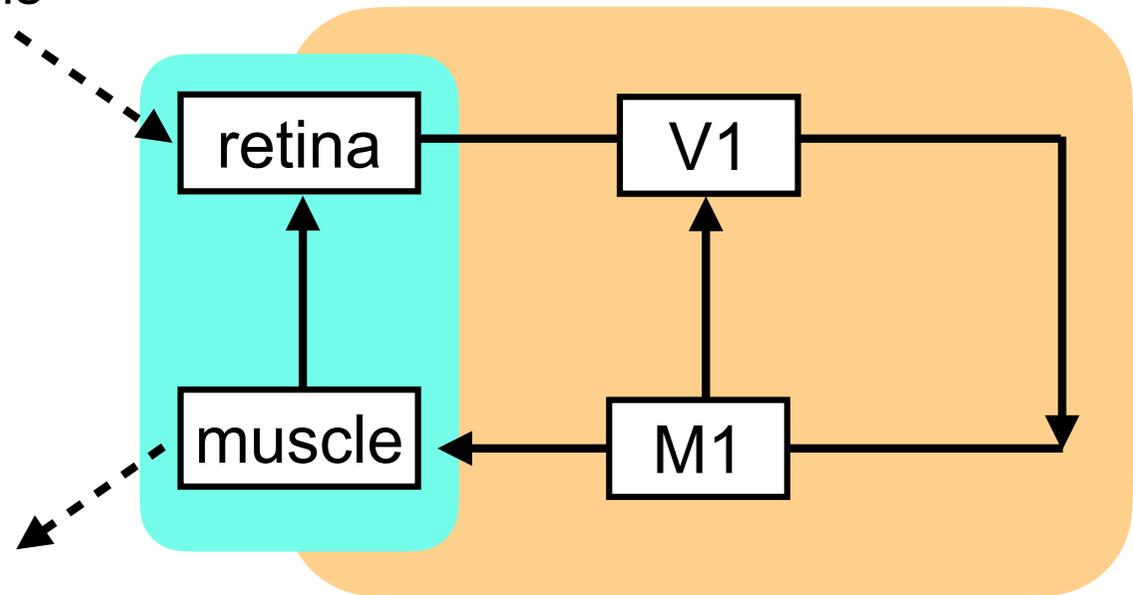


“Closed loop”

