

よくわかる  
フリストンの  
自由エネルギー原理

生理学研究所  
認知行動発達研究部門  
助教 吉田 正俊

# 本スライドの目的

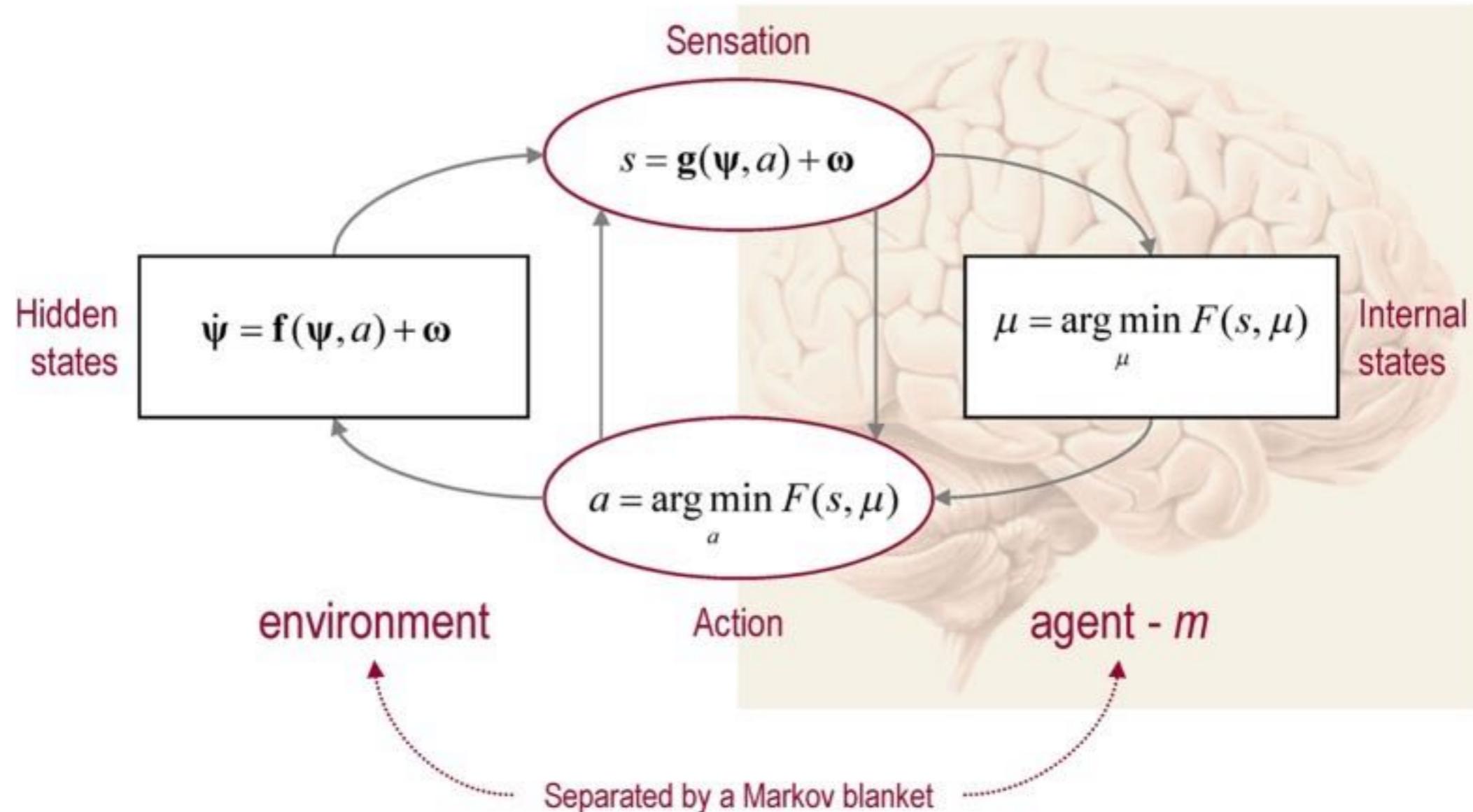
脳における知覚と行動を統一的に説明する  
「自由エネルギー原理」について  
「視覚シーンを眼で探索する」状況を  
題材にしてわかりやすく解説したい。

# 全体的な流れ

1. 自由エネルギー原理とは
2. レベル別に説明

# 自由エネルギー—最小化原理

Exchange with the environment



Friston K, Breakspear M, Deco G. Perception and self-organized instability. Front Comput Neurosci. 2012 Jul 6;6:44. Fig.1 (CC BY 3.0) [http://www.frontiersin.org/files/Articles/23035/fncom-06-00044-r4/image\\_m/fncom-06-00044-g001.jpg](http://www.frontiersin.org/files/Articles/23035/fncom-06-00044-r4/image_m/fncom-06-00044-g001.jpg)

# 自由エネルギー原理 (FEP)

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

「意識についての理論」 そのものではない

# 自由エネルギー「原理」である意義

注意

シナプス可塑性

学習・記憶

神経のコード

Friston (2010) Fig.4の図  
SlideShare用資料では省略。

運動制御

価値判断

ミラーニューロン

自閉症

統合失調症

進化

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

FEPは脳についての理論を統一的に理解するための「原理」であり、これまでの理論を置き換えるためのものではない。

# 全体的な流れ

1. 自由エネルギー原理とは

2. レベル別に説明

レベル1: 認識による世界のモデル化 (perceptual inference)

レベル2: + 行動による世界の探索 (active inference)

レベル3: + 反実仮想による世界のモデル化と目的の発生

# 全体的な流れ

1. 自由エネルギー原理とは

2. レベル別に説明

レベル1: 認識による世界のモデル化 (perceptual inference)

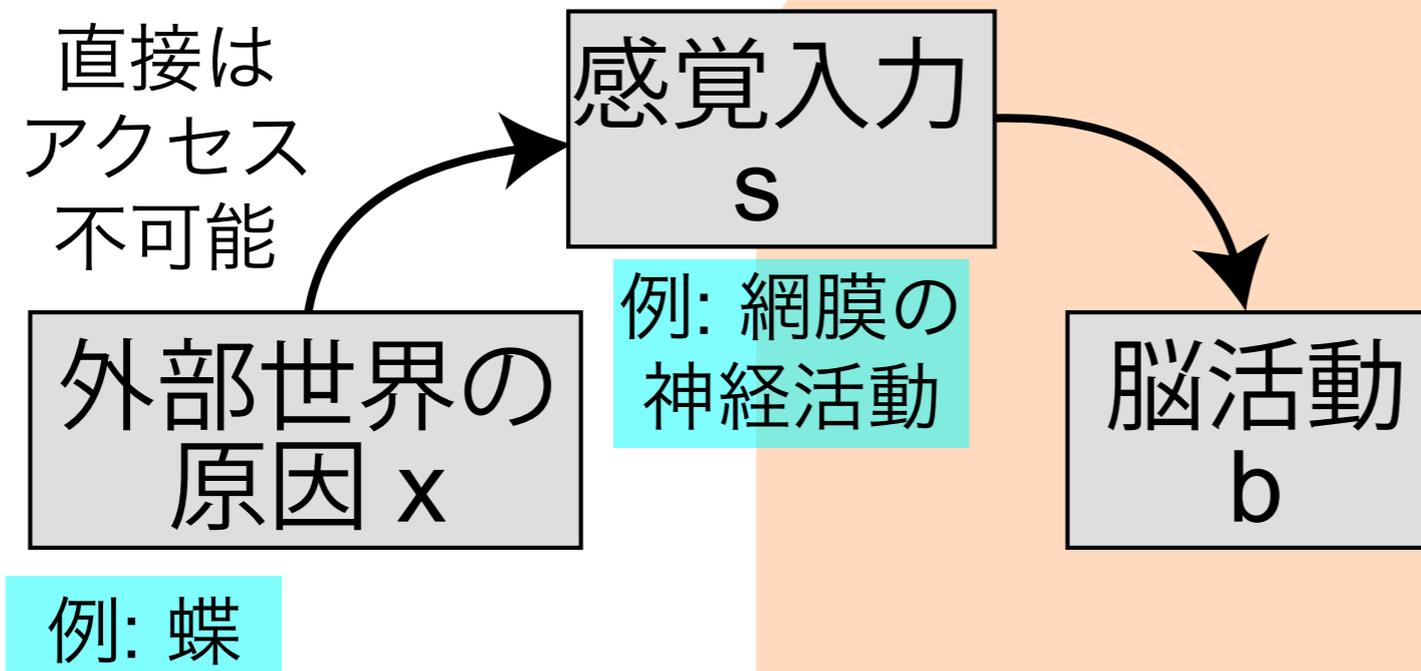
レベル2: + 行動による世界の探索 (active inference)

レベル3: + 反実仮想による世界のモデル化と目的の発生

# 自由エネルギー原理 (FEP)

外部世界

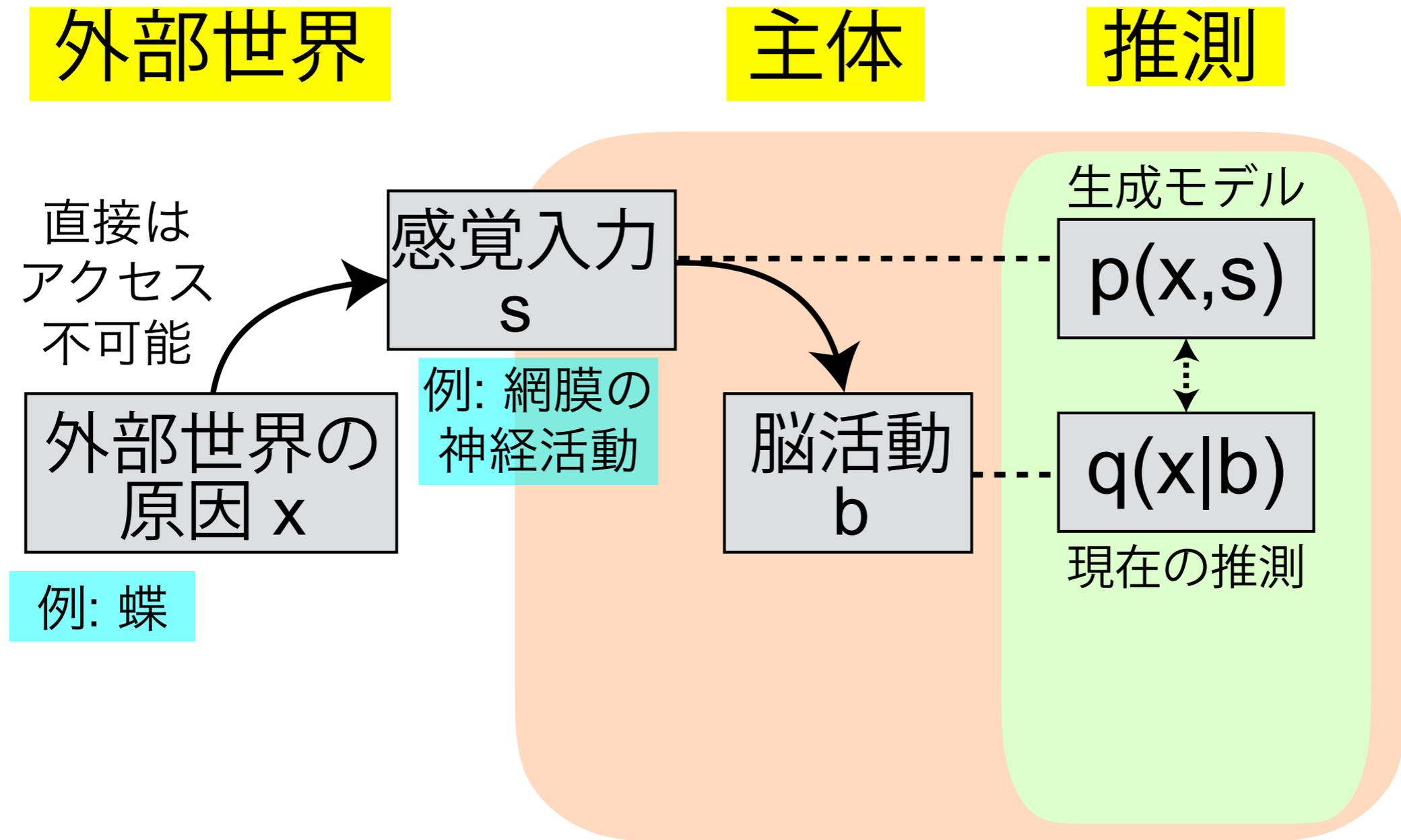
主体



**ヘルムホルツ的視覚観**：Agentは直接アクセス不可能な世界 $x$ を知るために感覚入力 $s$ から「無意識的推論」を行う。

**変分ベイズ推定**における観測値 $s$ から隠れ値 $x$ の推定と形式的に同じ。

# 自由エネルギー原理 (FEP)

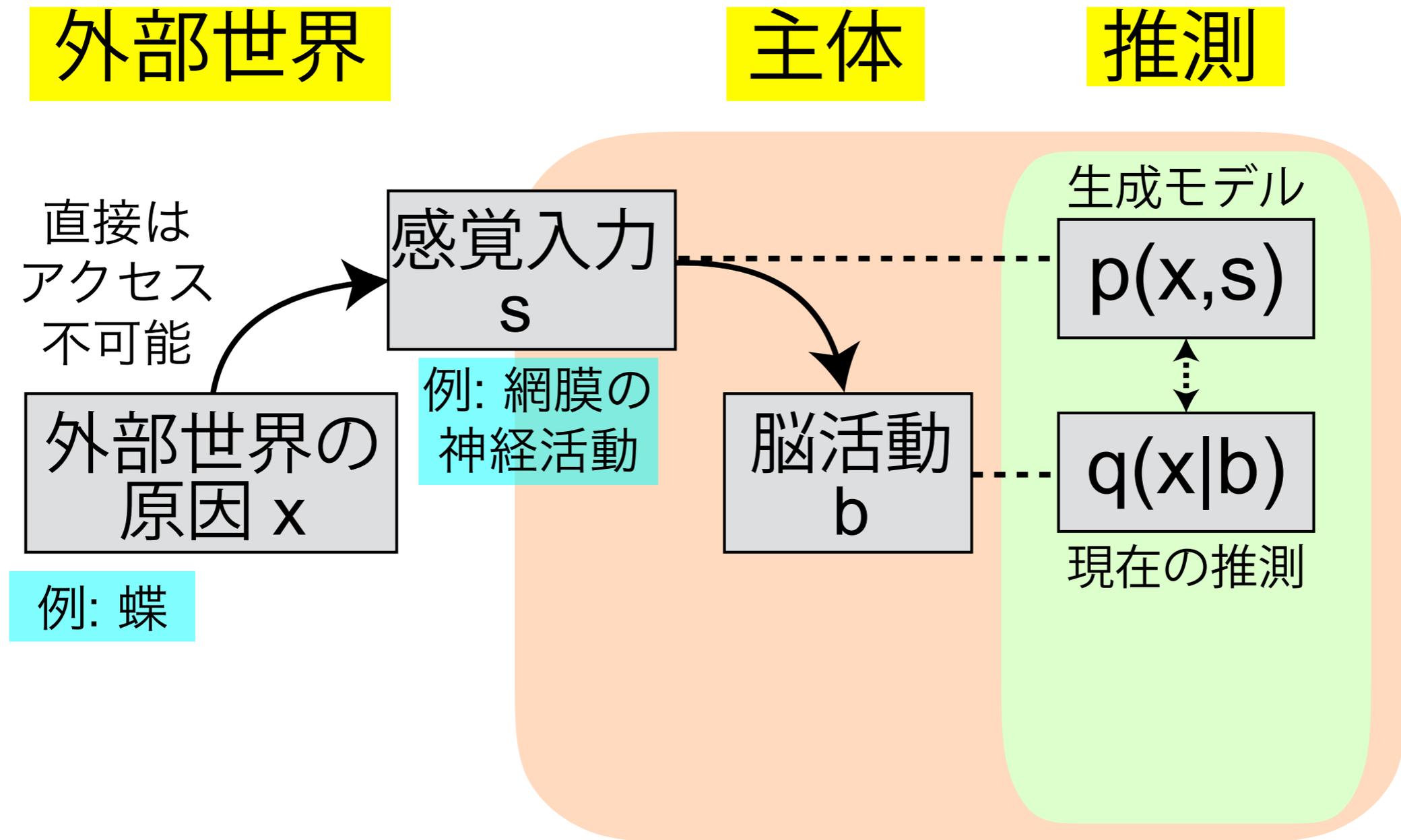


情報理論的自由エネルギーIFE:

生成モデル: Agentは感覚入力 $s$ とその原因 $x$ の関係を知識として獲得

現在の推測: イマココでの原因 $x$ の推測(脳活動からの事後分布)

