

# Title: 自由エネルギー原理入門 改め 自由エネルギー原理の基礎徹底解説

## 生理研・認知行動発達・吉田正俊

今年の生理研研究会は「認知神経科学の先端 脳の理論から身体・世界へ」と題して、自由エネルギー原理(Free-energy principle, FEP)をテーマに9/2に開催します。これに先立つ8/31-9/1には「脳の自由エネルギー原理チュートリアル・ワークショップ」というタイトルでFEP入門のためのレクチャーとハンズオンあり。

これらに向けてFEP入門の資料を作りました。ぜひこれを使って予習してみてください。

参加募集を5/17に開始しました。ぜひご応募ください。

---

## Sec.0 自由エネルギー原理を数式無しで説明する

### [0-1. 自由エネルギー原理の定義]

この文書では自由エネルギー原理とはなにか、について概念的な説明からスタートして、最終的には数式を用いた詳細な理解と批判が可能になるところまでたどり着くことを目的としている。

まず自由エネルギー原理(Free energy principle,以下FEPと呼ぶ)とはなにか。脳イメージングの解析ソフトSPMの作者として著名な、University College Londonの研究者Karl Fristonが提案している、知覚と行動と学習の統一原理だ。2005年のPhilos Trans R Soc Lond B Biol Sci.論文で最初に提案されてから、現在まで理論的にも進歩を続けている。

Friston自身の定義を見よう。自由エネルギー原理とは「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」というものだ。また別の表現では「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」とある。

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

順番に言葉を追っていこう。まず「原理」というだけあって、「もしxxであるなら、yyでなければならない」という形式になっていることがわかる。ベイズ脳仮説のような「仮説」でもなければ、予測符号化理論のような「理論」とも区別した表現になっている。

「いかなる自己組織化されたシステム」とあるけれど、われわれ人間にかぎらず、様々な生物に当てはまることが想定されている。「システム」とあるので、生物でなくても成り立つけれども、たとえば氷の結晶には当てはまらないだろう。氷の結晶はシステムとして安定する境界を持っていないから。

そうすると「環境内で平衡状態でありつづける」というのは、氷の結晶が氷の結晶であり続ける話ではなくて、生命のあるものが生き続ける条件のことを言っているようだ。

そして「そのシステムの(情動的)自由エネルギー」というものが定義されるものでなければならない。(情動的)自由エネルギーというのがなんなのかはこれから順番に説明をしてゆくとして、いま知っておくべきは、そのシステムと外界との間での情報のやり取りに関わる概念なので、システムと外界との境界が必要なのだ。

ではシンプルな例でもっと具体的に表現してみよう。

## [0-2. 知覚=現在の外界の状態の推定]

図0-1の例ではagentが外界に接している状態が表現されている。先程も書いたようにこの「自己組織化されたシステム」は人間や生物に限らない。そのことを示すために以降「agent」という呼び方で統一する。

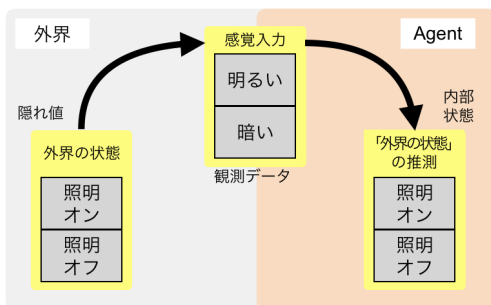


図0-1: 知覚の例

いま使う説明では、世界自体はわれわれの現実世界とまったく同じものなのだけど、agentが世界を切り分ける能力が低いので「外界の状態」は2つしか区別できない。「照明オン」と「照明オフ」だ。Agentは照明オンかオフかについて直接アクセスすることはできない。つまり「外界の状態」はagentにとって隠れ値だ。

Agentは「外界の状態」を推測するために「感覚入力」を用いる。ここでは網膜のような光センサーがあって、照明の明るさに従って「明るい」「暗い」という2つのどちらかの値を時々刻々観測している(agentの識別能力が低いので2つの値しか区別できない)。Agentはこの観測データにだけアクセスできる。いま「感覚入力」と書いたが、「感覚sensation」と「知覚perception」を区別するためにこの言葉を使っている。たとえばわれわれが写真を見たとき、網膜の視細胞の活動のようなセンサー値が「感覚入力」だ。そしてその写真になにが写っているか知ることが「知覚」だ。