「入出力関係
によっては
規定されない」

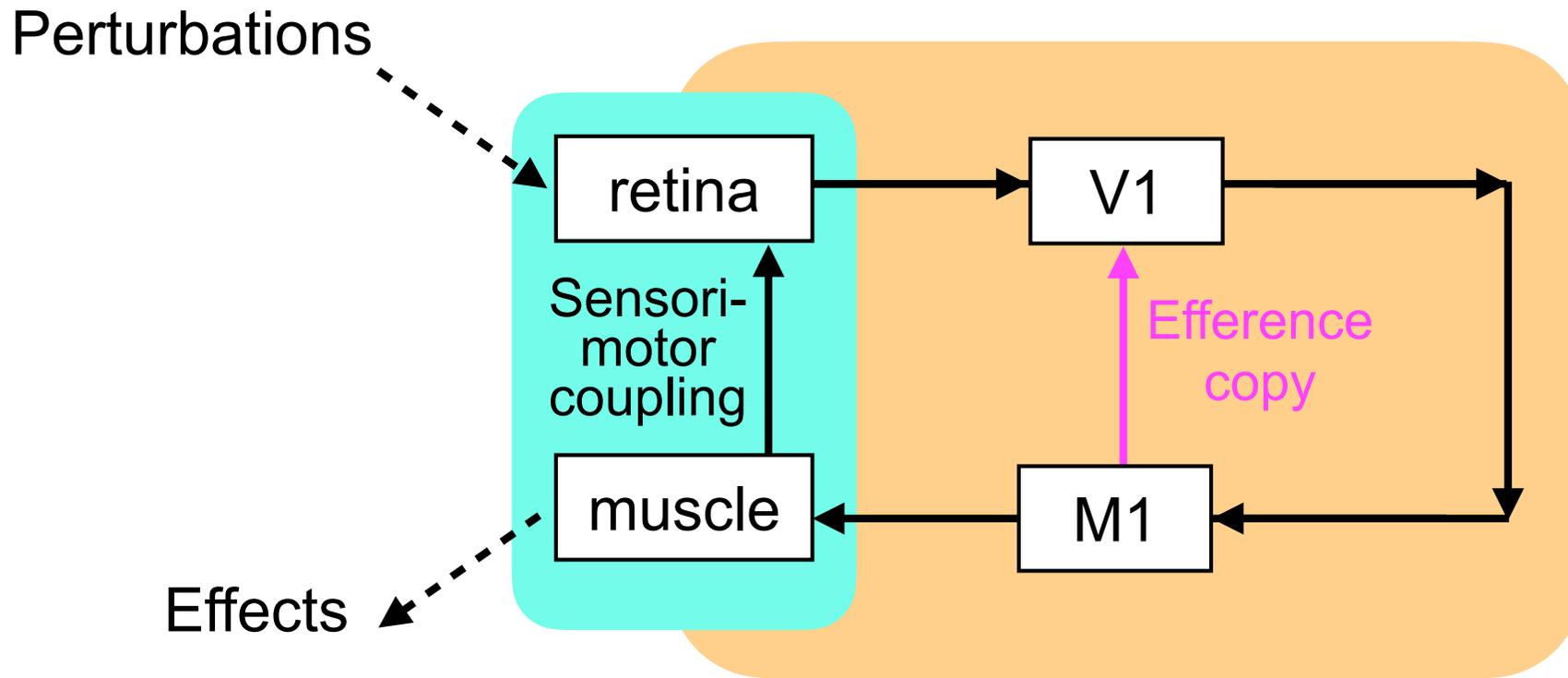
Perturbations

Effects



=>オートポイエーシス

オートポイエーシス = 操作的閉包 + 構造的カップリング



3) 構造的カップリング

外界とagentの内部状態は相互作用の履歴の結果、お互いの適合につながっている

1) 境界の自己決定

たとえば、マルコフブランケットによる実現

2) 操作的閉包

活動電位という「通貨」をやり取りしながら互いを維持している

エネルギーのやり取りはこのループとはカップルしていない

操作的閉包: 安静時脳活動

resting-state fMRI



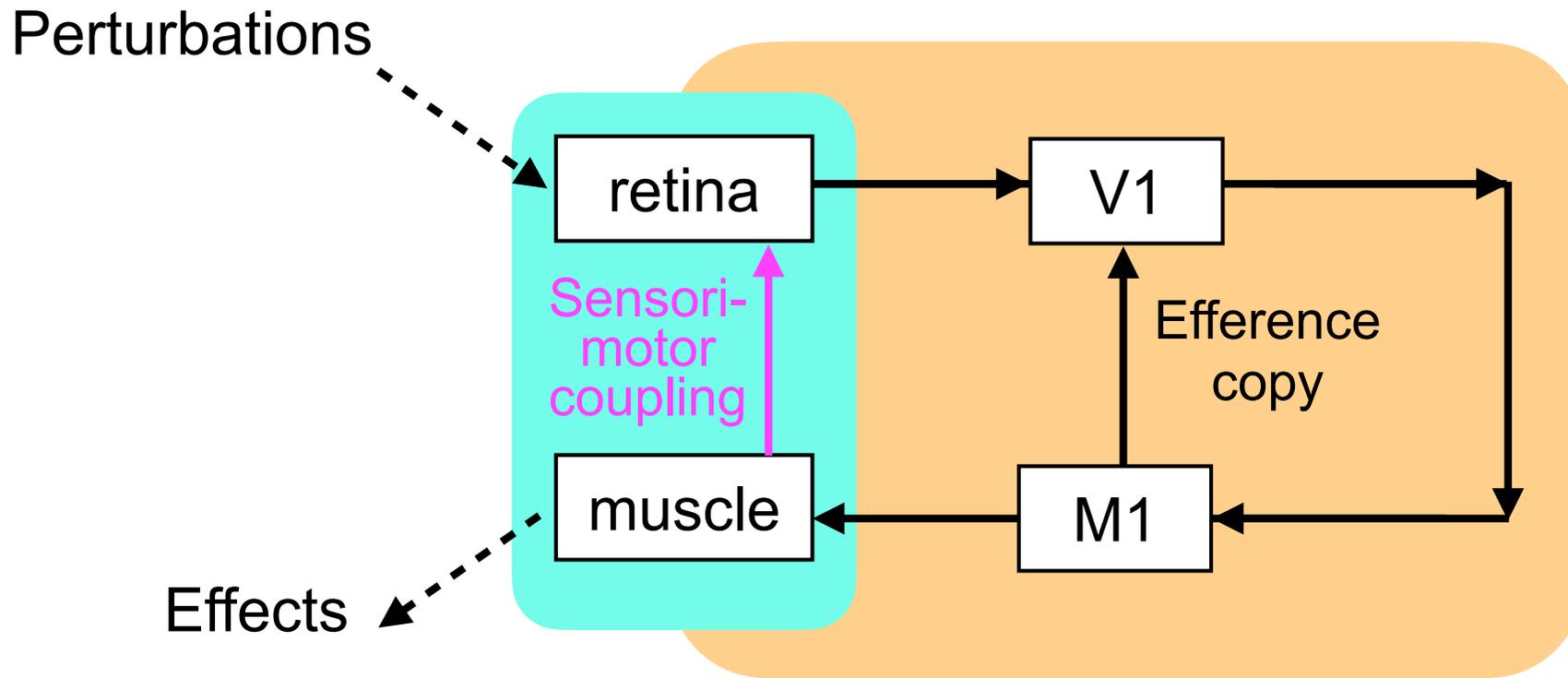
削除済み

Opportunities and limitations of intrinsic functional connectivity MRI.
Buckner RL, Krienen FM, Yeo BT.
Nat Neurosci. 2013 Jul;16(7):832-7.

脳は安静時でも活動している。

しかもその活動パターンは刺激が来たときと似かよっている。

オートポイエーシス = 操作的閉包 + 構造的カップリング



3) 構造的カップリング

外界とagentの内部状態は相互作用の履歴の結果、お互いの適合につながっている

1) 境界の自己決定

たとえば、マルコフブランケットによる実現

2) 操作的閉包

活動電位という「通貨」をやり取りしながら互いを維持している

エネルギーのやり取りはこのループとはカップルしていない

1) FEPの概念の深化

2) 自律性からエナクティヴィズムへ

- オートポイエーシス
- エナクティヴィズム

エナクティヴィズム Enactivism

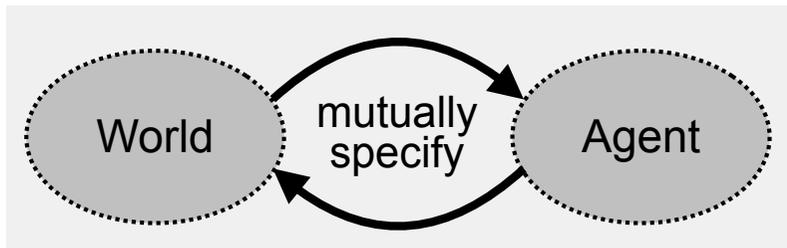
Varelaはそれまでに行ってきた
オートポイエーシスと自律性の考えを発展させて、
Enactivism、Enactiveアプローチ、を提案した。

Enact (役を演ずる) => 「行為からの産出」

Enactivismの定義 (p.246):

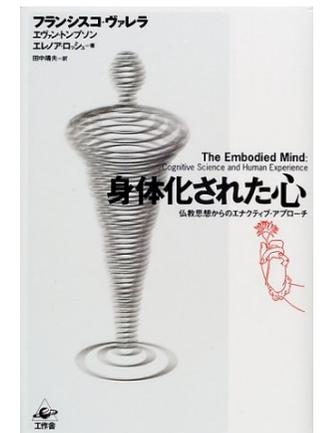
「知覚とは、知覚的に導かれた行為のことである」

Perception consists in perceptually guided action.