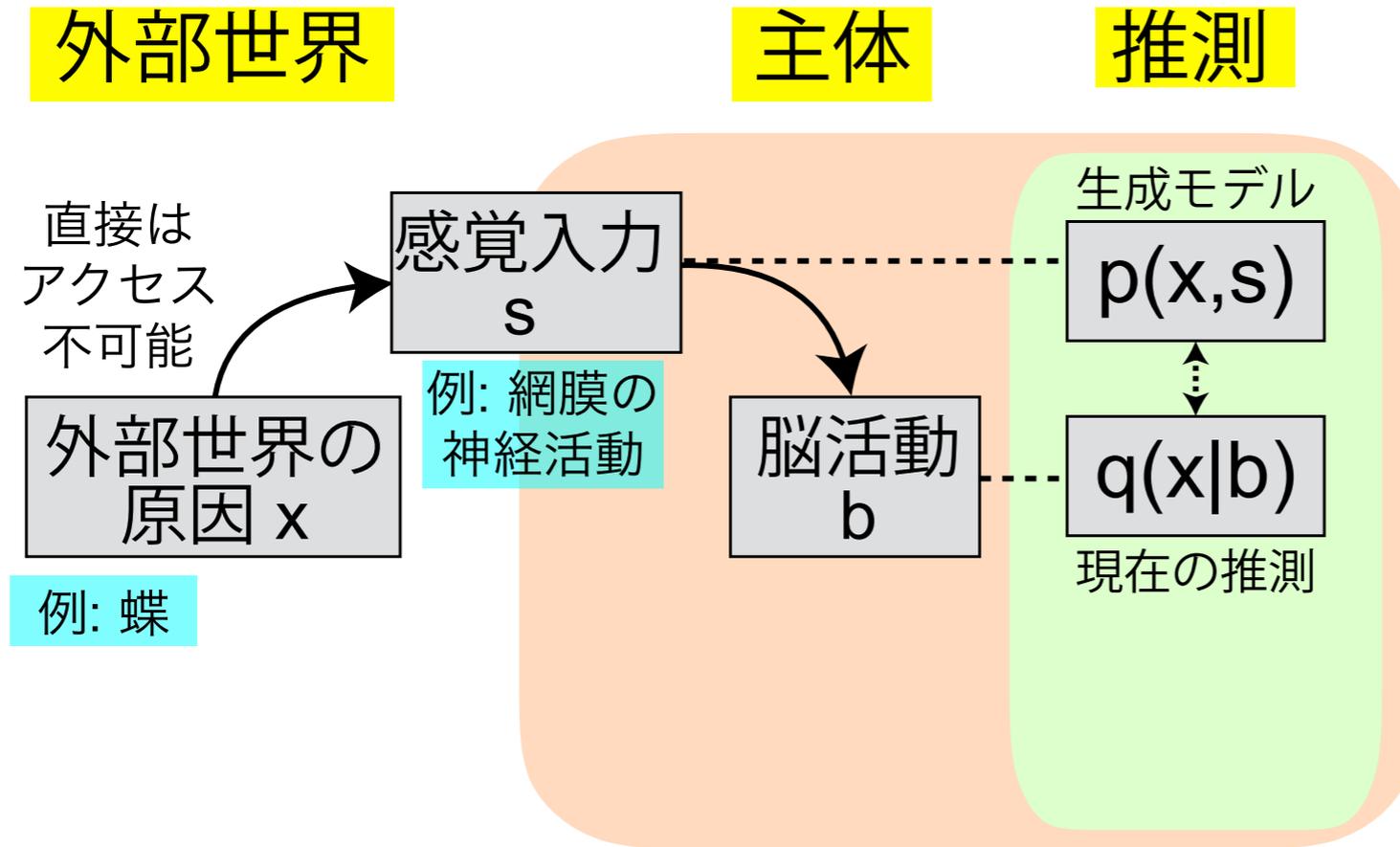
# 自由エネルギー原理 (FEP)



情報理論的自由エネルギー:  $F(s, b) = D_{KL}(q(x|b) || p(x, s))$

「現在の推測 $q$ 」から「生成モデル $p$ 」までの距離 = 予測誤差

# 自由エネルギー原理 (FEP)



$$F(s, b) = D_{KL}(q(x|b) \| p(x, s))$$

$$F(b) = D_{KL}(q(x|b) \| p(x|s)) + C_1(s)$$

Fを最小化するには  
どうすればいい？

(1) bを変える (sは固定) - Perceptual inference = 認識

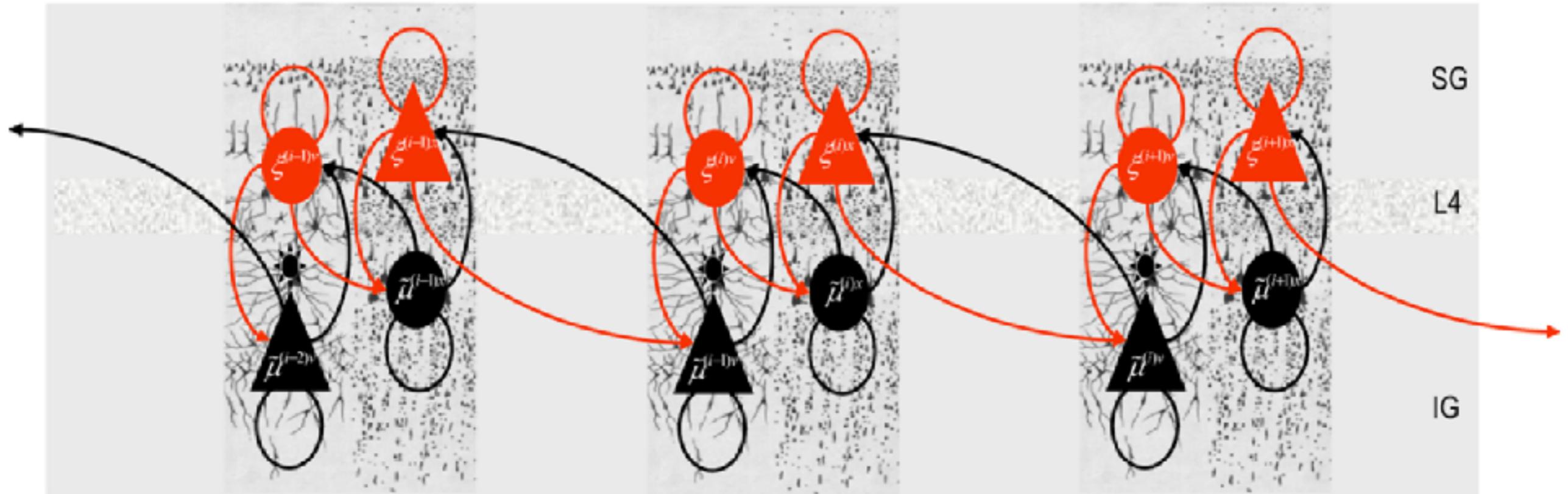
脳活動bを変えると「推測した原因q」が変わる。

「推測した原因q」が「真の原因p」と一致するとき(=正しい認識)、  
DKL=0となり、Fは最小になる。

# Predictive coding

予測誤差 = 視覚サリエンス

Forward prediction error



Backward predictions

予測モデル = Conscious perception

このような考え方は「Predictive coding」としてモデル化されてきた。つまり、レベル1の説明だけでは「自由エネルギー原理」を持ち出す必要はとくに無い。レベル2を考える必要がある。

# 全体的な流れ

1. 自由エネルギー原理とは

2. レベル別に説明

レベル1: 認識による世界のモデル化 (perceptual inference)

レベル2: + 行動による世界の探索 (active inference)

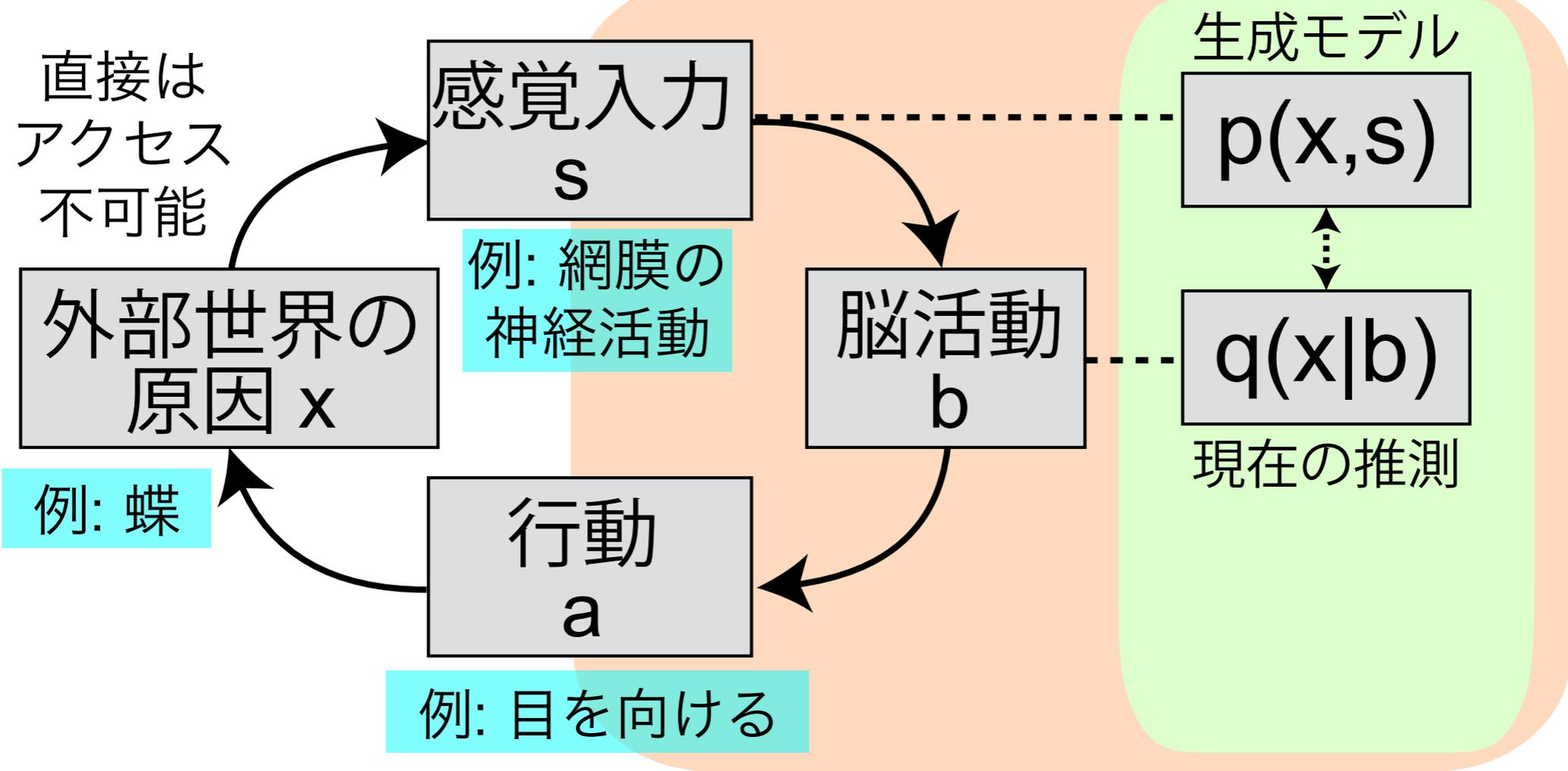
レベル3: + 反実仮想による世界のモデル化と目的の発生

# 自由エネルギー原理 (FEP)

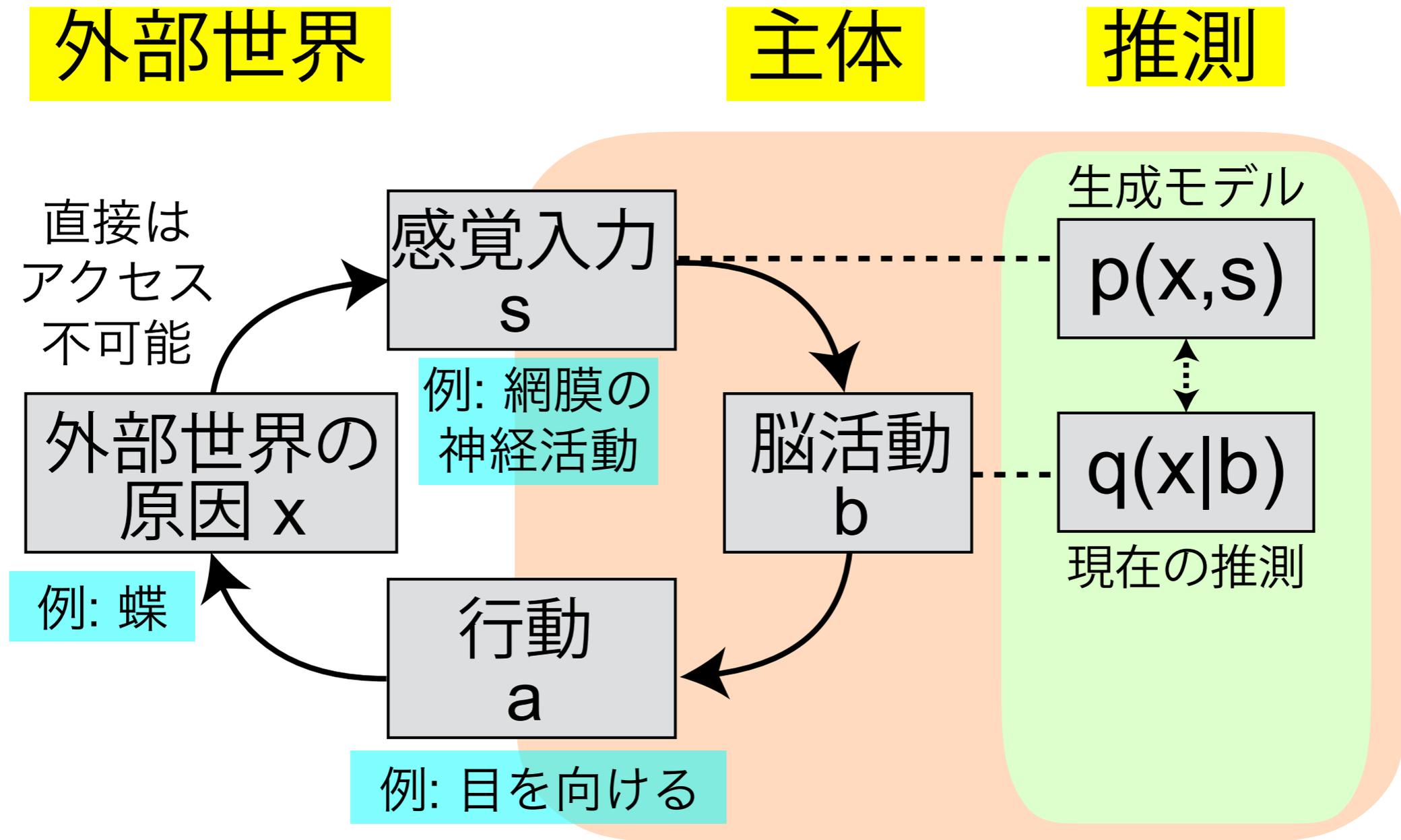
外部世界

主体

推測



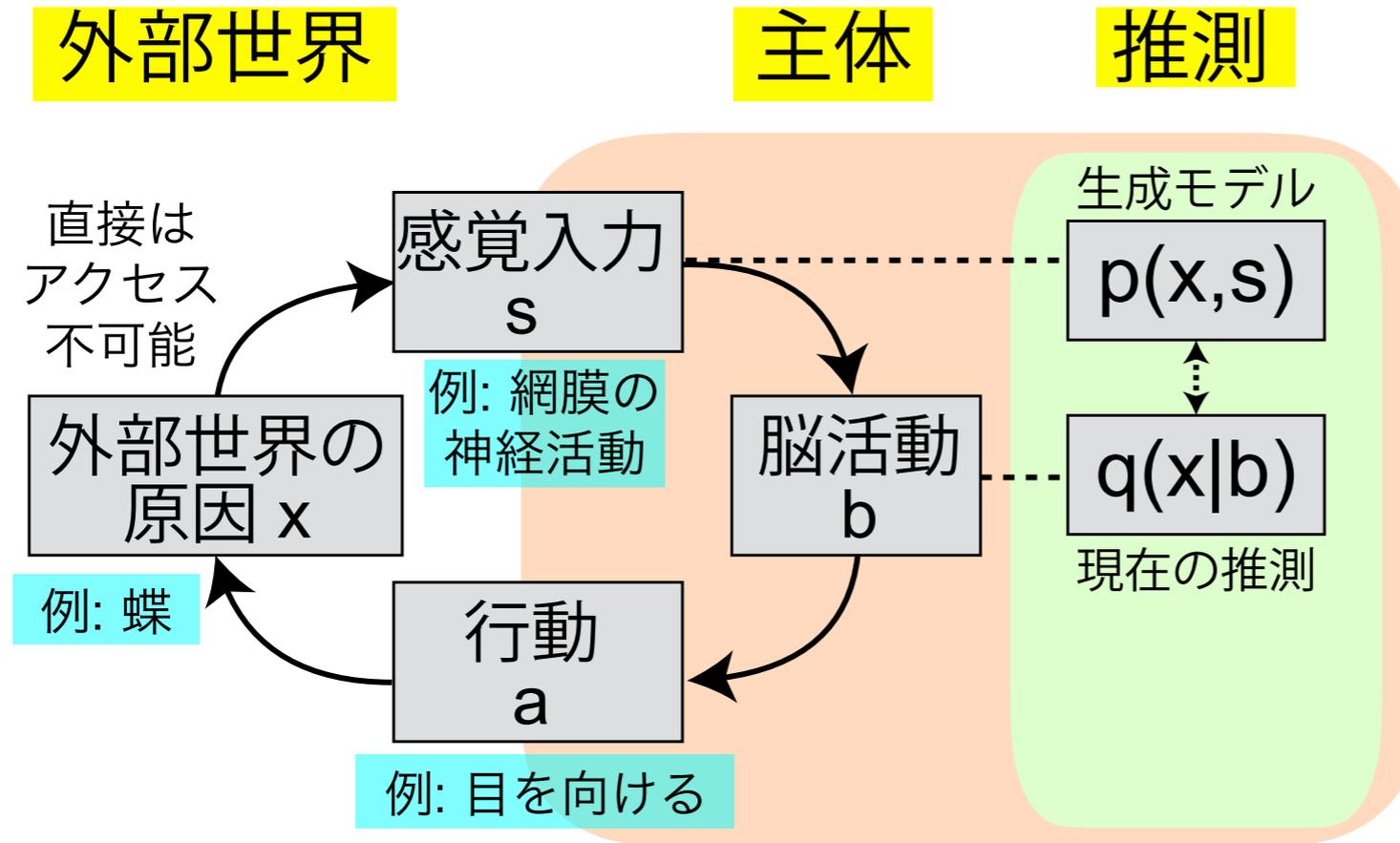
# 自由エネルギー原理 (FEP)



$$F(s, b) = D_{KL}(q(x|b) || p(x, s)) \quad \text{Fを最小化するには?}$$

- (1)  $b$ を変える ( $s$ は固定) - 現在の推測のアップデート = 認識
- (2)  $s$ を変える ( $b$ は固定) - 行動 $a$ で感覚 $s$ を変える = 行動