Agentは「感覚入力」を元にして、「外界の状態」がいまどうなっているかを推定する。これが「知覚」だ。たとえば「感覚入力」が「明るい」を観測したなら、「外界の状態」は「照明オン」である可能性が高いだろう。100%高いとは言えないことはわれわれは経験上知っている。「照明オン」でも「感覚入力」が「暗い」を示すこともありうるだろう(照明とセンサーの間になにか邪魔なものがあるかもしれない)。よってこの設定では、知覚とは、「外界の状態の推定」とは、あくまで確率的に表現される。「外界の状態が照明オンである確率は90%」というように。確率で表すことによって、推定がどのくらい確実かというuncertaintyも表現することができる。

でもそもそもなぜ「感覚入力」から「外界の状態」を推定できるかといえば、agentはこれまでの経験から、「外界の状態」がどのように「感覚入力」に影響を与えるか、その因果関係について学習しているからだ。外界におけるこの関係を「生成過程」と呼び、agentが学習したこの関係を「生成モデル」と呼んで区別する。「生成過程」は外界の物理法則そのものだが、「生成モデル」はそれを写し取ったモデルでしかない。ゆえにモデルは間違っている可能性がある。今の場合も3次元の世界で照明からセンサーに光が届く生成過程があるのだけど、それをひとつの光センサーしかもたないagentは生成過程を1点に投射されたものとして生成モデルを獲得している。

### [0-3. 行動選択=未来の外界の状態の推定]

しかしこのような設定ではagentは「桶の中の脳」と同じで、外界の生成過程を正しく生成モデルとして維持する方法がない。ここで行動を考える必要が出てくる。行動を含めた世界設定の図を示す(図0-2)。

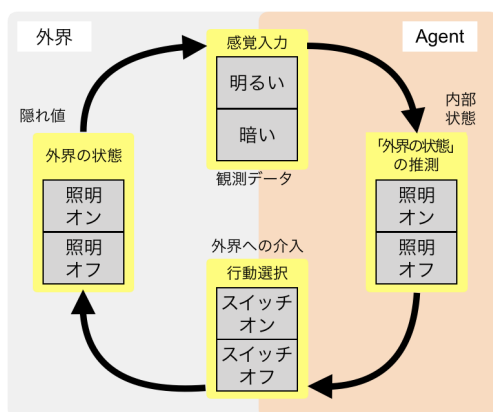


図0-2: 感覚運動ループ

ここでは「外界の状態」が「感覚入力」という観測データを生み出し、「外界の状態」を推定するagentの内部状態が「行動選択」という「外界の状態」への介入を行うというループが閉じている。これを感覚運動ループと呼ぶ。

このループを使うことで、agentの「生成モデル」は外界の「生成過程」と整合的であるように維持される。たとえばいまagentは「感覚入力」が「暗い」を観測していて、「外界の状態の推定」(=知覚)として「照明オフ」の確率90%としている。これを確かめるために、「行動選択」を「スイッチオン」にして、「感覚入力」が「明るい」になれば、現在の「外

界の状態」が「照明オン」であるという推定(=知覚)の根拠となったagentの「生成モデル」は正しく機能していることが確認されるので、そのまま維持すればよいということがわかる。

この「行動選択」では、これからする行動(スイッチオン)が「未来の外界の状態」を「照明オン」にして、「未来の感覚入力」が「明るい」になるという推定をしたうえで、別の行動(スイッチオフ)ではなくスイッチオンが選ばれる。つまり「行動選択は未来の外界の状態の推定」に基づいている。

さきほどの知覚の話のときには「知覚とは現在の外界の状態の推定」であると書いた。両者を合わせると、知覚も行動選択も「外界の状態の推定」をいかにうまく行うかが知覚の正確さ、行動選択の正しさを決める。このようにして知覚と行動選択とをまとめて捉えることができる、これが「自由エネルギー原理が知覚と行動選択を統一的に説明できる」ということの内実だ。

ここまで(情動的)自由エネルギーがなにかの説明はしてこなかったが、(情動的)自由エネルギーとは「外界の状態の推定」をするときにagentが(非明示的に)使っている指標だ。Agentが(情動的)自由エネルギーを減らすように(脳や身体といった)内部状態を変化させるとき、知覚は現在の外界の状態を正確に推定できるようになり、行動選択は未来の外界の状態を正確に推定するように選ばれる。