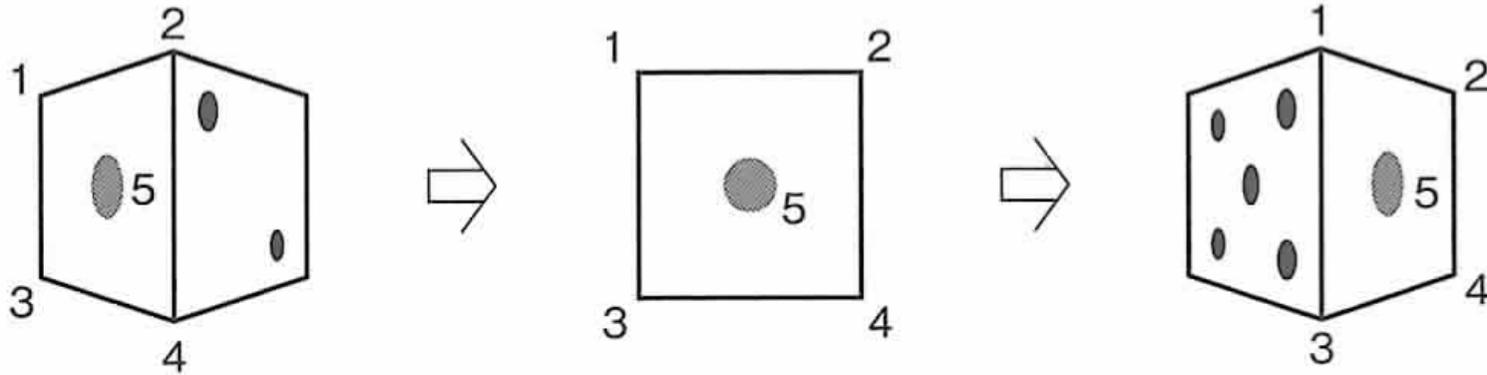


<=>オートポイエーシスは
操作的閉包による自己の産出、維持

フランシスコ・ヴァレラ、エヴァン・トンプソン、エレ
ノア・ロッシュ (1991) 身体化された心 仏教思想からの
エナクティブ・アプローチ. 工作舎 (日本語訳 2006)



フッサー現象学での視覚論



私が身体を左に動かせば、眼前のサイコロの
これまで見えていなかった左側の側面が見えてくる。
われわれが見ているものは、
つねに自分の身体の運動可能性
と結びつけられる仕方で経験されている。

Sensorimotor enactivism (SMC)

例えば、私たちが対象に向かって近づくと対象の姿が大きくなる。

私たちの知覚能力は、この種の感覚-運動随伴性の所有によって構成されている。

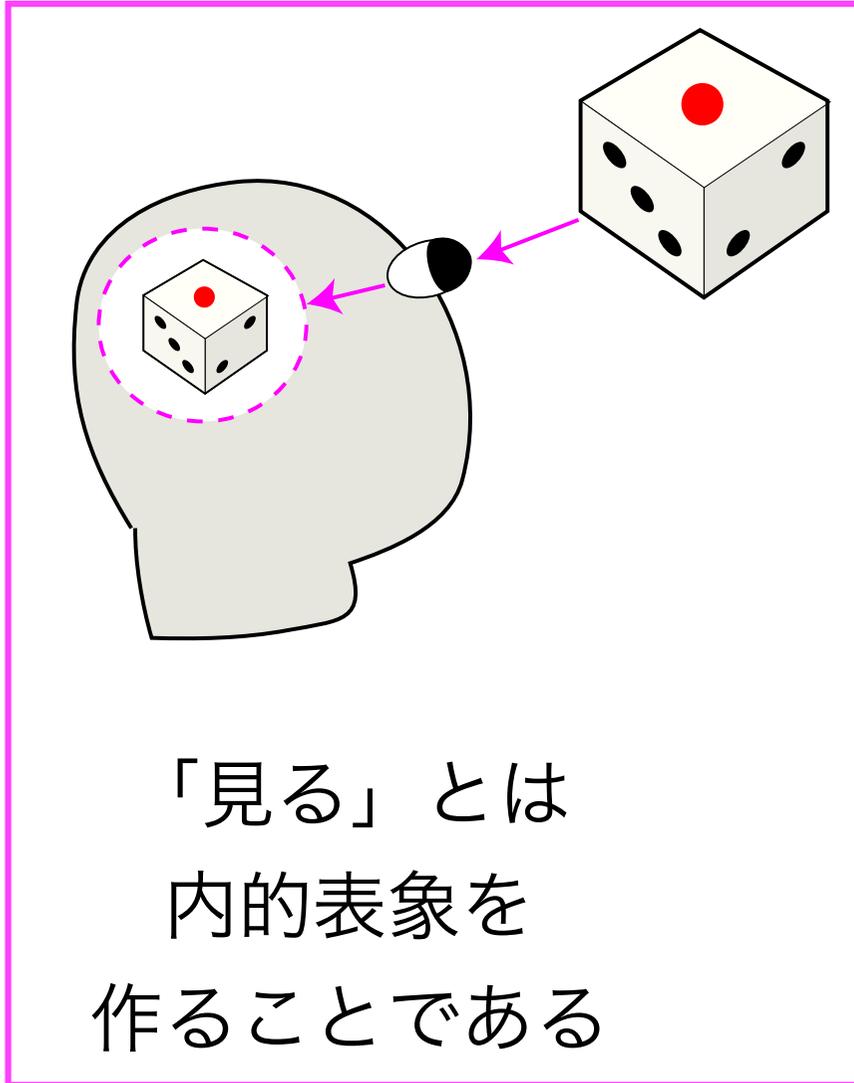
これによって、標準的な知覚説(表象説)では説明しにくい現象を説明できる

- 逆さメガネ
- 開眼手術

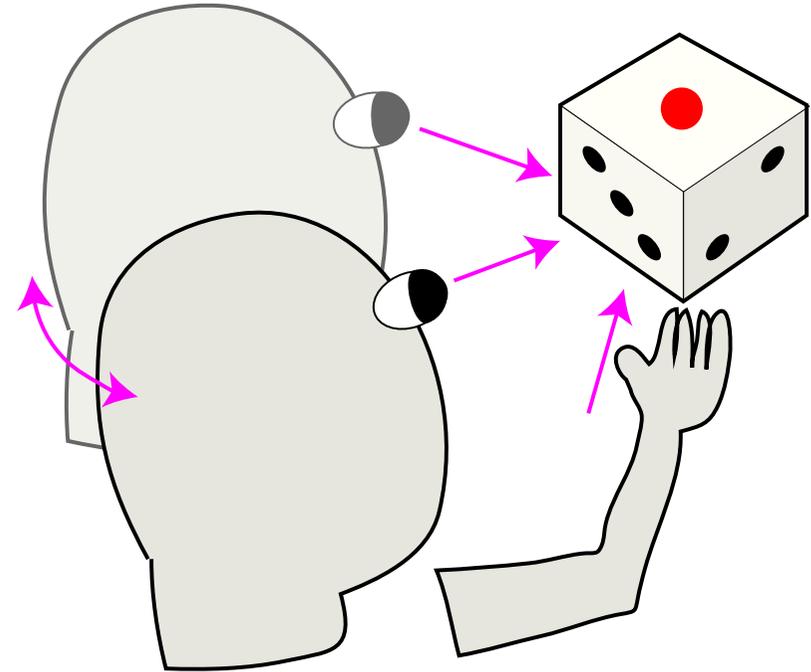


Sensorimotor enactivism (SMC)