# 自由エネルギー原理 (FEP)



accuracy

$$F(s) = C_2(b) - \sum_x \{q(x|b) * \log(p(s|x))\}$$

Fを最小化するにはどうすればいい？

(2)  $s$ を変える ( $b$ は固定) - Active inference = 行動

行動 $a$ によって感覚入力 $s$ を変える(例：対象に眼を向ける)。これによって「推測した原因 $q$ 」をより選択的に生み出す感覚入力 $s$ をサンプルしたとき、accuracyが最大化し、 $F$ が最小になる。

# 自由エネルギー原理 (FEP)

「行動する理由は  
周りの世界を知りたいから」

学者の欲望に忠実な世界観だな！

# FEPを実例を追って理解する



われわれは普段こういう視覚像を経験しているつもりであるが、じっさいには中心窩の1度くらいの範囲しか解像度が高くないので、これは正しくない。

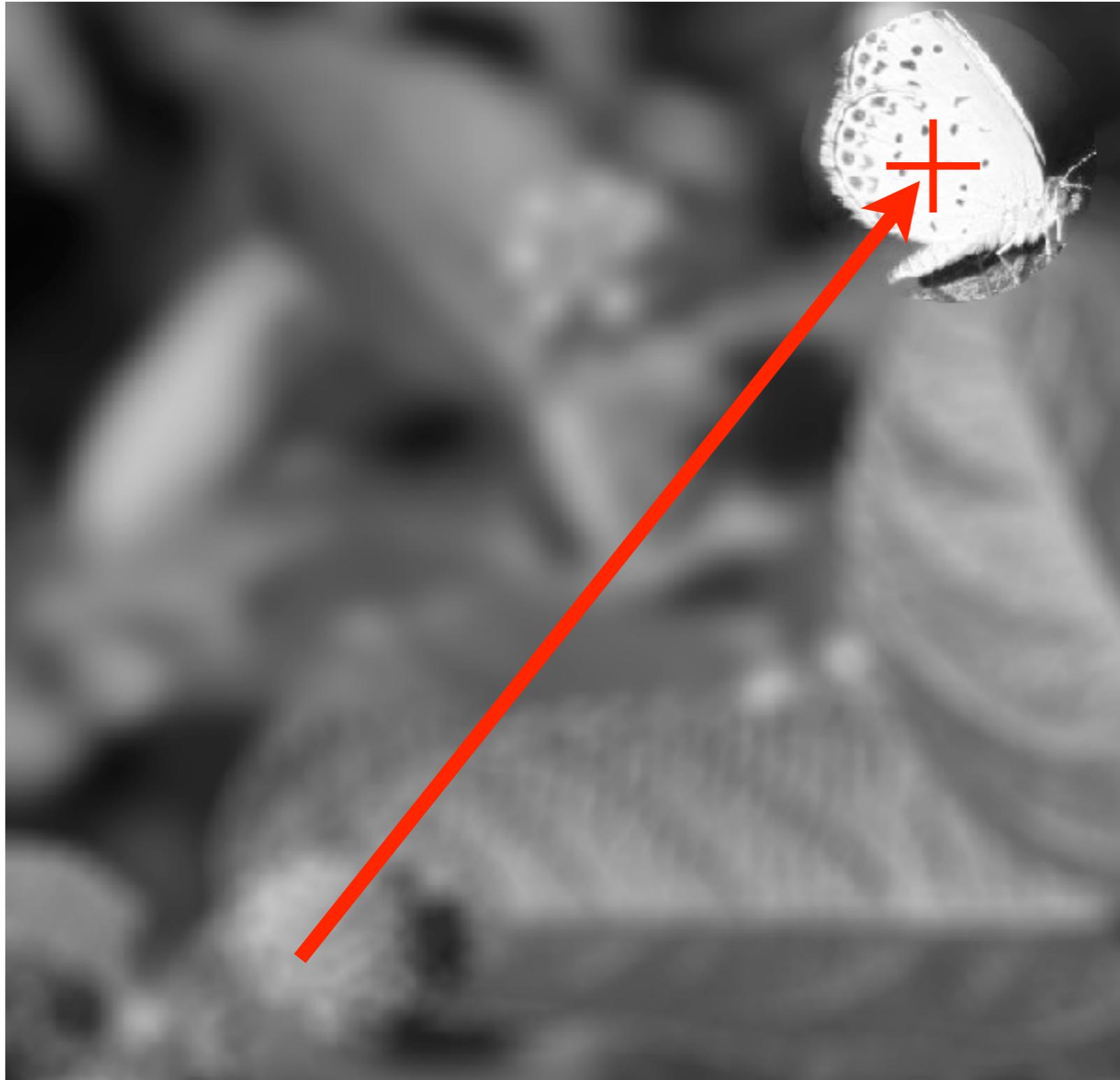
# FEPを実例を追って理解する



実際の私達に与えられている感覚入力はこんなかんじ。

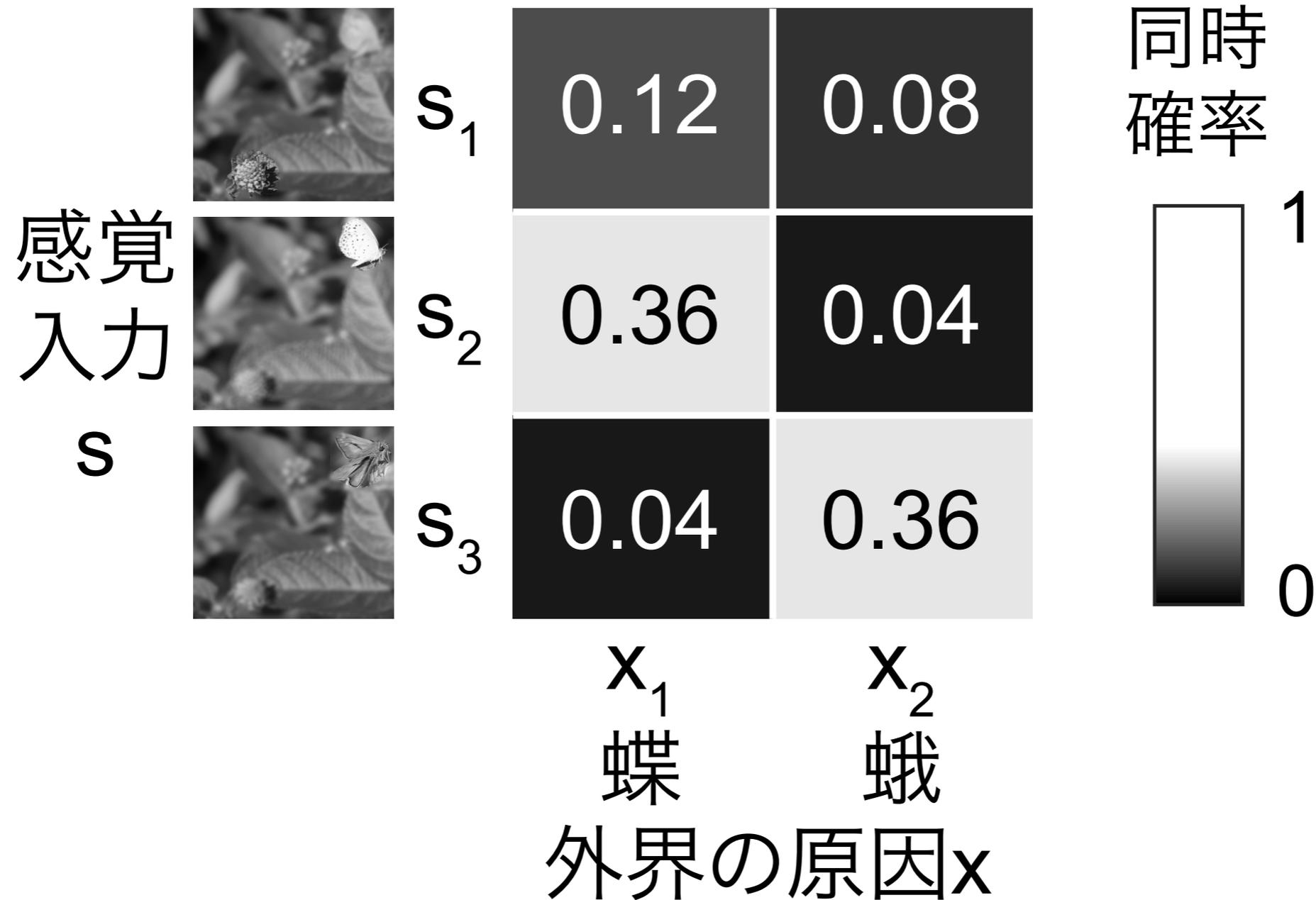
+に視線があるとき、右上にあるものはぼんやりしていてなんだかわからない。

# FEPを実例を追って理解する



視線を右上に向けると、これが蝶であることが判明した。

# 生成モデル $g(x,s)$

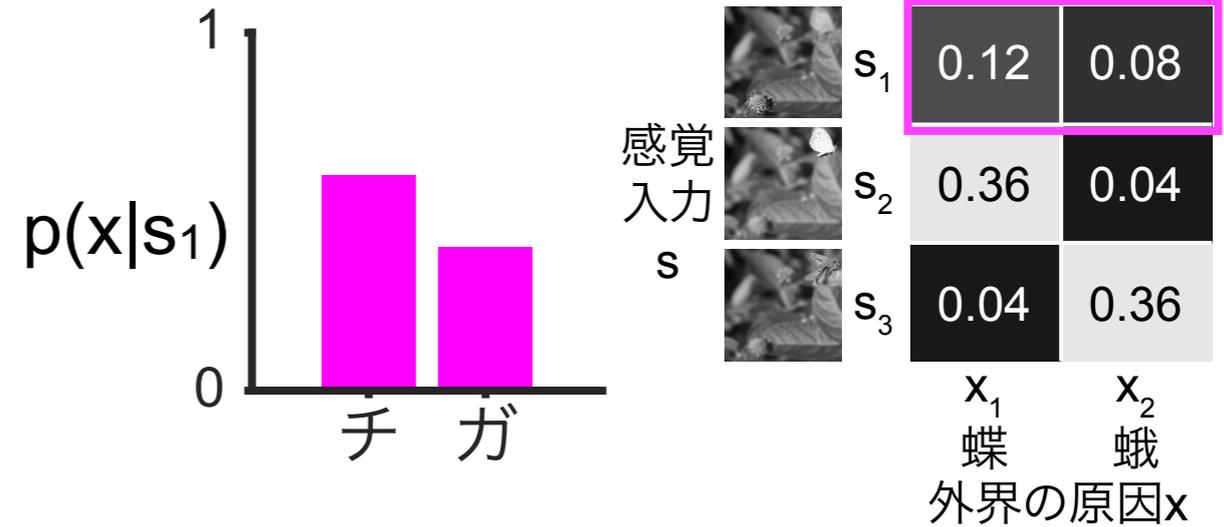


仮定：Agentはこれまでの経験から  
生成モデル(感覚入力 $s$ と原因 $x$ の同時確率)を持っている。

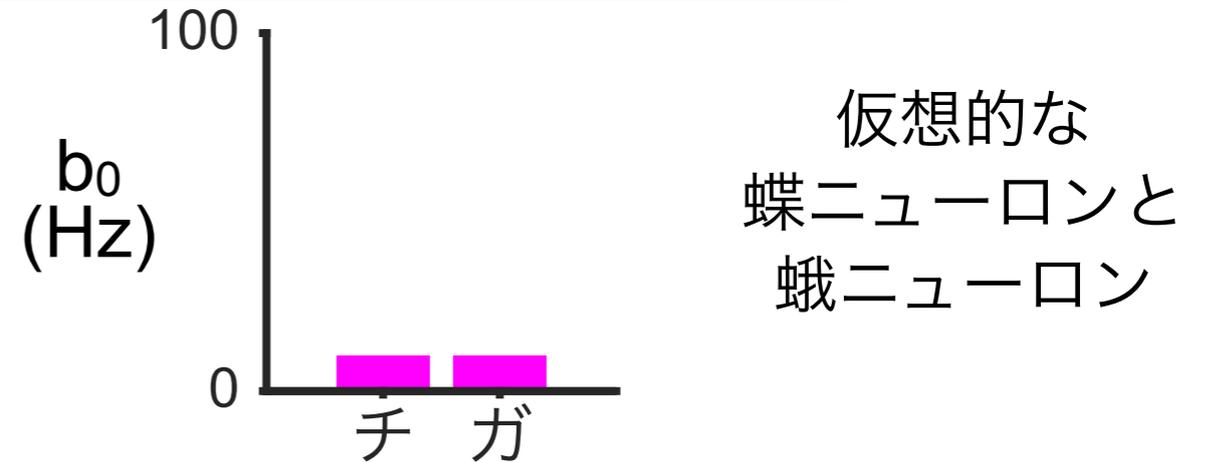
# FEPを実例を追って理解する



## 真のxの事後分布 (知識)



## 脳活動b

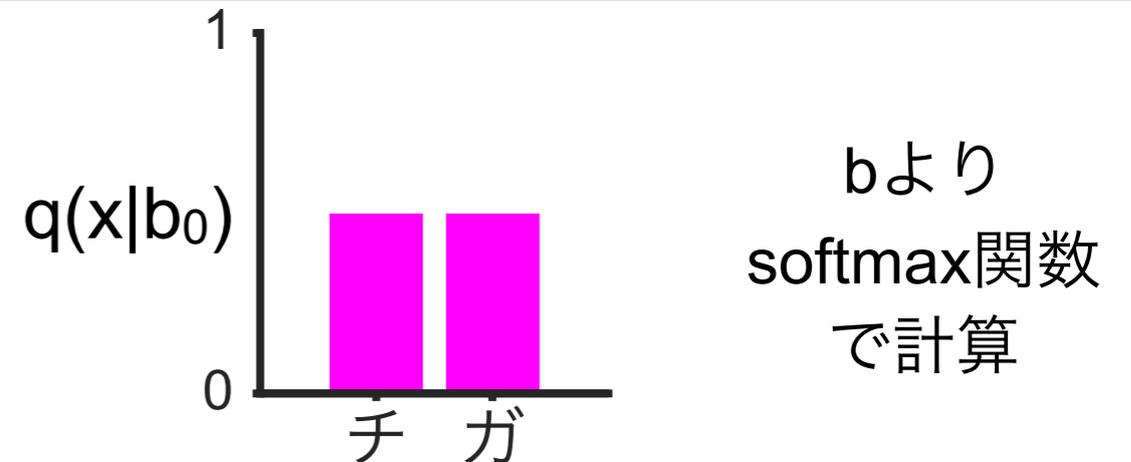


$$F(b) = D_{KL}(q(x|b) || p(x|s)) + C_1(s)$$

2.35 (bit)      0.02 (bit)