いったんまとめる。

- 知覚: 現在の外界の状態の推定
- 行動選択: 未来の外界の状態の推定(の帰結)

## [0-4. 学習=外界の状態を推定するモデルの更新]

冒頭にFEPとは「知覚と行動と学習の統一原理」だと書いた。では「学習」はどこに関わってくるか？

さきほどの例を用いれば、いまagentは「感覚入力」が「暗い」を観測していて、「外界の状態の推定」(=知覚)として「照明オフ」の確率90%としている。これを確かめるために、「行動選択」を「スイッチオン」にすれば、「感覚入力」が「明るい」になるだろうという予測を立てて行動選択をする。この予測が正しければ、推定の根拠となったagentの「生成モデル」は正しいのでそのまま維持すればよい。しかしこの予測が裏切られたとき、つまり「スイッチオン」にしたのに「感覚入力」が「暗い」を観測した。このときが「学習」の出番だ。

予想外のことが起きたときのまず最初の対処法は、繰り返しスイッチオン、オフを繰り返して行動選択から予測のサイクルを回す方法だろう。しかしもしこの予想外が続くのであれば、「生成モデル」が間違っている、現在の状況に合わなくなったということなので、生成モデルをアップデートしなければならない。これが学習だ。

発達や老化も同じように捉えることができる。Agentは発達により明るさセンサーの特性が変わると、どういう状況でも「明るい」を観測するようになるかもしれない。このようにして発達においても生成モデルの改変が必要になる。

同じことは進化にもあてはまるだろう。気球規模の変動で新たな環境に対応しなければならなくなったagentは、新たな環境(火山噴火によって照明は常に暗く観測されるかもしれない)に合わせた新しいセンサー特性へのアップデートが必要になるだろう。

このようにしてFEPは「知覚と行動と学習」について「生成モデルを元に外界の状態の推定する」という単一の枠組みで統一的に説明することができる(と主張している)。以上をまとめるとこうなる：

- 知覚: 現在の外界の状態の推定
- 行動選択: 未来の外界の状態の推定(の帰結)
- 学習: 生成モデルのアップデート

## [0-6. まとめ、以降の方針]

これでFEPとはなにか、ということについて数式を用いない範囲で言えることをだいたいうことができた。FEPのような原理が本当にあるのかはわからないけど、まずはこの理論について知ってみよう、そのうえで、知覚と行動選択と学習とを統一的に説明できる理論というものがあろうか考えてみたい、これが私のFEPに対する態度だ。

よって以下の説明でも、FEPのような原理はあろうのかという観点から、実際の神経科学的データを説明するためにFEPを使うテクニク的なところには入り込まないようにして、なるだけFEPの本質的なところだけ抜き出して理解することに注力するという方針を取る。

(ところでこのFEPという考えはずいぶんキャラが立ってる。なんせ、知覚も行動も世界のことを知るためにあり、行動することで世界のことを理解できる、というのだから、これは学者的な世界観ではないだろうか？われわれは世界のこと全部わかってなくても不安ではないし、わからないなりにもなんとか生きてるし、わかったからってなんともならんことが多いよなと思うわけで。なんてことはどうでもいい。FEP人生論はここでストップ。)

---

## Sec.1 変分自由エネルギー VFE の定義

このセクションでは、変分自由エネルギー VFE について、教科書的な定義と基礎についてまとめてみた。ちゃんとした説明のためには「パターン認識と機械学習(PRML)」の10章とかの機械学習の教科書を読んでみてほしい。

だいたい知っている人のためにまとめておくところとおり：以降のセクションを読むために理解してほしいことはたった二つだ。(a) 変分推定を使うためには、問題となっている状



況の生成モデルがどういう構造になっているかを因果グラフ、因子グラフを用いてきっちり決めてやる必要がある。(b) そのうえで、変分自由エネルギー  $F$  を定義するためには、なにが観測データで、なにが潜在変数を理解して、あとはそれは式(5)に代入すれば一意に決まる。

## [1-1. ベイズ推定とは]

そもそも変分自由エネルギー variational free energy VFE  $F$  とはなにかかというと、変分推定を行うときに使われる値だ。変分推定というのはある種間接的な推定法なので、その前にもっと直接的な推定(ベイズ推定)について説明する。ベイズ推定というのは確率論的な推定法だ。