表象説

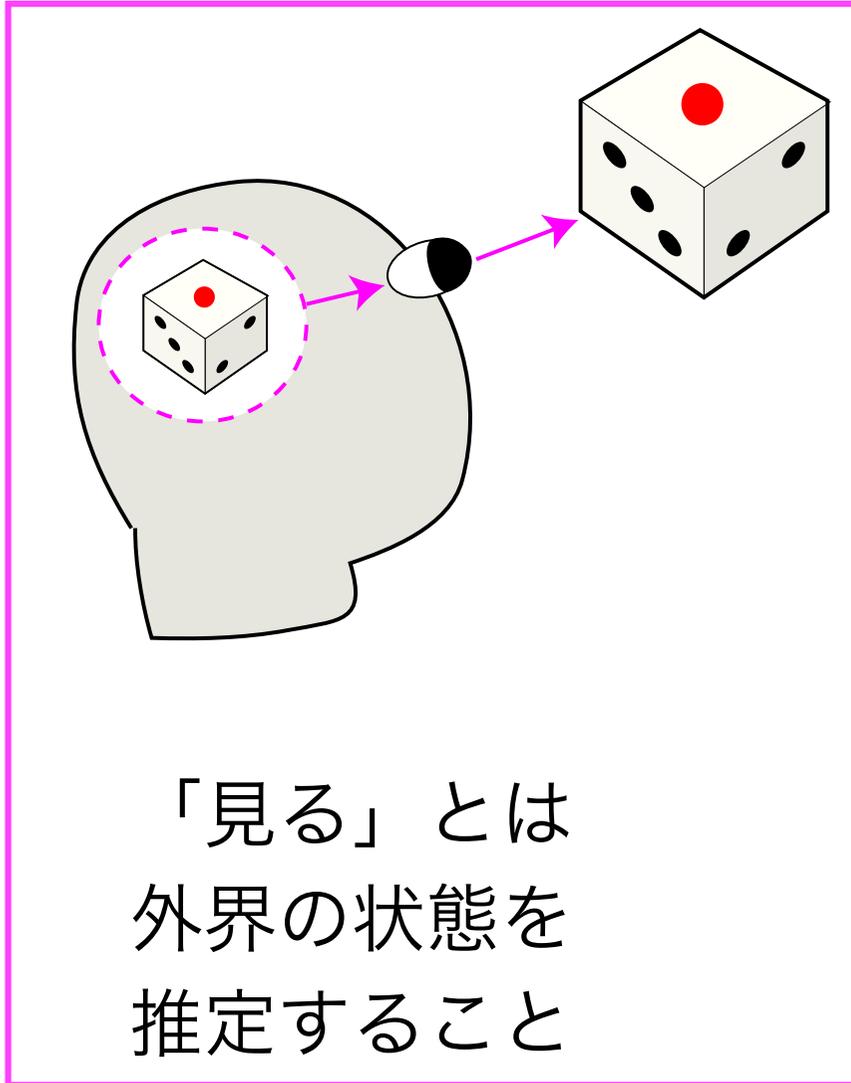


SMC 説

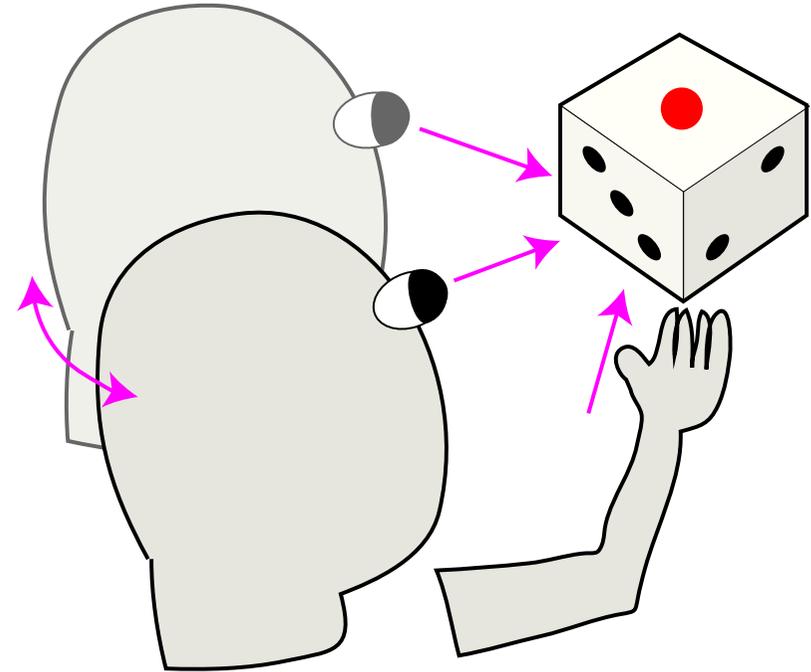


Sensorimotor enactivism (SMC)