画像が提示された直後なので  
まだ蝶と蛾の判別は5:5

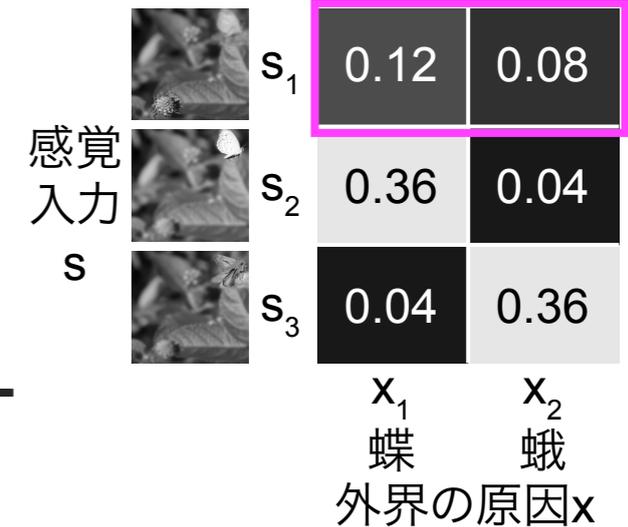
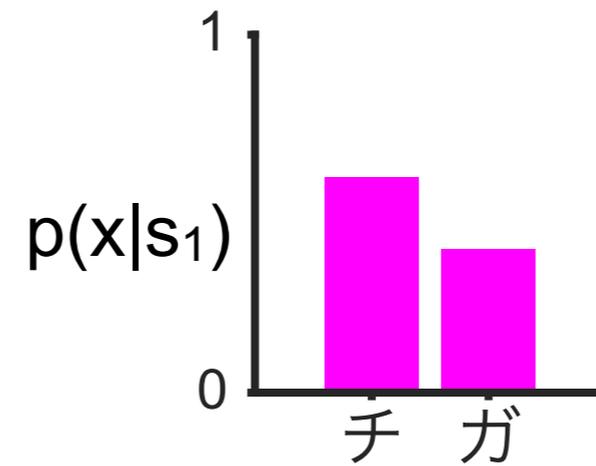
## 推測したxの事後分布 (現在の認識)



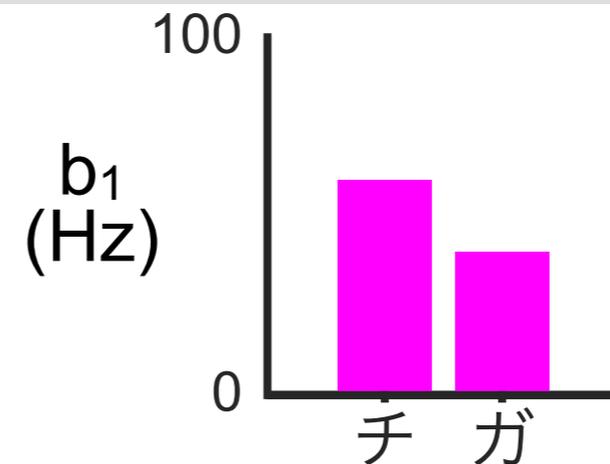
# FEPを実例を追って理解する



## 真のxの事後分布 (知識)



## 脳活動b



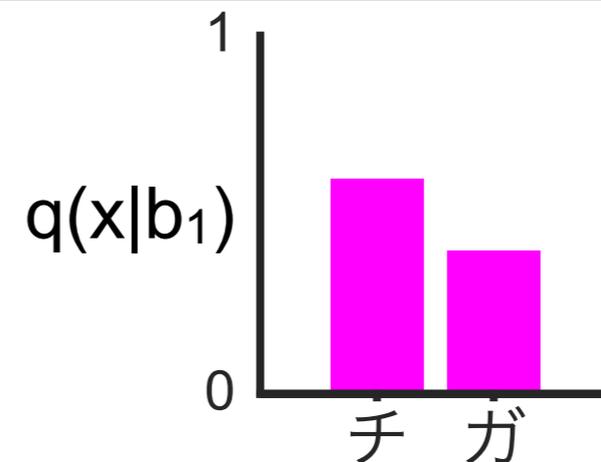
仮想的な  
蝶ニューロンと  
蛾ニューロン

$$F(b) = D_{KL}(q(x|b) || p(x|s)) + C_1(s)$$

2.32 (bit)      0.00 (bit)

Perceptual inference: 脳活動をb0からb1に変えると「推測した原因q」が変わる。「推測した原因q」が「真の原因p」と一致するとき、DKLは0になる。

## 推測したxの事後分布 (現在の認識)

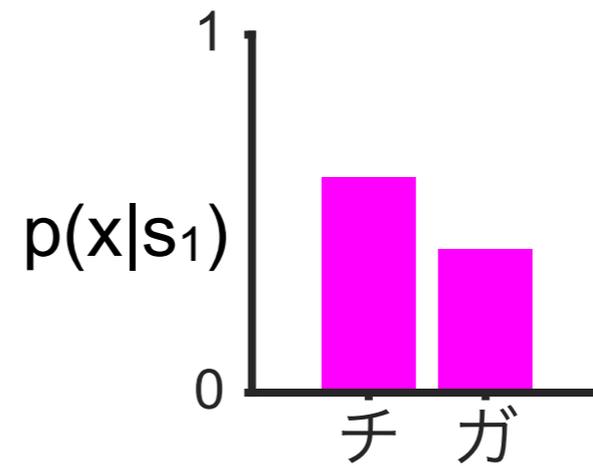


bより  
softmax関数  
で計算

# FEPを実例を追って理解する



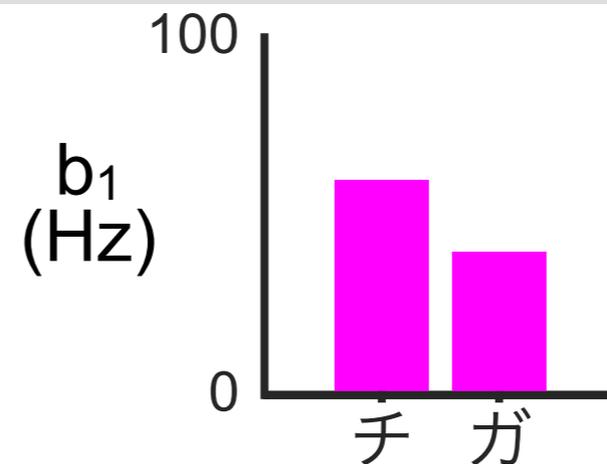
## 真のxの事後分布 (知識)



感覚入力 s	真のxの事後分布 (知識)	
	x <sub>1</sub> 蝶	x <sub>2</sub> 蛾
s <sub>1</sub>	0.12	0.08
s <sub>2</sub>	0.36	0.04
s <sub>3</sub>	0.04	0.36

外界の原因x

## 脳活動b



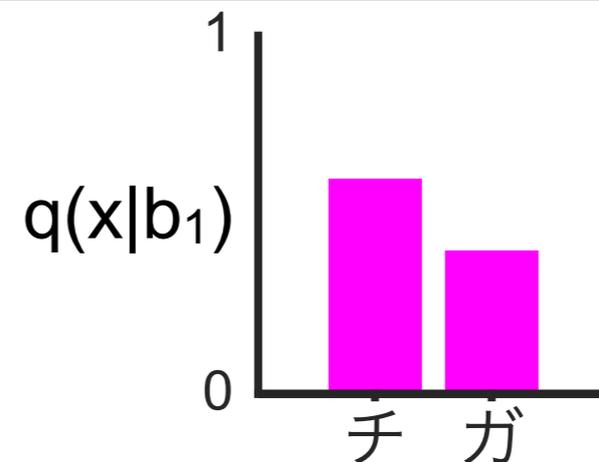
仮想的な  
蝶ニューロンと  
蛾ニューロン

$$F(b) = D_{KL}(q(x|b) || p(x|s)) + C_1(s)$$

2.32 (bit)      0.00 (bit)

原因xについてもっと良い推定  
をするためには、  
行動によってsを変えるしかない

## 推測したxの事後分布 (現在の認識)



bより  
softmax関数  
で計算

# FEPを実例を追って理解する



感覚入力 $s_1$

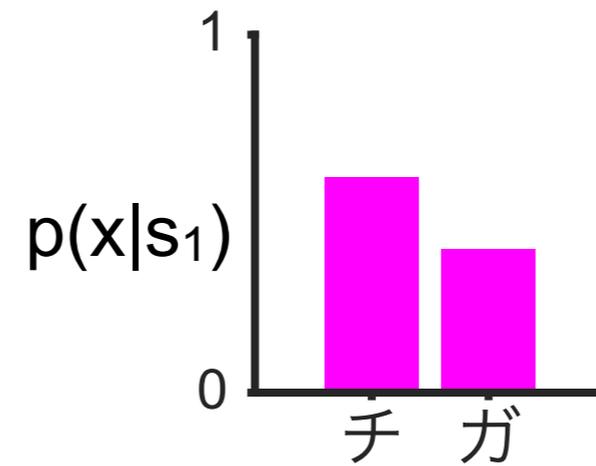
accuracy

$$F(s) = C_2(b) - \sum_x \{q(x|b) * \log(p(s|x))\}$$

2.32 (bit)                      -2.30 (bit)

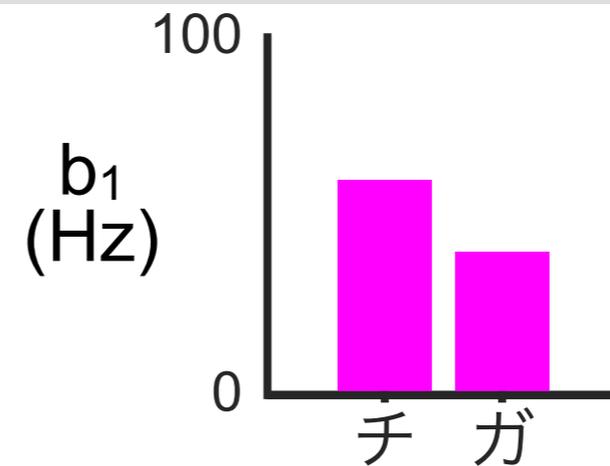
原因 $x_1$ (蝶)をより高い確率で  
サンプルできるのは感覚入力 $s_2$

真の $x$ の事後分布 (知識)



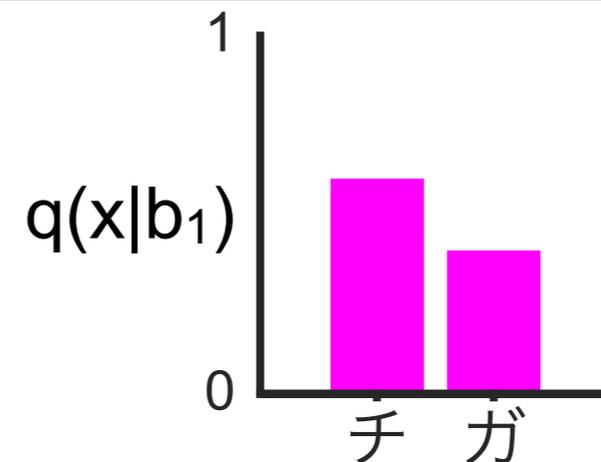
感覚 入力 s	$s_1$	0.12	0.08
	$s_2$	0.36	0.04
	$s_3$	0.04	0.36
		$x_1$ 蝶	$x_2$ 蛾
		外界の原因x	

脳活動b



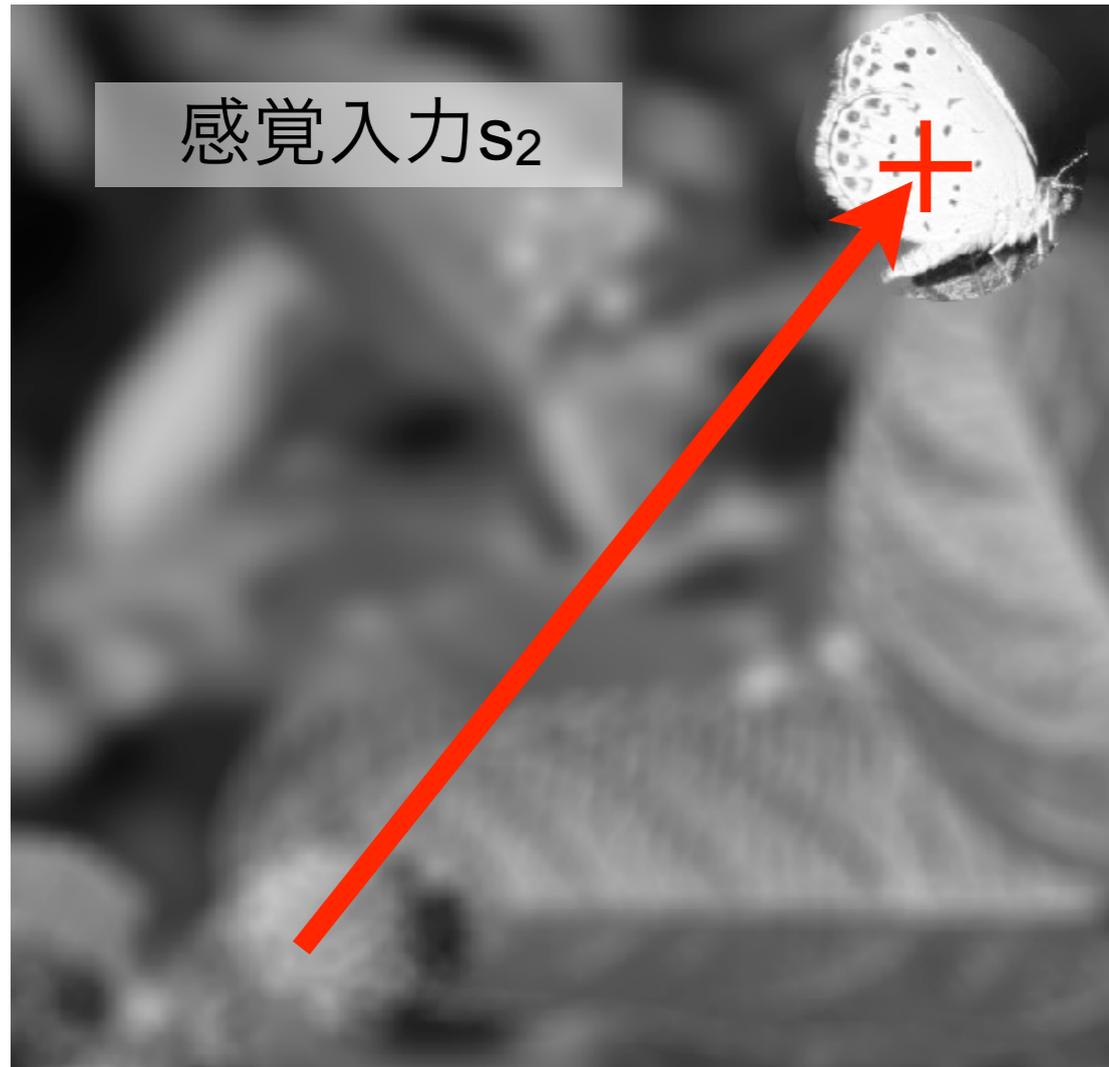
仮想的な  
蝶ニューロンと  
蛾ニューロン

推測した $x$ の事後分布 (現在の認識)