まず問題設定はこうだ。観測データ  $Y (= \{y_1, y_2\})$  ( $y_1$ : 窓ガラスが濡れている or  $y_2$ : 窓ガラスが濡れていない) と 潜在変数  $Z (= \{z_1, z_2\})$  ( $z_1$ : 外は雨が降っている or  $z_2$ : 外は雨が降っていない) がある。いま窓ガラスが濡れている、を観測した( $Y = y_1$ )。このとき外が雨が降っている確率  $Prob(Z = z_1)$  を推定したい。これだけだと問題の解きようがないから、潜在変数  $Z$  (雨が降っているかどうか) という原因が観測データ  $Y$  (窓ガラスが濡れているかどうか) という結果を引き起こすときの関係を、これまでの経験から、両者の同時確率  $Prob(Z, Y)$  として持っている。この同時確率のことを生成モデル  $p(Z, Y)$  と表記する。

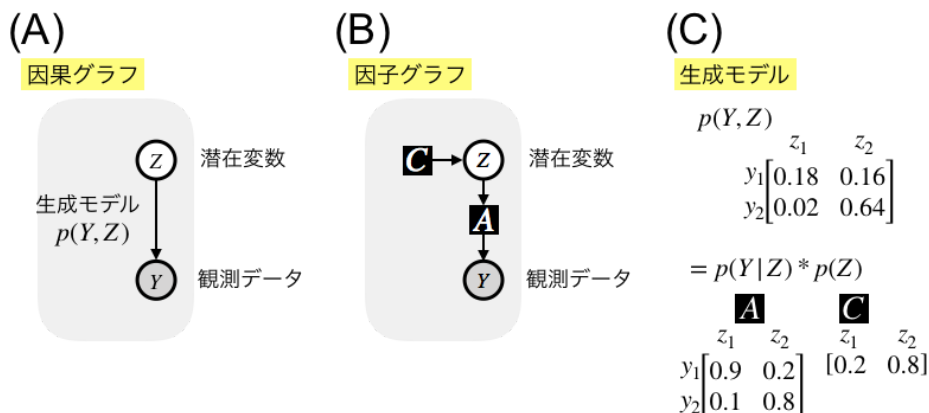


図1-1: 窓ガラスYから降雨Zを推測する

このときの潜在変数  $Z$  と観測データ  $Y$  の因果関係を有向グラフで表したのが図1-1Aの因果グラフ。両者をつなぐ関係が生成モデル  $p(Z, Y)$  となっている。生成モデルは同時確率なので、具体的な例としては図1-1Cのような2x2の行列で表すことができる。同時確率なので、4つのセルの確率を全部足すと1になる。

しかし図1-1Aでは生成モデル  $p(Z, Y)$  と  $Z$  および  $Y$  に対する関係がよくわからないので、それを明示したのが図1-1Bの因子グラフ。このような表現にすると、生成モデル  $p(Z, Y)$  は  $Z$  の事前分布である  $C = p(Z)$ 、それから  $Z$  から  $Y$  を生成する観測モデル  $A = p(Y|Z)$  の二つに分解(因子化)できることができる。こうすると生成モデルの意味も理解しやすくなる(図1-1C)。事前分布  $C$  からわかるように、そもそも雨が降る確率は低い

$(p(z_1) = 0.2)$ 。また、観測モデル  $A$  からわかるように、雨が降ってなくても  $(z_2)$ 、窓ガラスが濡れている  $(y_1)$  という可能性はある  $(p(y_1|z_2) = 0.2)$ 。

いま知りたいのはある  $Y = y_1$  のときの  $Z$  の確率分布  $p(Z|y_1)$  だから、ベイズの法則が使える。

$$p(Z|y_1) = \frac{p(Z, y_1)}{\sum_Z p(Z, y_1)} \quad (1-1)$$

このようにして、生成モデル  $p(Z, Y)$  から事後分布(以下true posteriorと呼ぶ)  $p(Z|y_1)$  を計算するのがベイズ推定だった。

(なお。以下のすべての説明で、総和の記号  $\sum$  のみで説明できる状況の話だけをする。積分記号  $\int$  は出てこない。つか積分記号嫌い。)

## [1-2. 変分推定とは]

しかし、式(1)は分母で全ての可能な潜在変数  $Z$  で足し算をするという作業が入っている。そんな事できないときもあるし、現実的な場面での応用では生成モデル  $p(Z, Y)$  は変数が多くて計算量的に難しいという事情もある。そこで使われる近似的方法が変分推定だ。