無意識的推論



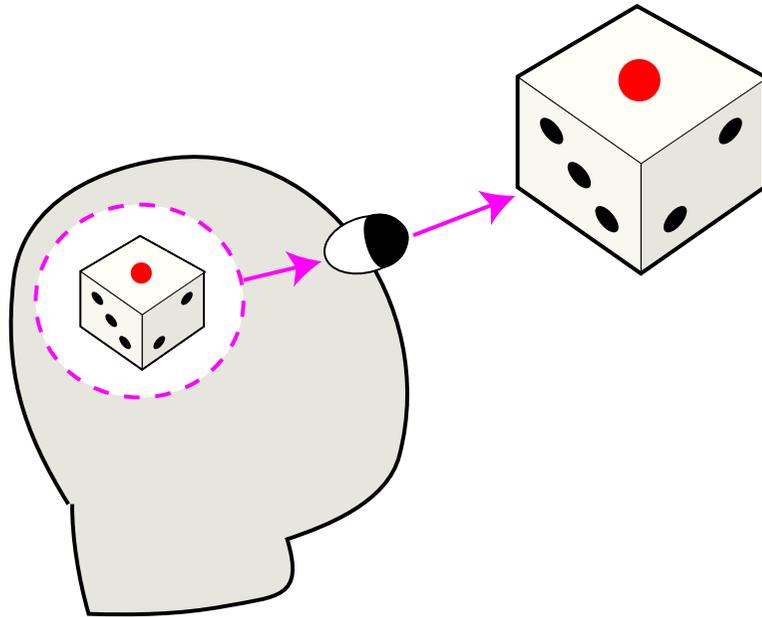
SMC 説



「見る」とは
感覚運動随伴性の習熟に媒介
された探索的活動である

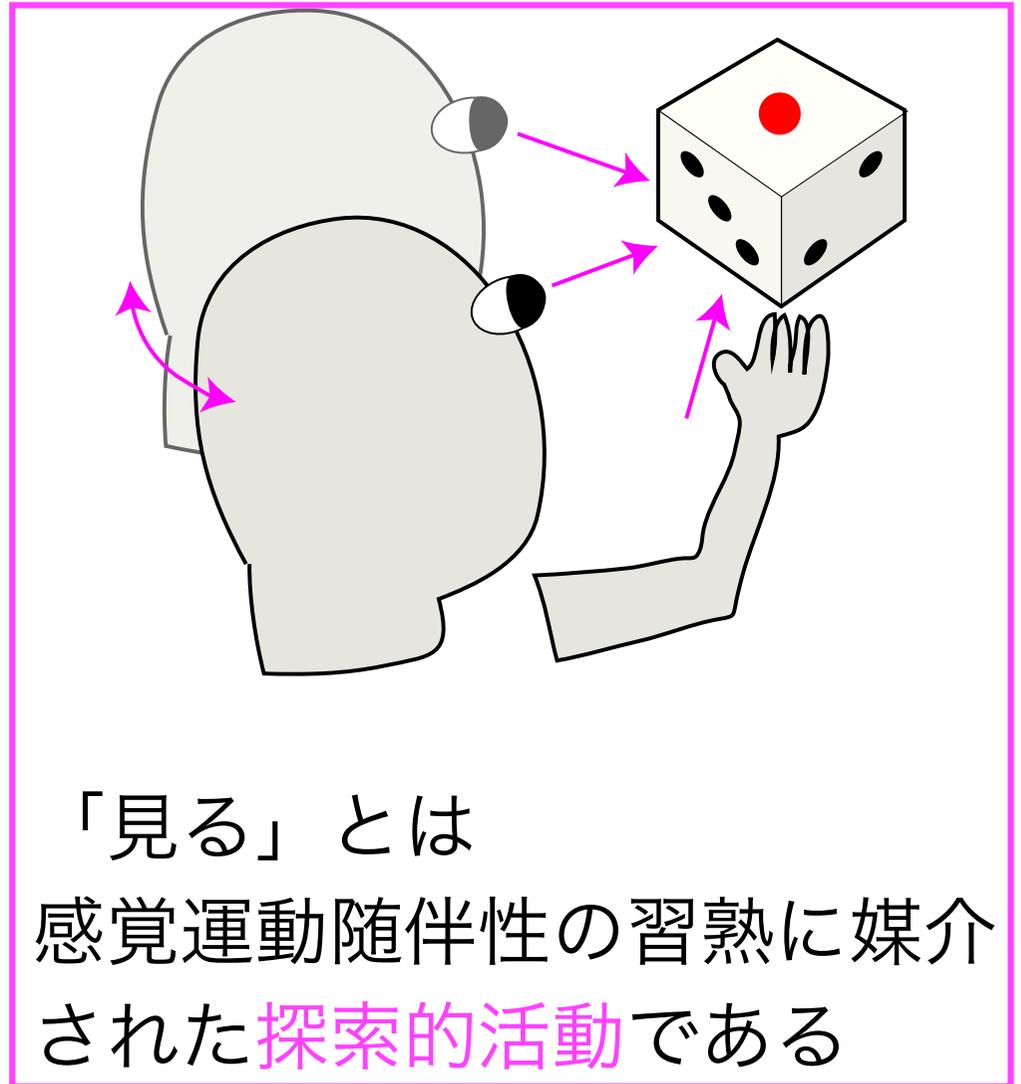
Sensorimotor enactivism (SMC)

無意識的推論



「見る」とは
外界の状態を
推定すること

SMC 説



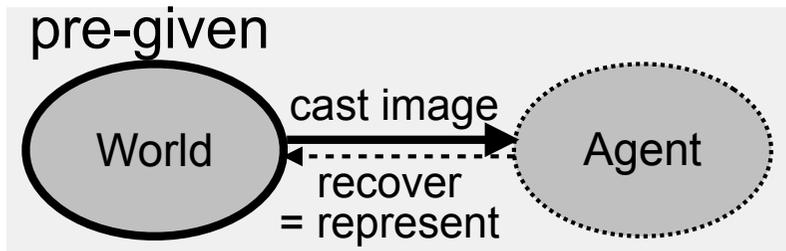
「見る」とは
感覚運動随伴性の習熟に媒介
された探索的活動である

エナクティヴィズムとは

追加

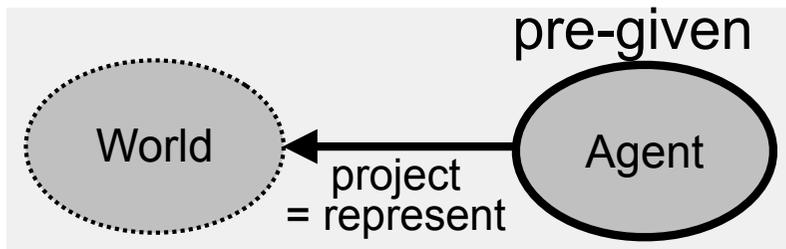
表象主義

古典的表象主義



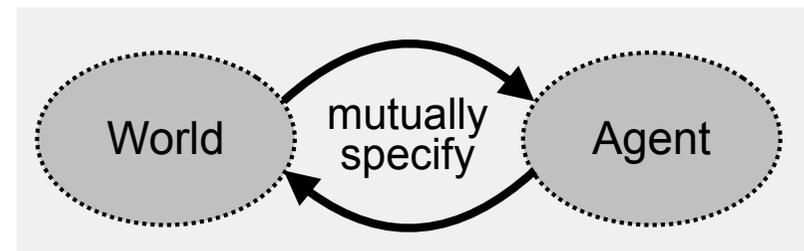
Cognition as the recovery of a pre-given outer world

ヘルムホルツの無意識的推論



Cognition as the projection of a pre-given inner world

エナクティヴィズム



Cognition as embodied action

Perception consists in perceptually guided action.