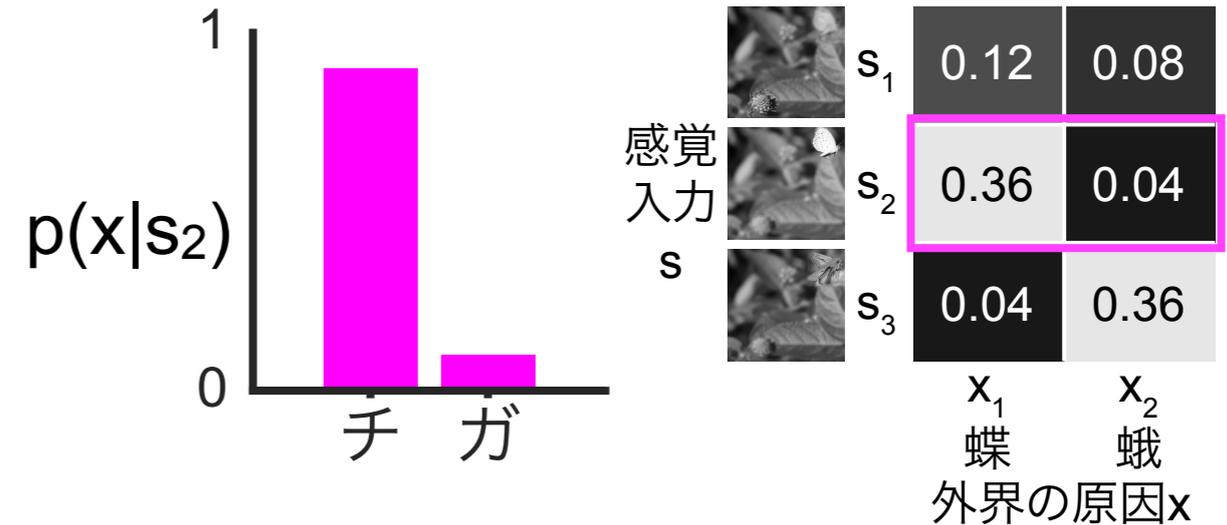


bより  
softmax関数  
で計算

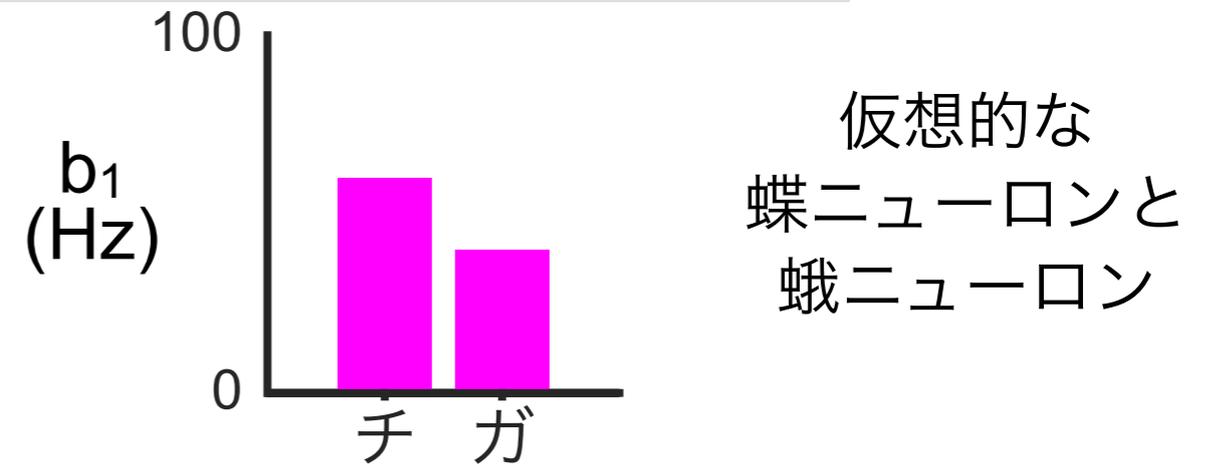
# FEPを実例を追って理解する



## 真のxの事後分布 (知識)



## 脳活動b



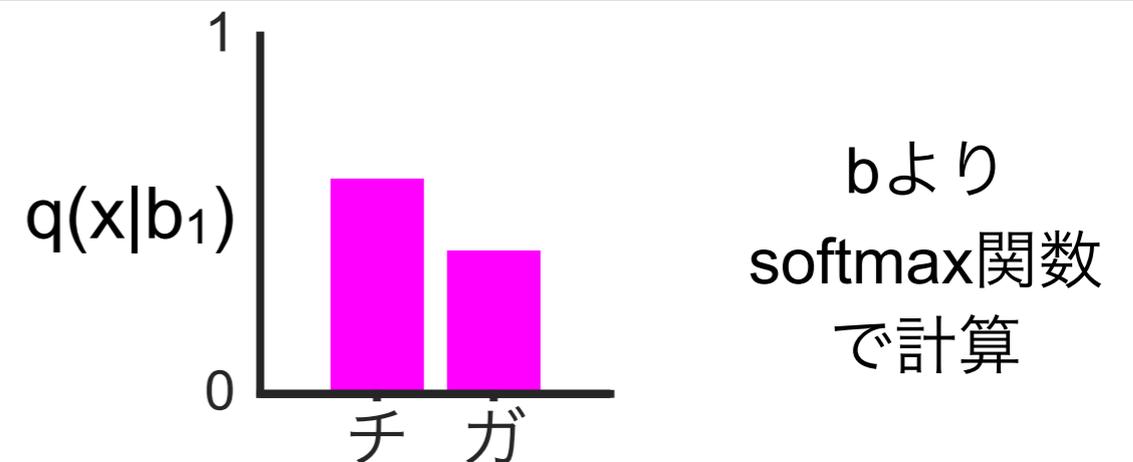
accuracy

$$F(s) = C_2(b) - \sum_x \{q(x|b) * \log(p(s|x))\}$$

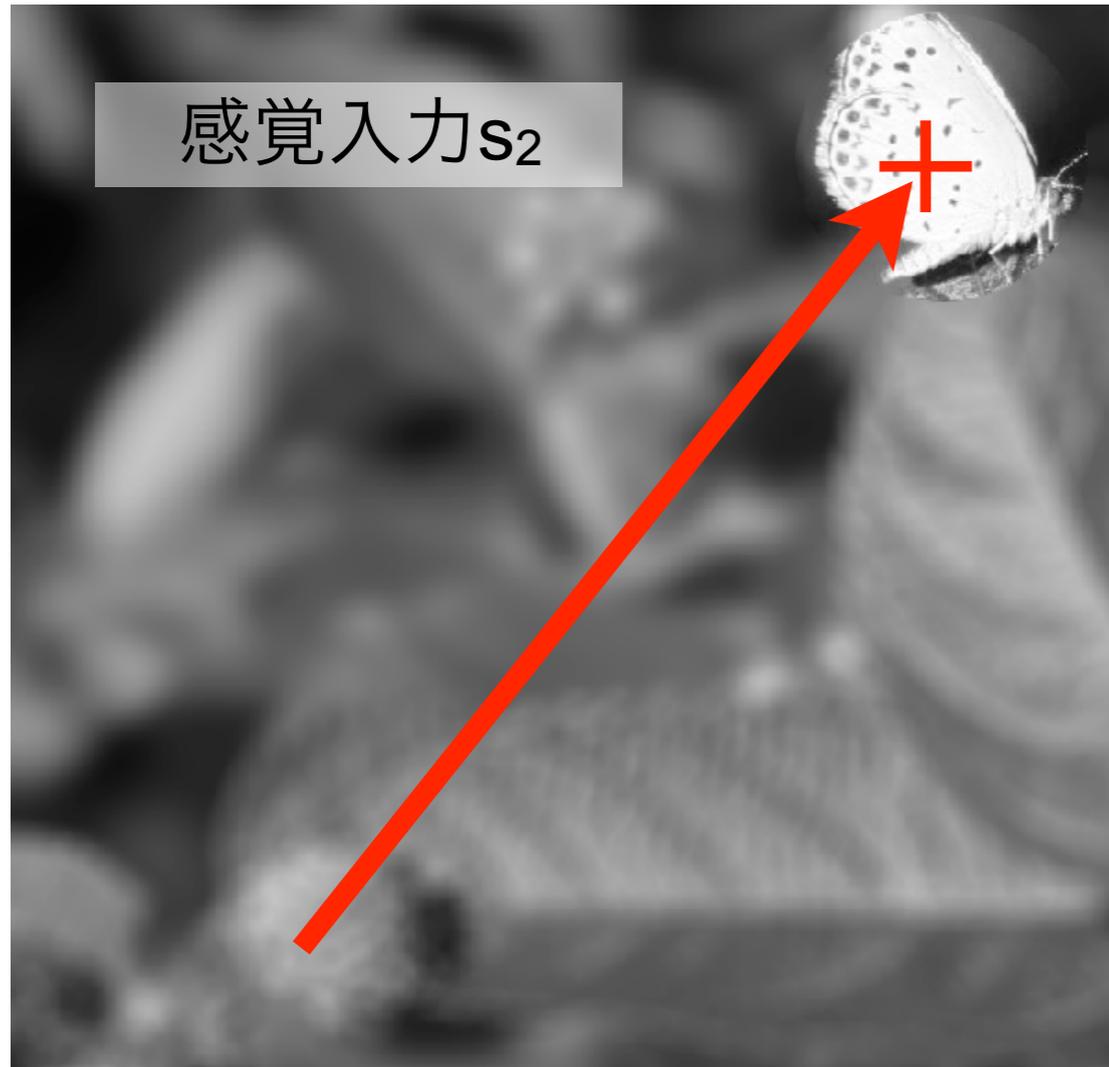
1.77 (bit)                      -1.75 (bit)

行動aで感覚入力s<sub>2</sub>をサンプルしたらaccuracyが上がった。  
(Active inference)

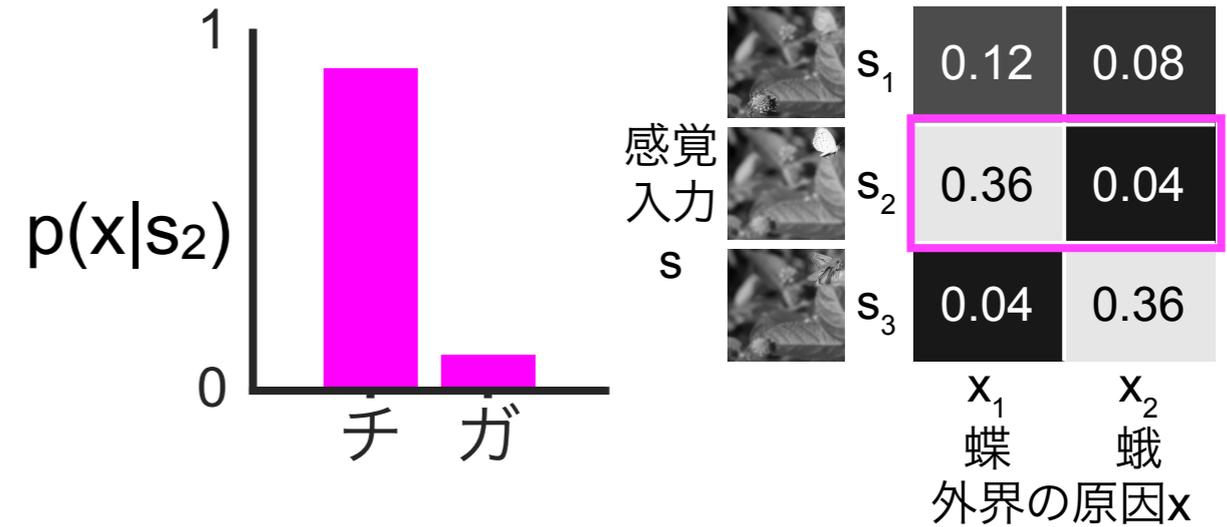
## 推測したxの事後分布 (現在の認識)



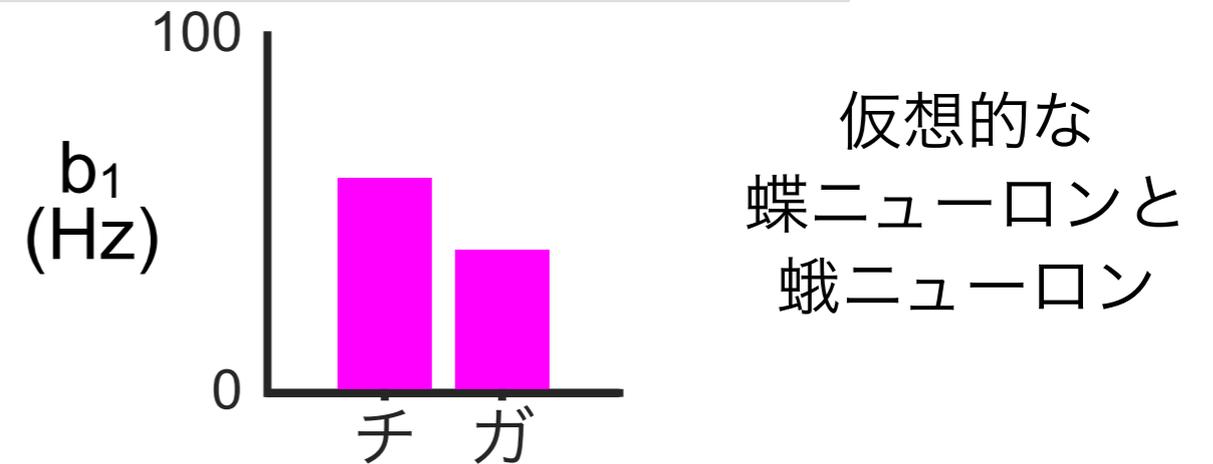
# FEPを実例を追って理解する



## 真の $x$ の事後分布 (知識)



## 脳活動 $b$



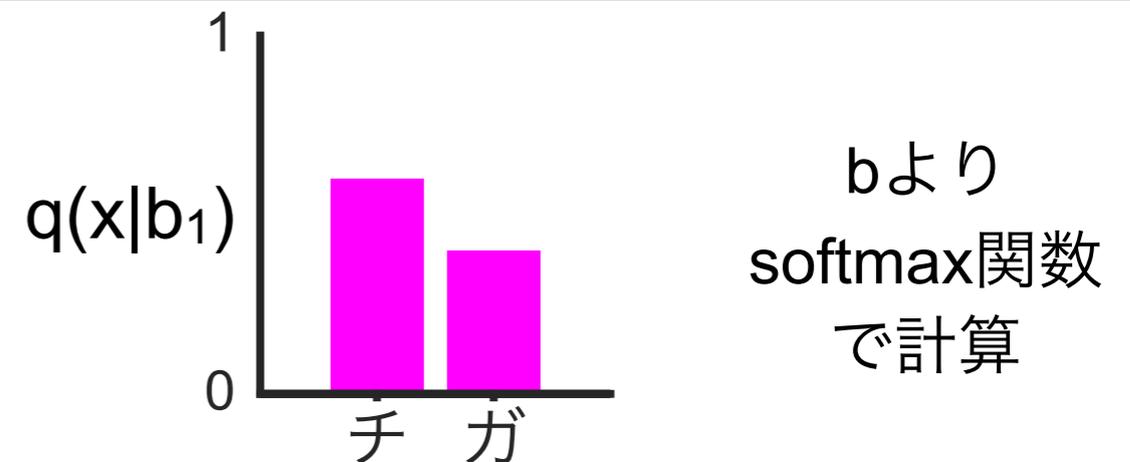
$$F(b) = D_{KL}(q(x|b) || p(x|s)) + C_1(s)$$

1.77 (bit)

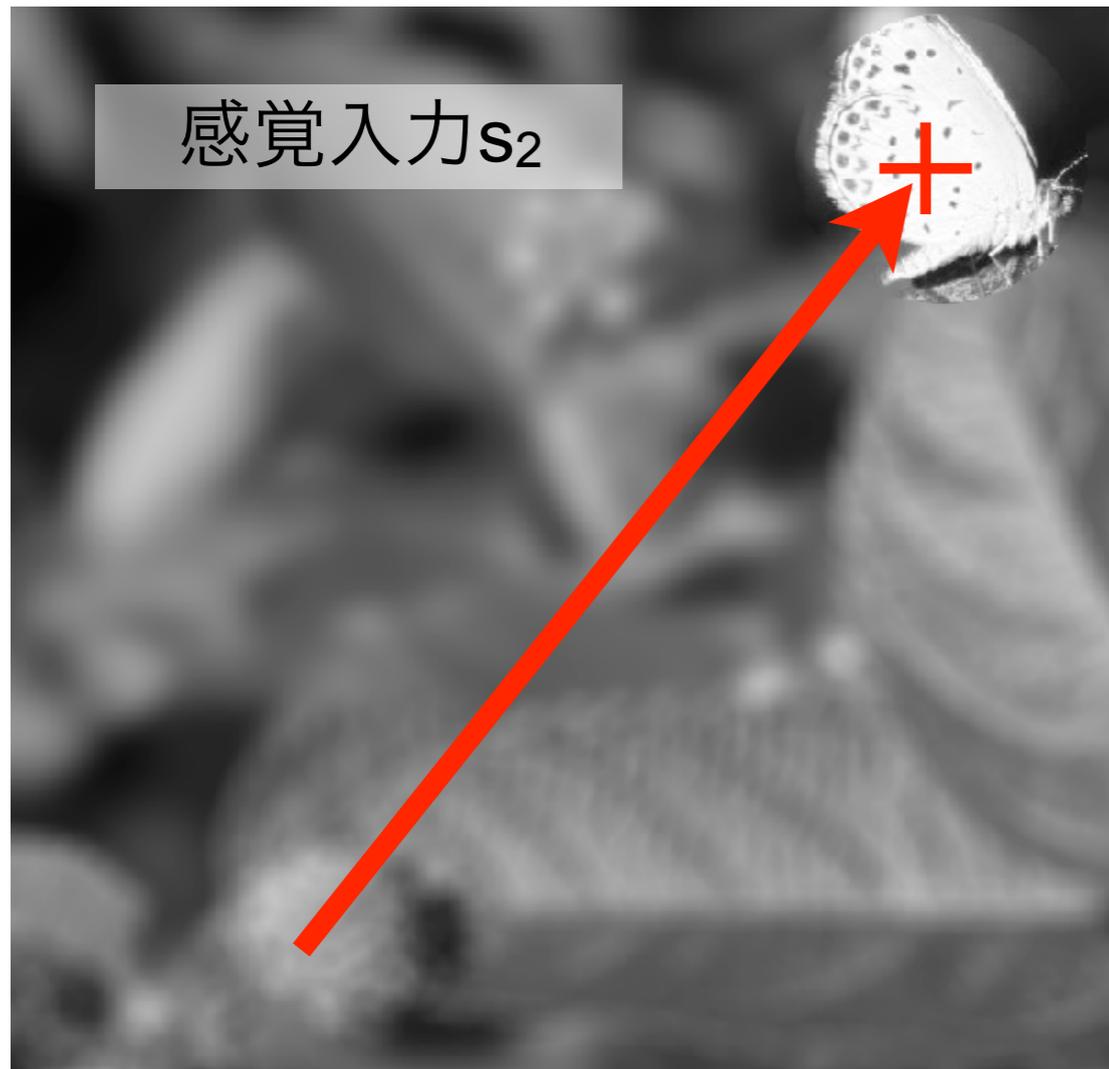
0.44 (bit)