True posterior  $p(Z|y_1)$  を直接計算する代わりに、それを近似する確率分布として推測  $q(Z)$  というものを設定する。推測  $q(Z)$  の分布の形を変えて、true posterior  $p(Z|y_1)$  に一致させることができれば、true posterior  $p(Z|y_1)$  を計算できたのと同じことだ。

(この文書では  $q$  のことは一貫して「推測」と呼ぶことにする。より正確な表現ではapproximate posteriorと呼ぶべきだろう。この呼び方だと、true posteriorと同じ変数が入っているということがわかりやすい。True posteriorの変数は潜在変数  $Z$  だから、推測  $q =$  approximate posteriorの変数も潜在変数  $Z$  だ。どちらも全ての  $Z$  で和を取れば1になる。)

(それでもここで  $q$  を「推測」と呼ぶのは、今回の説明では近似法(平均場近似やラプラス近似)を使わないexactな計算だけに話を絞るので、 $q$  を近似によって得られたものと混同しないようにしたいという意図がある。機械学習での変分ベイズとは近似推定であり、exactな計算ができるならこんな回り道をする必要はない。しかし自由エネルギー原理においての変分推定の本質とは、 $q$  という汎関数functionalを操作することにある、というのが私の理解。なのにFEP論文の説明のほとんどは平均場近似やラプラス近似をしたあとのテクニックに終始していて、FEPの正確な理解が阻まれている、というのが私の現状認識。)

## [1-3. KL距離とは]

このために、推測  $q(Z)$  とtrue posterior  $p(Z|y_1)$  というふたつの確率分布の近似度を計算する指標としてカルバック・ライブラー距離  $D_{KL}$  (以下KL距離, KLDと略する) というのを使う。

$$D_{KL}[q(Z)||p(Z|y_1)] = \sum_Z q(Z) \ln \frac{q(Z)}{p(Z|y_1)} \quad (1-2)$$

$y_1$  はすでに観察されて確定している値で、 $Z$  で総和をとっているから、KL距離は  $q(Z)$  の分布の形だけによって決まる定数だ。もし推測  $q(Z)$  と true posterior  $p(Z|y_1)$  が完全一致していたらKL距離は0になる。それ以外は>0となっている。

(距離とはいけど逆向きは同じ距離ではない。つまり、 $D_{KL}[q(Z)||p(Z|y_1)] \neq D_{KL}[p(Z|y_1)||q(Z)]$ 。ではここでなぜ前者を使うのかというと、KL距離と交差エントロピーとの関係から説明できるはずだが、後回しで。)

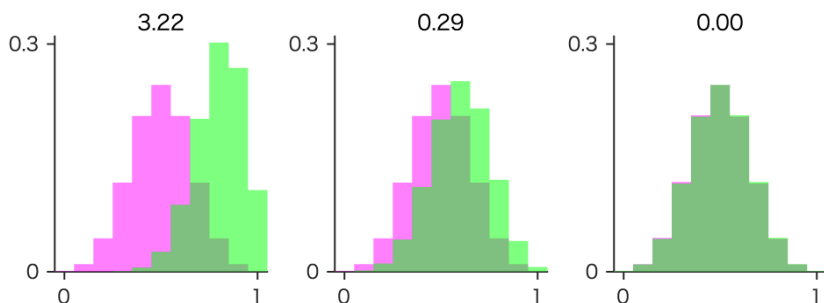


図1-2: KL距離の例

図1-2にKL距離の例として、二つの確率分布(二項分布から作成)のヒストグラムを書いて、そのKL距離(単位はbits)を表示した。離れた分布ほどKL距離は大きく(図1-2左)、二つの確率分布が完全に一致するときはKL距離=0なることがわかる(図1-2右)。

### [1-4. KL距離の最小化]

しかしこのKL距離が直接計算できるならいいのだが、それができるなら直接true posteriorを計算すればいいだけだ。違いはどこに現れてくるかというと、KL距離を式変形すると、式(1-3)のようになる。(式変形で  $p(Y, Z) = p(Z|Y)p(Y)$  を使用。)

$$\underbrace{\ln p(y_1)}_{\text{Log evindecce}} = - \underbrace{\sum_Z q(Z) \ln \frac{q(Z)}{p(Z, y_1)}}_{\text{ELBO}=-F} + \underbrace{D_{KL}[q(Z)||p(Z|y_1)]}_{\text{KLD}} \quad (1-3)$$