これは世界や知覚者の実在性 (形而上学)の話ではなくて、どのようにそれらが分節されるかの話と捉えるべき。

Varela F. J., Thompson E. & Rosch E. (1991) The embodied mind: Cognitive science and human experience. MIT Press, Cambridge. p.172-173

エナクティヴィズムとは

追加

4E cognition: Embodied, Embedded, Extended, and **Enactive**

身体性テーゼEmbodiment thesis:

身体は認知にとってなくてはならないものである

The **body** is **crucial** for **cognition**.

(2)

(3)

(1)

= a self-individuatingなシステム(継続的な自己の生成と維持)
= 自律的autonomousかつsense-makingするシステム

= sense-making
= (agent自身の生存に関わる面で) agentの状態と相互作用を適応的に調整regulateすること

= 身体なしにはsense-makingできない
+ sense-makingとは身体的な過程である
= 身体と認知の関係は因果的かつ構成的constitutiveである

Enactive view

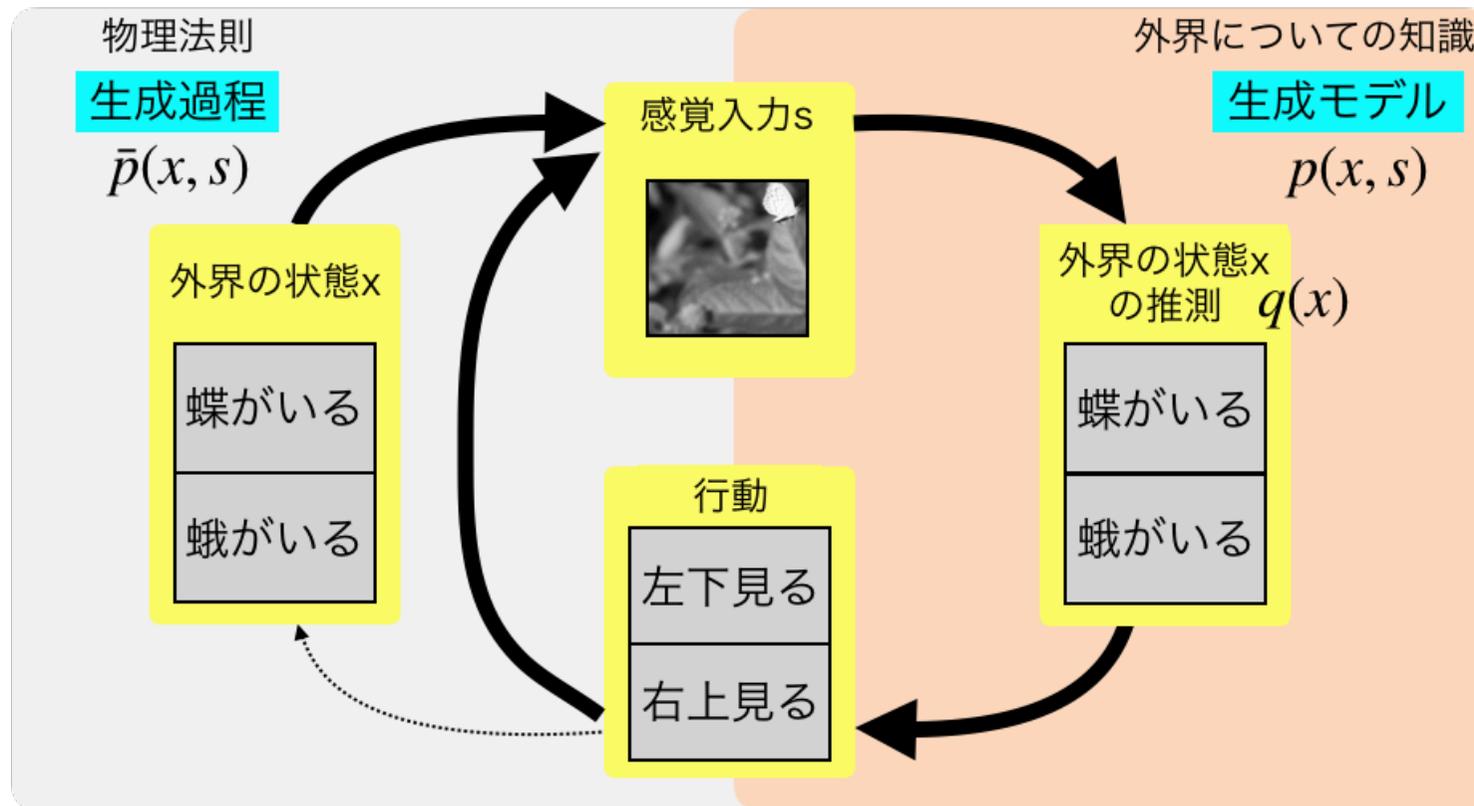
Di Paolo, E., & Thompson, E. (2014). The enactive approach. In L. Shapiro (Ed.), Routledge handbooks in philosophy. The Routledge handbook of embodied cognition (p. 68–78). Routledge/Taylor & Francis Group.