ふたたび perceptual inference で  
DKLを下げる余地ができた。

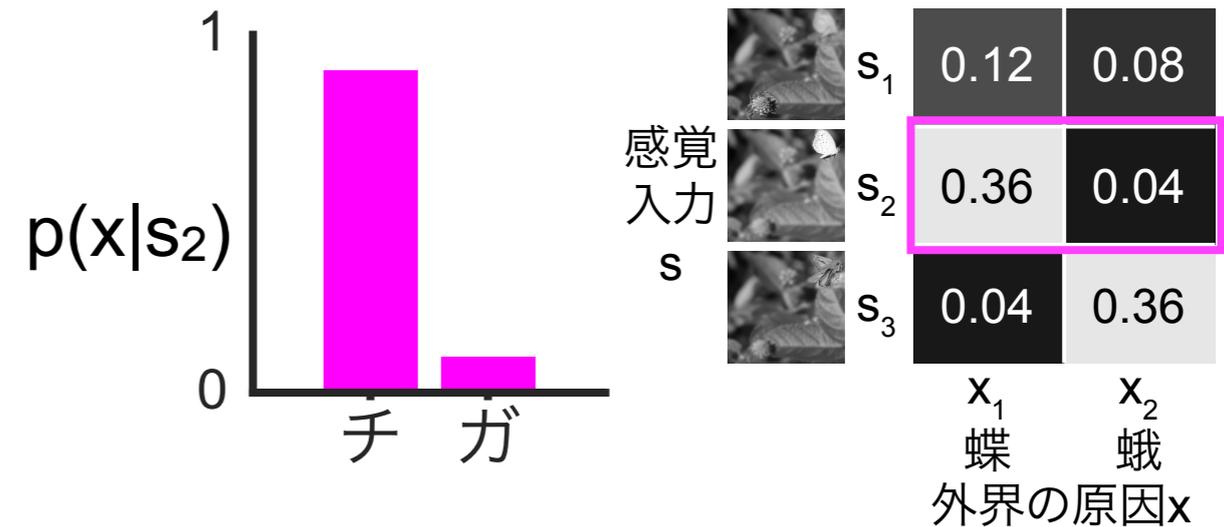
## 推測した $x$ の事後分布 (現在の認識)



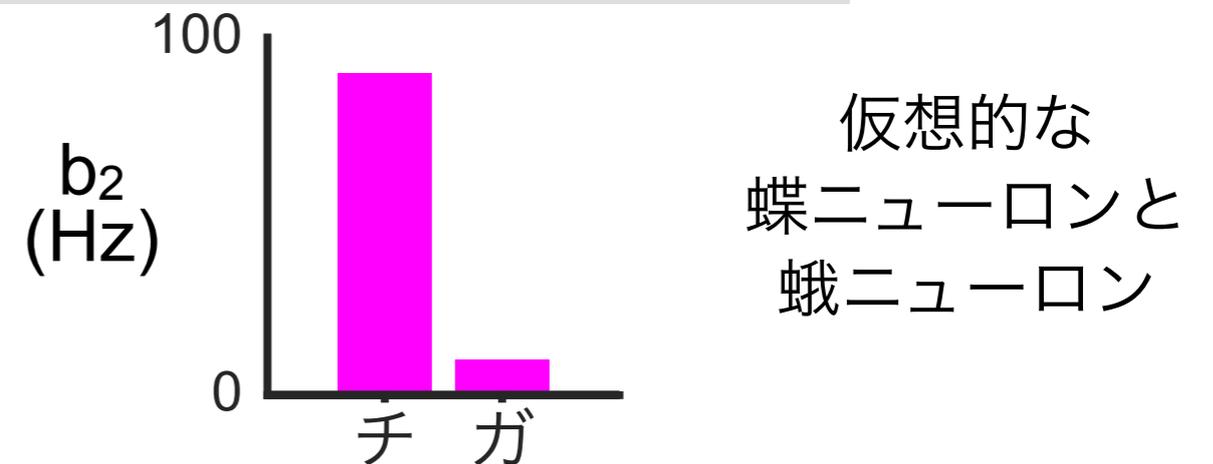
# FEPを実例を追って理解する



## 真の $x$ の事後分布 (知識)



## 脳活動 $b$



$$F(b) = D_{KL}(q(x|b) || p(x|s)) + C_1(s)$$

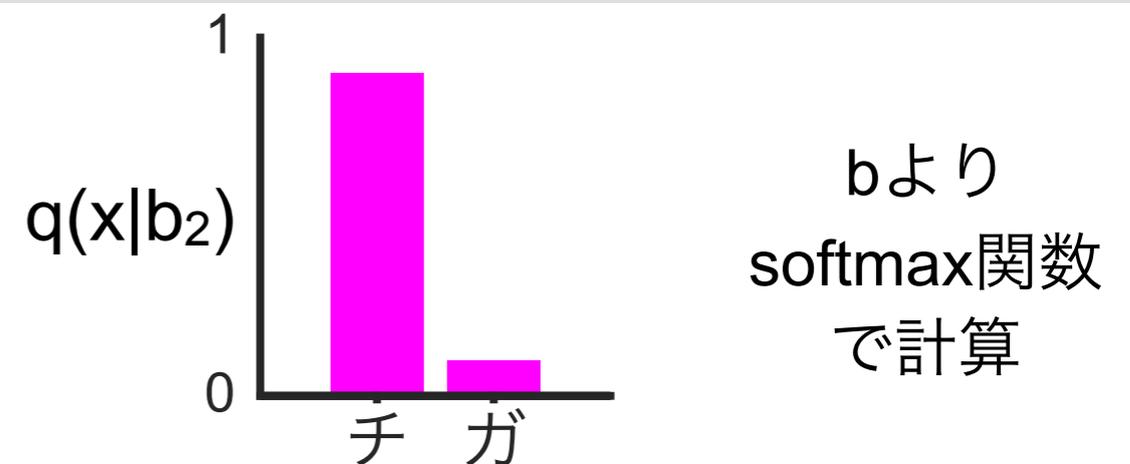
1.32 (bit)

0.00 (bit)

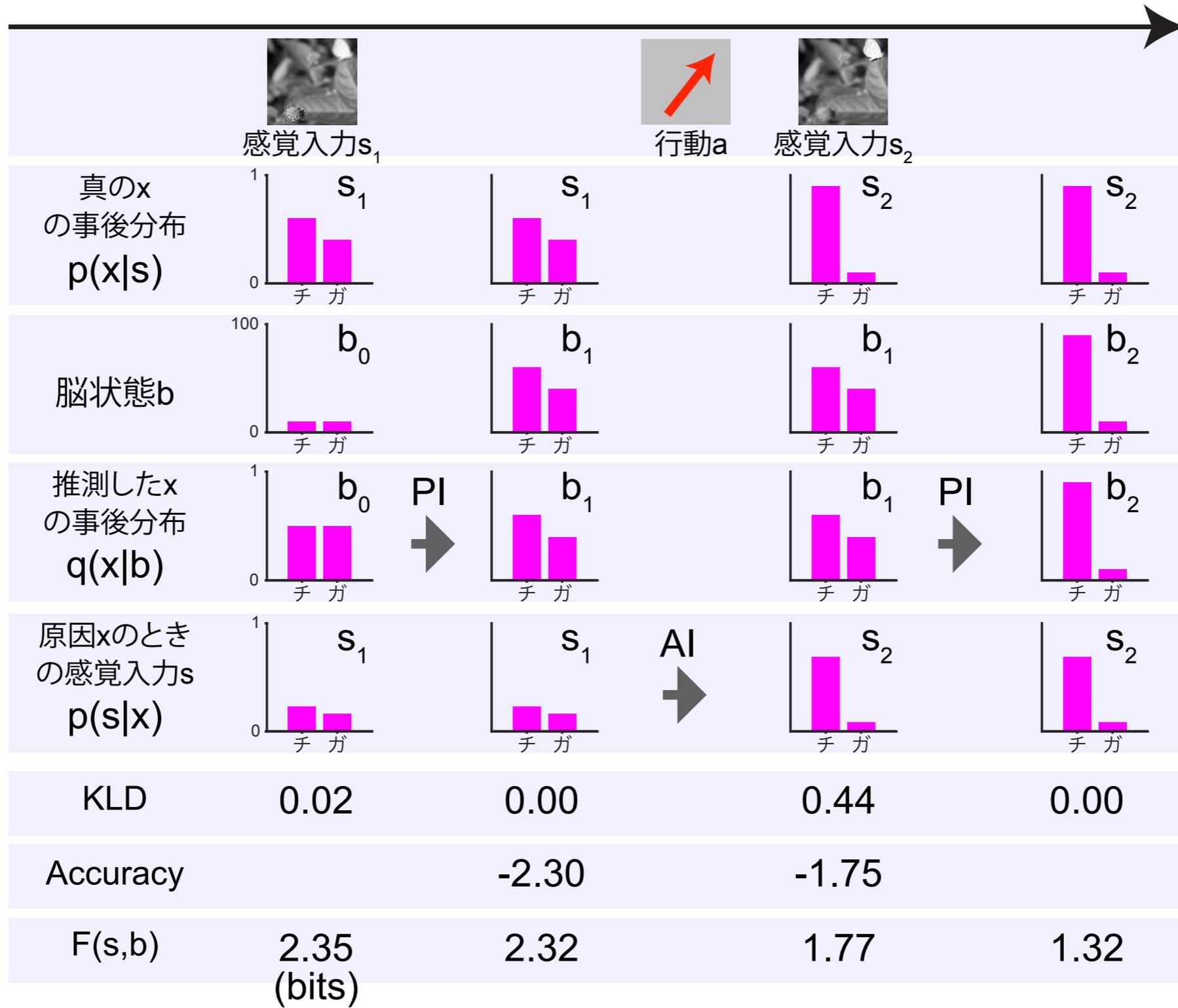
Perceptual inference: 脳活動を $b_1$ から $b_2$ に変えるとDKLは0になる。

$F$ は最小化した。

## 推測した $x$ の事後分布 (現在の認識)



# FEPを実例を追って理解する

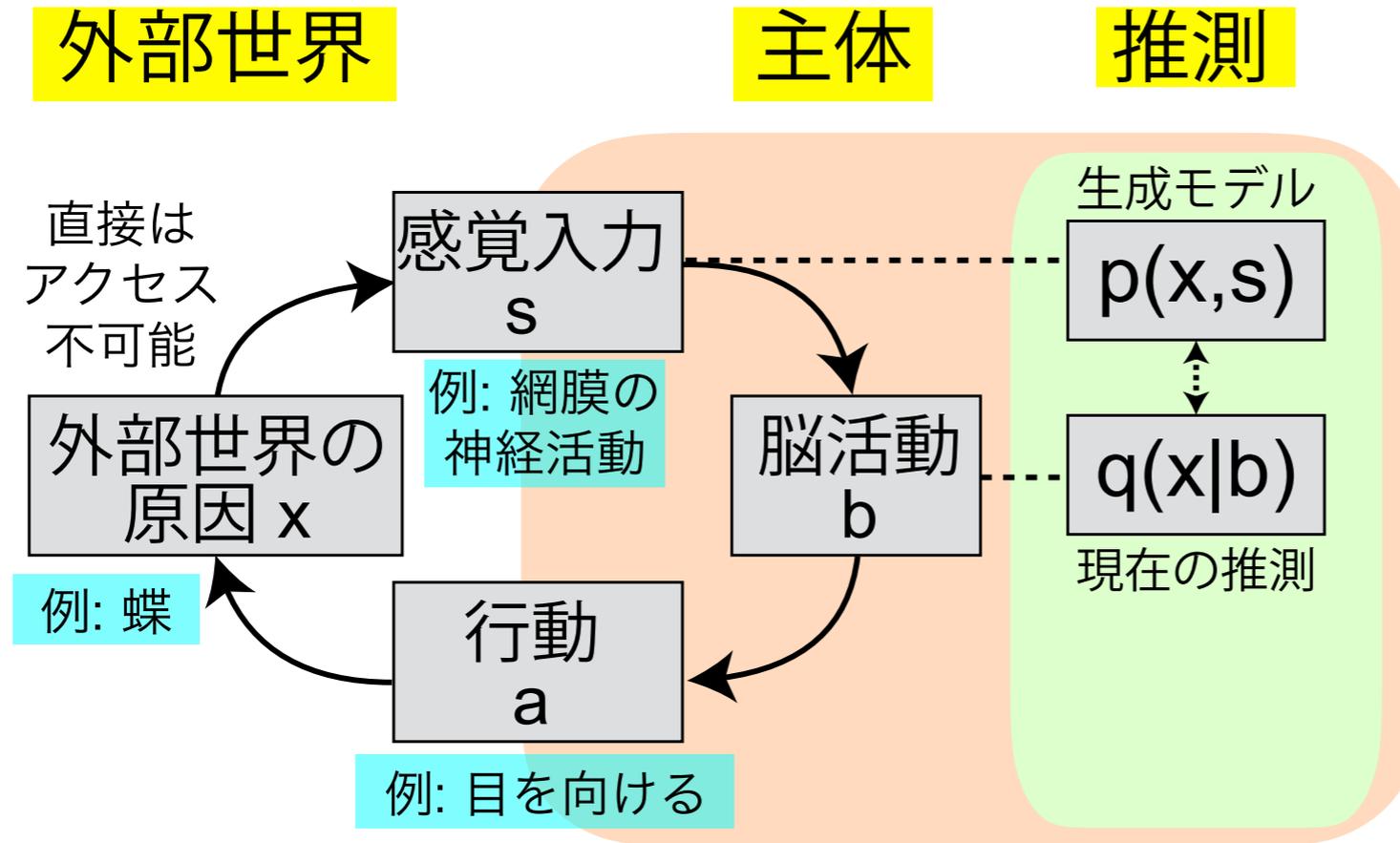


たしかに、Fは時間をおって小さくなっている。

$$F(b) = D_{KL}(q(x|b) || p(x|s)) + C_1(s) \quad \text{KLD} \quad \text{accuracy}$$

$$F(s) = C_2(b) - \sum_x \{q(x|b) * \log(p(s|x))\}$$

# 自由エネルギー原理 (FEP)



$$F(s, b) = D_{KL}(q(x|b) \| p(x, s)) \quad F \text{を最小化するには?}$$

## (1) Perceptual inference: 認識

脳活動  $b$  を変える  $\Rightarrow$  「推測した原因  $q$ 」を「真の原因  $p$ 」に近づける

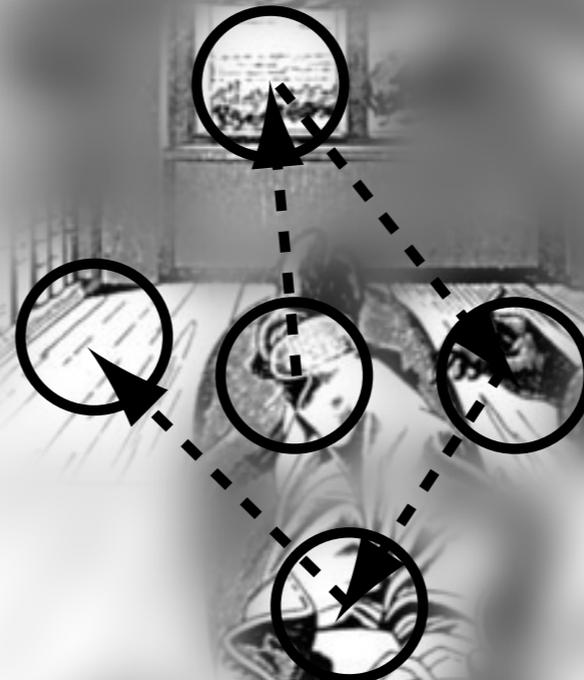
## (2) Active inference: 行動

感覚入力  $s$  を変える  $\Rightarrow$  「推測した原因  $q$ 」をより選択的に生み出す  
感覚入力  $s$  をサンプルする

# Active vision



エルンスト・マッハ「感覚の分析」より  
<https://publicdomainreview.org/collections/self-portrait-by-ernst-mach-1886/>