左辺は観測データ  $Y$  が生成モデル  $p$  のもとでどのくらいの確率であり得るかという evidence  $p(Y) = \sum_Z p(Y, Z)$  のlogをとったものなので、log evidence (周辺対数尤度)と呼ぶ。

(なお、 $p(Y)$  はただの観測データ  $Y$  の出現確率  $Prob(Y)$  ではないことに注意。 $p(Y)$  はいまある生成モデルに基づいたうえでデータの出現確率だから、尤度likelihoodなのだ。このことは後述の「暗い部屋問題」に関わってくる。)

(このことを明示するためにこの生成モデルに  $m$  という名前をつけて、 $p(Y|m)$  もしくは  $p_m(Y)$  と表示することもある。ここではその表示は採用しない。そのかわりに、生成モデル  $p()$  とただの確率  $Prob()$  を厳密に分けて表示している。)



でもって、このlog evidenceというやつはlogの中身が確率だから、かならず負の値を取る。(実際に観測されたデータがあるのだから、この値は決して  $-\infty$  にはならない。) そして右辺の第2項は  $KLD \geq 0$  だから、第1項はlog evidenceよりも必ず小さい。

$$\underbrace{\ln p(y_1)}_{\text{Log evidence}} \leq - \underbrace{\sum_Z q(Z) \ln \frac{q(Z)}{p(Z, y_1)}}_{\text{ELBO} = -F} \quad (1-4)$$

よって、この第1項はlog evidenceの下限になっている。そこでこの第1項はEvidence lower bound, ELBOと呼ばれる。そして変分自由エネルギー  $F$  とは  $F = -\text{ELBO}$  と単に符号逆転させただけのものだ。(ELBOはlog evidenceよりも小さいのだから、必ず負の値だ。ということは変分自由エネルギー  $F$  はかならず正の値。)

## [1-5. ちょっと脇道]

ちょっと脇道。といつつFEPの理解に関しては、ここからが面白い。以上のことからわかるのは、以下の(1)-(7)が全部ひとつながりの現象であるという点だ。

- (1) ベイズ推定をする
- (2) true posteriorを計算する
- (3) 推測  $q$  をtrue posteriorに一致させる
- (4) KL距離を0にする
- (5) ELBOを最大化する
- (6) 変分自由エネルギー  $F$  を最小化する
- (7) 周辺対数尤度  $\ln p(Y)$  を最大化する

ただし、生物が扱うような複雑な状況においては、これらすべてを完全に実現することはできない。よって、どこかで近似が入ってくることになる。

自由エネルギー原理についての議論で、どれが目的でどれが結果か、どれがなにを近似しているのか、といった問題を理解するためには、この関係について考えればいい。

たとえば暗い部屋問題というのがある。自由エネルギー原理では(6)  $F$  を最小化することが生物の目的だというが、それだったら、(7) 一番ありうる観測データをサンプルすればいいのだから、暗い部屋にじっとしていればいいじゃん、という議論。でもこれは(7)の解釈が間違っている。(7)は生成モデルの元での尤度なのだから、暗い部屋という観測データの出現確率が高い生成モデルをあらかじめ獲得していない限り、これは成り立たない。(だから、冬眠する動物は暗い部屋を選ぶ。我々だって夜は暗い部屋を選ぶ。)

それよりは、ネット言論がタコソバ化することについて当てはめる方がまだ尤もらしいだろう。つまり、我々は自分の持論(=世界についての生成モデル)を持っていて、それをより補強するデータばかり観測しようとする。だからネット言論ではみな自分が聞きたい意見ばかり集めるようになって、分断はより強調されていく。これは自由エネルギー原理から説明できる。なんだってー！

あと、ベイズ脳と自由エネルギー原理の関係についてもどちらが原因でどちらが結果かは自明でない。

- 議論A: 「(1) われわれ生物はベイズ的に情報処理をしているんだ」というのが先にあって、そのための近似法として「(6) 自由エネルギー最小化 をするように脳と身体を進化せさせてきた」と主張。
- 議論B: なんらかの生物学的拘束条件から「(6) われわれ生物は自由エネルギー最小化をするようになっていく」が先にあって、その結果として「(1) 行動や知覚でベイズ推定をしているように見えている」と主張。