1) FEPの概念の深化

2) 自律性からエナクティヴィズムへ

- オートポイエーシス
- エナクティヴィズム
- **Back to FEP**

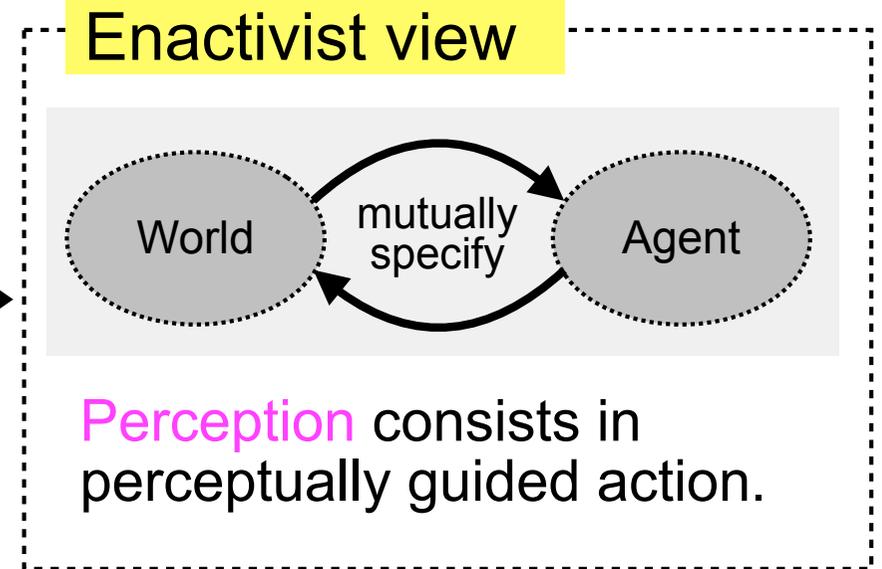
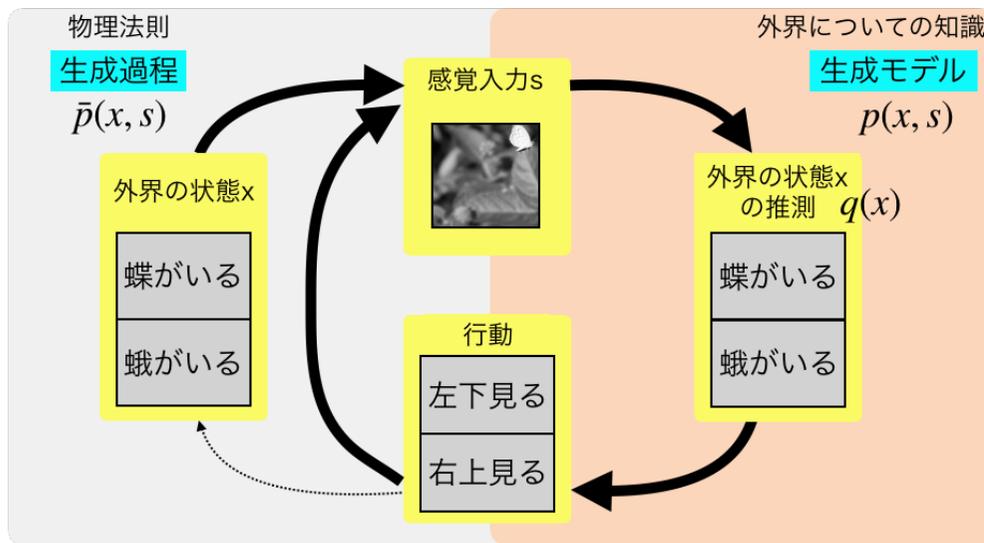
FEPはenactivismと整合的



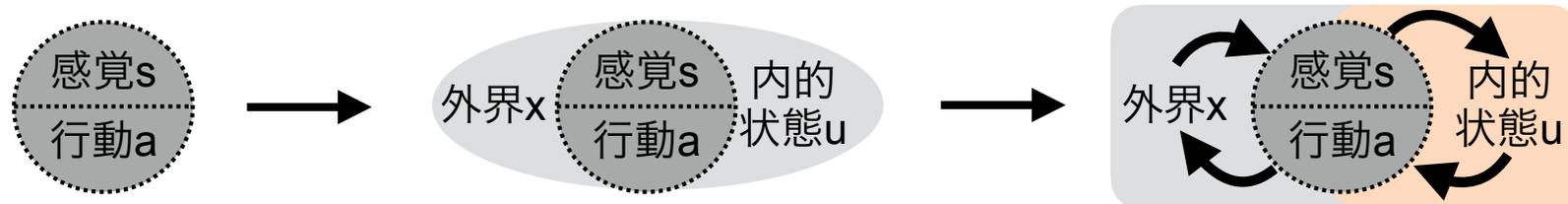
- FEPにおける
 - 生成モデル $p(x, s)$: 世界についての知識 (観察モデル*事前分布)
 - 近似的推測 $q(x)$: イマココでの推測
- SMC説における感覚運動随伴性への習熟とは生成モデルの確立、維持と表現できる。