我々は絶えず  
視線を移動  
させながら  
視覚シーンを  
構成している。

Active vision: 視覚とは受け身での表象形成ではなくて、行動(例えば眼球運動)によって主体が視覚情報をサンプルすることである。

FEPはactive visionをPI-AIを組み合わせることで感覚入力sから原因xを明らかにしてゆく過程として整合的に説明している

# 全体的な流れ

1. 自由エネルギー原理とは

2. レベル別に説明

レベル1: 認識による世界のモデル化 (perceptual inference)

レベル2: + 行動による世界の探索 (active inference)

レベル3: + 反実仮想による世界のモデル化と目的の発生

# 行動は常に賭けである

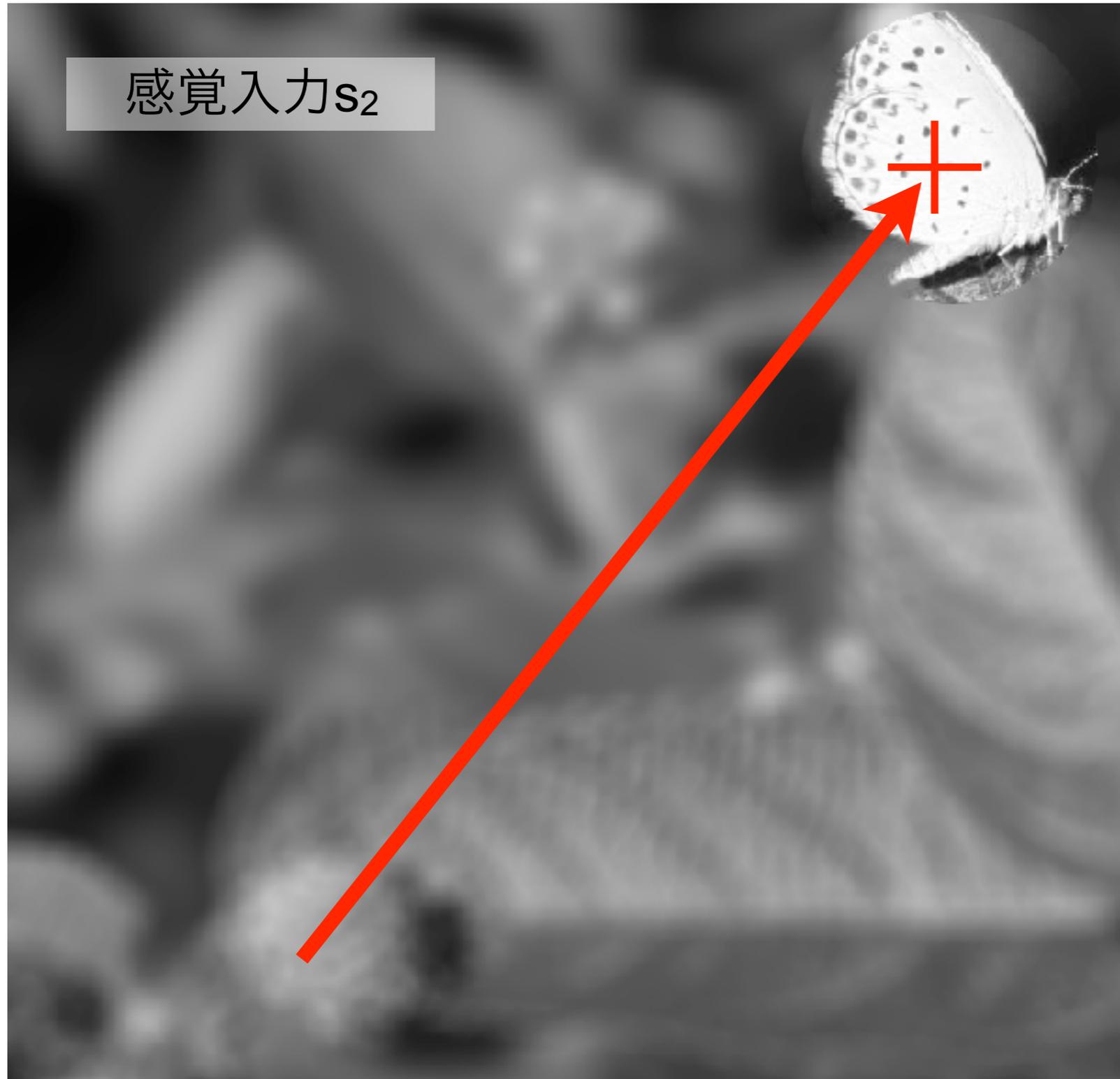
感覚入力 $s_1$



じつはレベル2での説明  
にはゴマカシがあった。

というのも、

# 行動は常に賭けである



「視線を右上に向けてみたら、蝶だった。」  
という状況をさきほどは例に挙げたわけだけど、

# 行動は常に賭けである



行動は賭けだから  
行動aの結果が  
必ずしも蝶だとは  
限らない。

感覚 入力 s		s <sub>1</sub>	0.12	0.08
		s <sub>2</sub>	0.36	0.04
		s <sub>3</sub>	0.04	0.36
			x <sub>1</sub> 蝶	x <sub>2</sub> 蛾
			外界の原因x	

$$P(x_1|s_1) = 0.6$$

$$P(x_2|s_1) = 0.4$$

なのだから、行動aの  
結果が蝶なのは60%  
の場合でしかない。

# 行動は常に賭けである



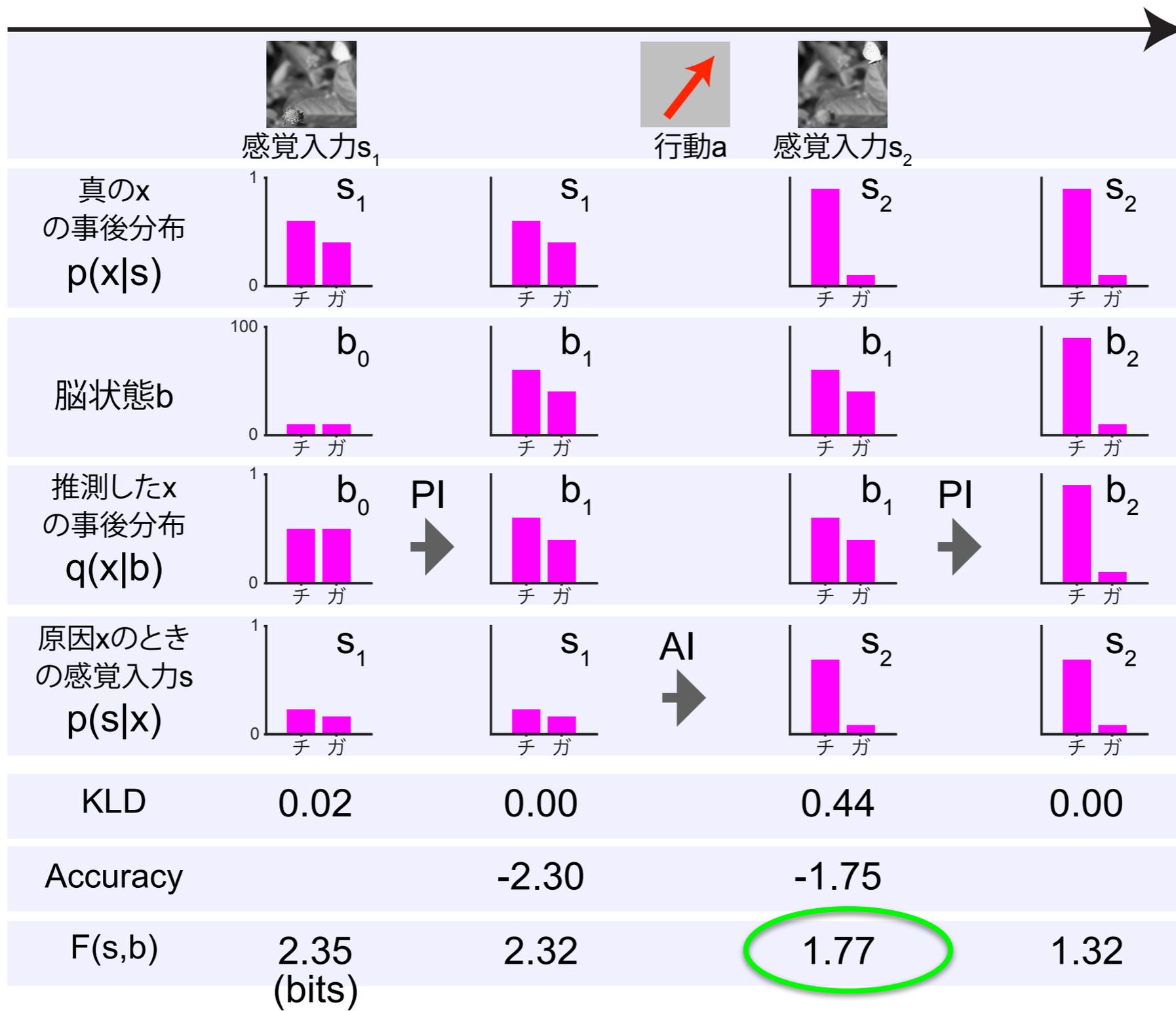
視線を右上に向けて  
みたら、じつは  
蛾だった。

感覚 入力 s		s <sub>1</sub>	0.12	0.08
		s <sub>2</sub>	0.36	0.04
		s <sub>3</sub>	0.04	0.36
			x <sub>1</sub> 蝶	x <sub>2</sub> 蛾
			外界の原因x	

$$P(x_1|s_1) = 0.6$$

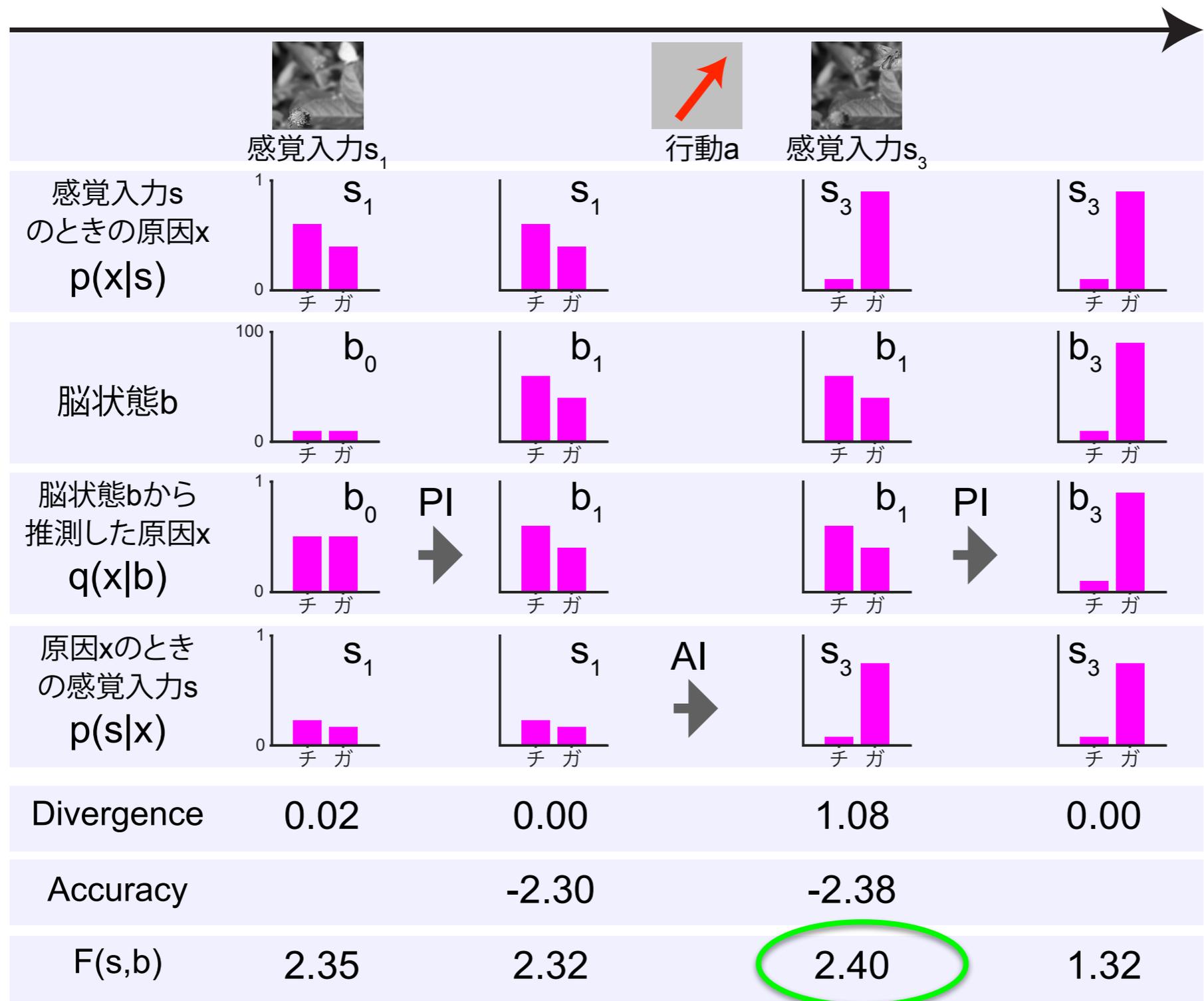
$$P(x_2|s_1) = 0.4$$

ということが40%の  
確率で起こる。



視線を右上に向けてみたら、蝶だった。

この場合は行動  $a$  によって  $F$  が下がったけど、



視線を右上に  
向けてみたら、  
蛾だった。

こちらの場合は  
行動aによってF  
が上がっている。

予想していた原因 ( $x_1$ =蝶)を反証する感覚入力 $s_3$ をサンプルしたため、  
accuracyは低下し、Fはかえって大きくなった。

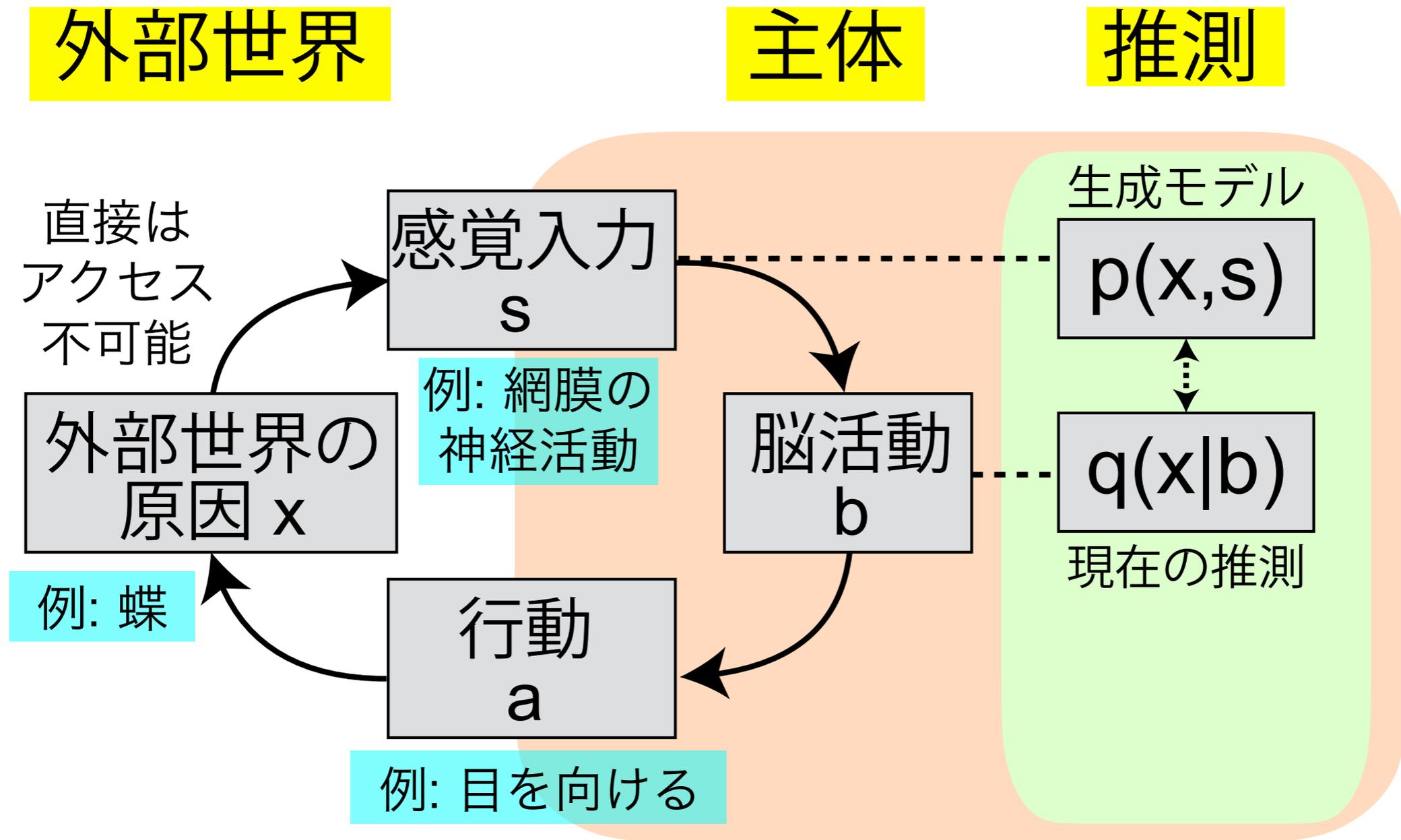
# FEPは期待値でのみ成立する



FEPはふたつの可能性(蝶と蛾)の期待値においては成立している。

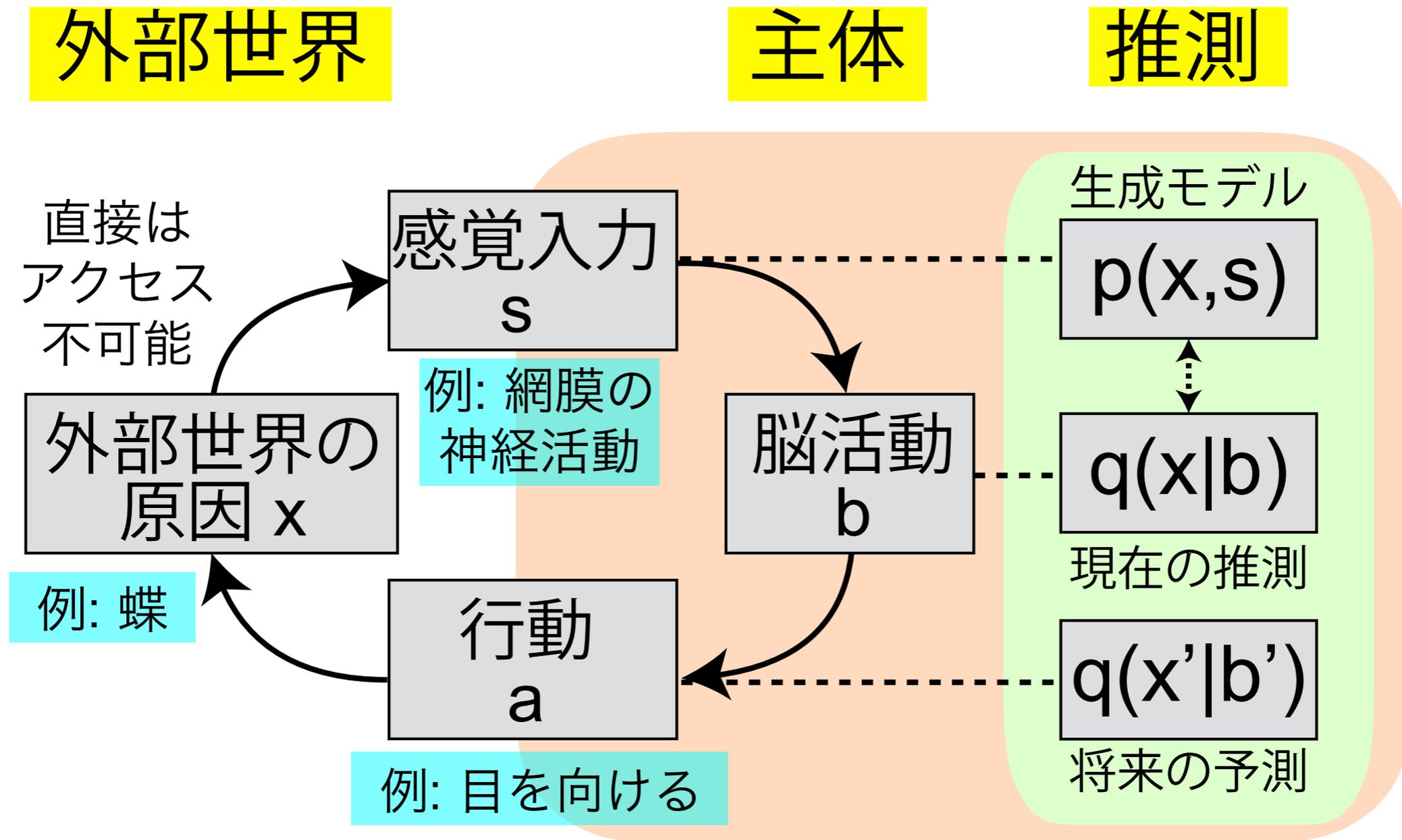
つまり、Agentは将来平均的にFを下げるのが期待される行動aを選択している。これは反実仮想 (まだ実現していない行動aによって蝶という原因を期待する)だ。

# 自由エネルギー原理 (FEP)



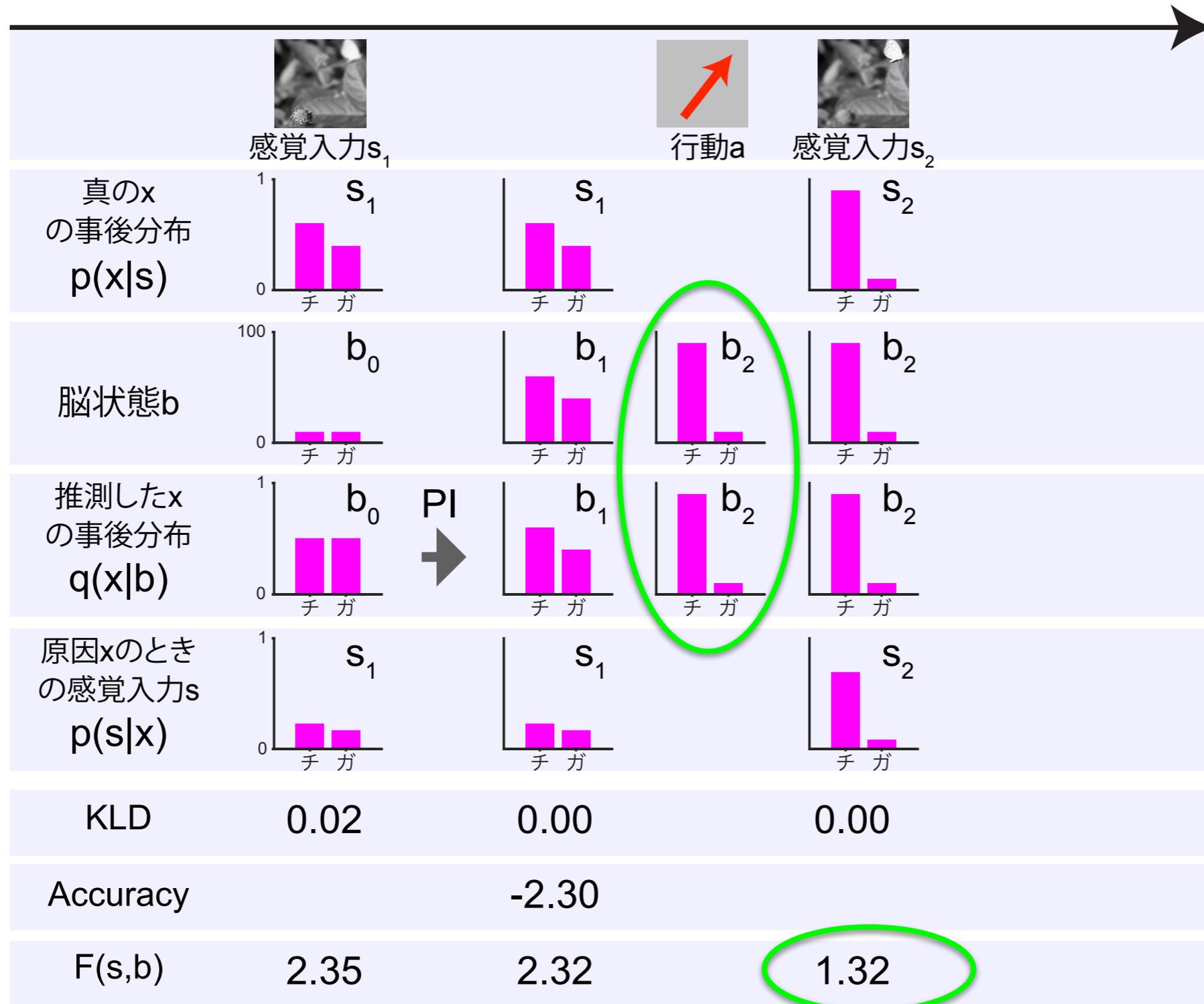
これまでの説明の図では「現在の推測」だけが考慮されていたけれど、

# 自由エネルギー原理 (FEP)



「現在の推測 $q$ 」だけでなく、  
「反実仮想的な将来の予測 $q$ 」も生成している。  
(ここでは $x'$ ,  $b'$ は将来の $x$ ,  $b$ についての表示)

# 反実仮想的な推測



行動の時点で脳活動が未来を先取りするのは、神経生理で知られている presaccadic remapping によく似ている。

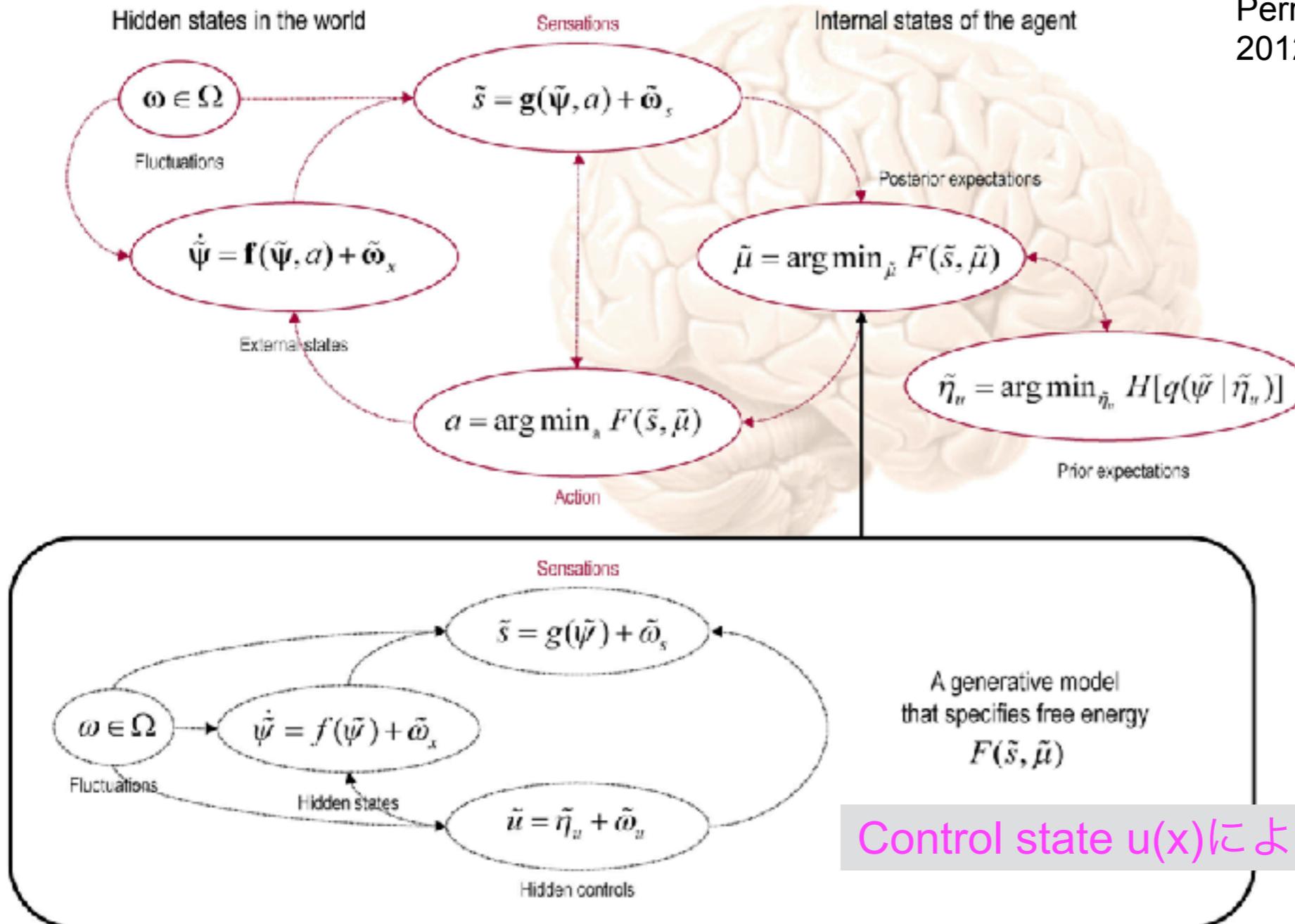
Fが下がっている。

反実仮想的な将来の予測  $q(x'|b')$  を生成することで行動の開始段階で脳状態が未来を先取りしている。

# 自由エネルギー原理 (FEP)

x (or psi)とsは時間的なシーケンス

Perceptions as hypotheses: saccades as experiments. Friston K, Adams RA, Perrinet L, Breakspear M. Front Psychol. 2012 May 28;3:151.



視覚サリエンスが高い  
= そこを見ればqのエントロピーを最小化する場所 => その場所を control state u(x)のprior として採用する。

Control state u(x)によって運動aを計画

実際のモデルはこの図のようにもっと複雑なのだけど、エッセンスはこれでつかめたと思う。

# 参考文献

- Friston K, Adams RA, Perrinet L, Breakspear M. (2012) Perceptions as hypotheses: saccades as experiments. *Front Psychol.* 3:151. (視線移動のモデル及び反実仮想の概念の初出)
- McGregor, Simon; Baltieri, Manuel; Buckley, Christopher L. (2015) A Minimal Active Inference Agent. arXiv:1503.04187 <https://arxiv.org/abs/1503.04187> (このスライドの説明の元ネタ。離散バージョンのFEPの説明などがあり、これがもっともわかりやすい。)
- Buckley, Christopher L.; Kim, Chang Sub; McGregor, Simon; Seth, Anil K. (2017) The free energy principle for action and perception: A mathematical review. arXiv: 1705.09156 (省略無しで丁寧な説明。)
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138. (FEPの全体像について網羅的な記述)
- Friston K, Rigoli F, Ognibene D, Mathys C, Fitzgerald T, Pezzulo G. (2015) Active inference and epistemic value. *Cogn Neurosci.* 2015;6(4):187-214. (視線移動を epistemic value と捉えて、他の value と対置して扱っている)
- Ryota Kanai (2017) Creating Consciousness. <https://www.slideshare.net/ryotakanai/creating-consciousness> (FEPの反実仮想についての言及あり。)

# 背景説明

このスライドは栢森情報科学振興財団 第17回 Kフォーラム2017(<http://www.kayamorif.or.jp/forum/f.html>)での発表をもとにして作成した。

以前作成したスライド「アクティブビジョンと フリストン自由エネルギー原理」 <https://www.slideshare.net/masatoshiyoshida/20170111-70980954> の更新版として作成された。以前のスライドについてはブログ記事(「セミナー「アクティブビジョンと フリストン自由エネルギー原理」スライドをアップロードしました」 <http://pooneil.sakura.ne.jp/archives/permalink/001607.php>)にさらに情報あり。

以下に記したように、FEPを正確に理解するためには機械学習(変分ベイズ)、そしてニューラルネットワーク(確率的勾配降下法)についての知識が必要となる。しかしそもそもそこまでしてFEPを理解する意義があるか、判断の材料を提供するために、神経科学者、心理学者、哲学者に向けてFEPの最小限の理屈を説明したい、というのが本スライドを作成した動機。

ゆえに、この説明では(世界一)単純化した例を用いているということを承知しておいていただきたい。私自身はFEPをどのくらい重要だと思っているかというと、批判的に継承したいと考えている。このあたりについては私のブログ内の記事を参照してほしい。

## FEPを正確に理解するためには:

金井さんのASSC2017チュートリアルにも書かれているように、フリストンの書いたものをいきなり読むのはハマり道(表記の統一などにも難があるし)。まずは、機械学習、とくに変分ベイズについてPRMLなどの教科書でよく理解することが大事。そのうえで、SPM12のDEM toolboxにあるmatlabコードを走らせ、コードを理解するという手順がよいと思われる。