FEPをもっとenactiveにするには？

Agentと環境の境界は行為によってそのつど決まるべき



=> 感覚運動ループが境界を作り、
学習と発達と進化が独立性を高める



FEPをもっとautopoieticにするには？

Ver.4 FEP: 「マルコフ・ブランケットを持っているNESSでは、
内部状態があたかも外部状態の推測をするようにふるまう」

「内」と「外」との
情報の独立性

$\frac{dp^*(\mu)}{dt} = 0$ 非平衡定常状態

マルコフ
ブランケット

表現系が
安定して存在する

サプライズ最小化
による更新則

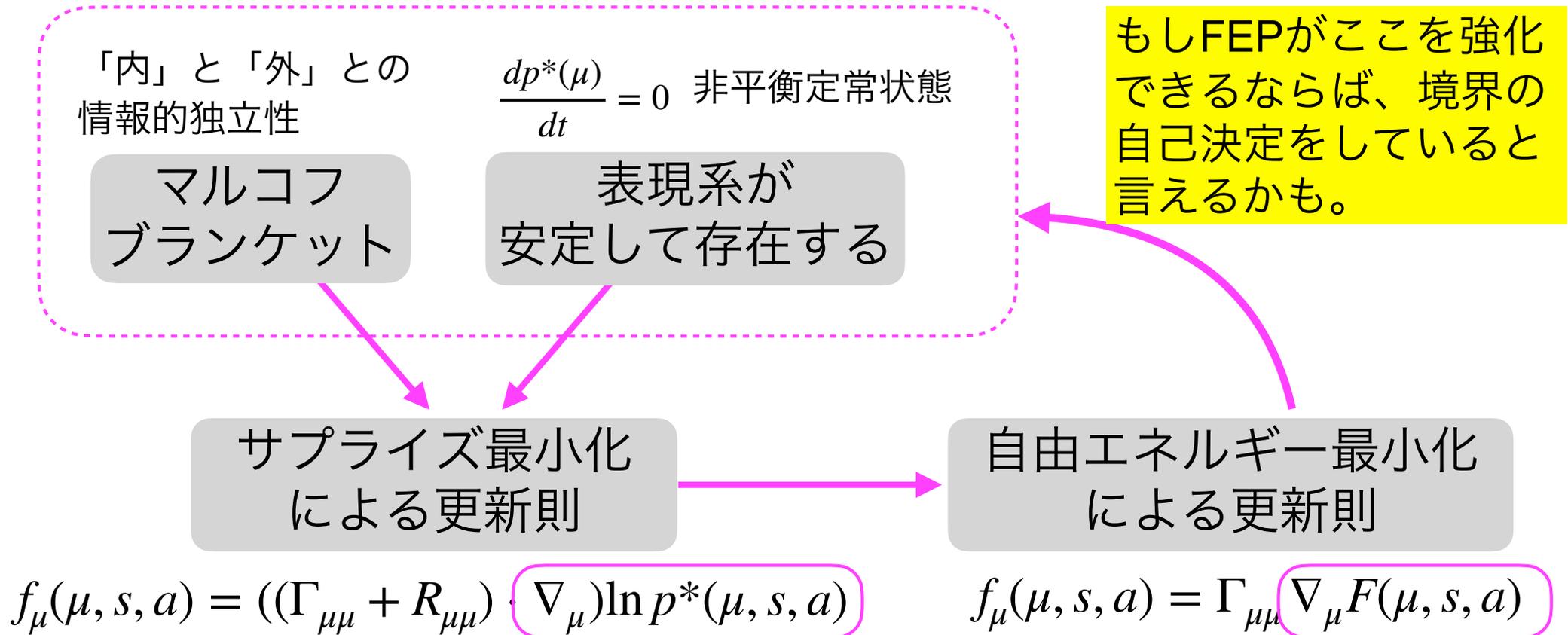
自由エネルギー最小化
による更新則

$$f_{\mu}(\mu, s, a) = ((\Gamma_{\mu\mu} + R_{\mu\mu}) \cdot \nabla_{\mu}) \ln p^*(\mu, s, a)$$

$$f_{\mu}(\mu, s, a) = \Gamma_{\mu\mu} \nabla_{\mu} F(\mu, s, a)$$

FEPをもっとautopoieticにするには？

Ver.4 FEP: 「マルコフ・ブランケットを持っているNESSでは、内部状態があたかも外部状態の推測をするようにふるまう」



FEPをもっとenactiveにするには？

BIC的なベイズ脳仮説およびFEP

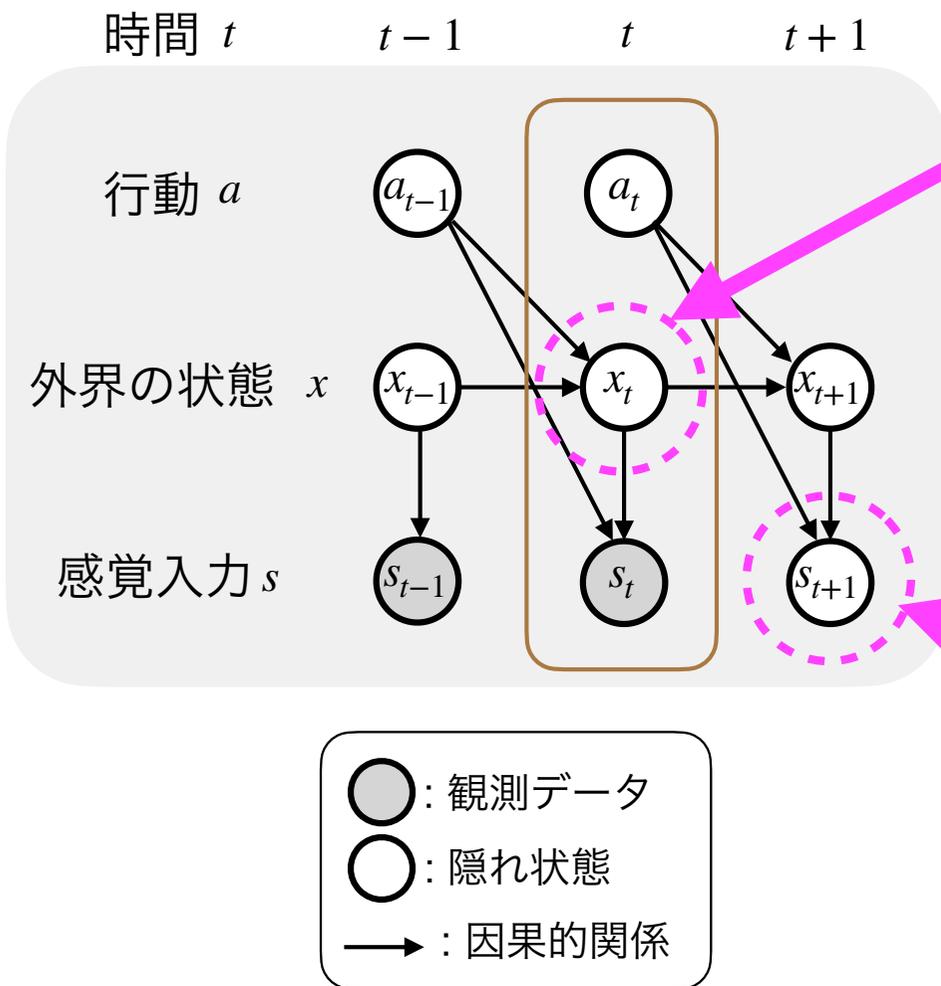
古典的なベイズ主義(xの实在論):
 隠れ値xについての(主観的)信念
 事後分布 $p(x_t | \{s_1, s_2, \dots, s_t\})$

AIC的なモデル選択理論

予測と汎化 (xの道具主義):
 新しい観察データsの予測

事後予測分布 $p(s_{t+1} | \{s_1, s_2, \dots, s_t\})$

エリオット・ソーバー (2012) 「科学と証拠 —統計の哲学 入門」 名古屋大学出版会. p.151-156 を元に



(BIC的な)現状のFEPは、xの实在論に基づいており、enactiveでない。
 AIC的なFEPがもしあるならば、それはenactivismと整合的。

END