どちらも議論としてはありえる。Friston自身は(6)を自由エネルギー「原理」と呼んだうえで、(1)をベイズ脳「仮説」と呼んでいるので、議論Bに基づいていると考えるのが筋が通っていると私は思うのだが、Fristonの発言自体はそのつど言っていることがブレているように思う。

## [1-6. 変分自由エネルギーとは]

そういうわけで、変分自由エネルギー variational free energy VFE  $F$  は式(1-5)のとおり定義される。

$$\begin{aligned} F &= \mathbb{E}_{q(Z)} [\ln q(Z) - \ln p(Y, Z)] \\ &= \underbrace{D_{KL}[q(Z) \| p(Z|Y)]}_{\text{KLD}} + \underbrace{-\ln p(Y)}_{\text{Surprisal}} \end{aligned} \quad (1-5)$$

(ここまでは観測データには具体例での  $y_1$  を使ってきた。しかしもちろんこの式は  $y_2$  でも成り立つわけで、ここからは観測データは  $Y (= \{y_1, y_2\})$  を使って表示する。しかしこれまでのことからわかるように、 $Y$  はすでに確定した一つのデータであり、一方で  $Z$  はとりうる値が全部並んだ変数になっている。同じ変数みたくに見えてべつものなのだ。慣れた人にとっては当たり前だろうけど、私はこれが馴染むまで相当時間がかかった。この違いを明示するために  $p(Z|y)$  や  $p(Z|Y = y)$  のように表示するのが正確だと思うが、式が長くグロくなるので、短く書ける場所以外ではその表記はしてない。)

(ここでは期待値の記号として  $\mathbb{E}$  を使っている。下付き文字の期待値で重み付けして、その期待値の変数  $Z$  で和をとる。たとえば  $\mathbb{E}_{q(Z)} p(Z) = \sum_Z q(Z)p(Z)$  というふうに。)

ここで第2項は(sensory) surprisalと呼んでいる。この値はlog evidenceの符号逆転で、必ず正の値になっている。得られた観測データが現在の生成モデルに基づいて予想外であれば、 $p(Y)$  はより小さくなるので、surprisalは大きくなる。Surprisalはシャノンのself informationとも呼ばれる。これをsurpriseと呼ぶこともあるが、次に出てくるBayesian surpriseとの区別のために、この文書ではsurprisalという言葉で統一させてもらう。

(なお、surprisalが必ず正になるのは離散分布を考慮したときだけ。)

## [1-7. Bayesian surpriseとは]

$F$  は式変形によって式(6)の形でも表現できる。(  $p(Y, Z) = p(Y|Z)p(Z)$  を使用。)

$$F = \underbrace{D_{KL}[q(Z) \| p(Z)]}_{\text{Bayesian surprise}} + \underbrace{-\mathbb{E}_{q(Z)} \ln p(Y|Z)}_{\text{Uncertainty}} \quad (1-6)$$

第1項のBayesian surpriseは、現在の推測  $q$  と生成モデルにおける事前分布  $p$  との間の距離なので、推測  $q$  が観測データによってアップデートされると、より大きくなる。つまり、情報獲得の大きさと捉えることができる。

第2項の *Uncertainty* は、Fristonの表記では  $-Accuracy$  となるが、*Accuracy* は必ず負なので、ここではわかりやすさ優先で、必ず正となる *Uncertainty* という表記を採

用している。この項については後回しで。(なお、ここでのuncertaintyは推定  $q$  の分散とは別ものであることに注意。現在の例では推定  $q$  は点推定をしているので、分散ではないことはわかると思う。今回の文書では推定の分散があるような例は扱わない。)

第1項の意味は、次の近似(いまこの文書で初めて近似を使った!)を考えるとさらに具体的になる。もし推測  $q(Z)$  を true posterior  $p(Z|Y)$  に完全に一致させることができた場合には、式(7)の2行目が成り立つ。

$$\begin{aligned}
 F &= \underbrace{D_{KL}[q(Z)||p(Z)]}_{\text{Bayesian surprise}} + \underbrace{-\mathbb{E}_{q(Z)} \ln p(Y|Z)}_{\text{Uncertainty}} \\
 &\approx \underbrace{D_{KL}[p(Z|Y)||p(Z)]}_{\text{BS}'} + \underbrace{-\mathbb{E}_{q(Z)} \ln p(Y|Z)}_{\text{Uncertainty}} \quad (1-7)
 \end{aligned}$$

ここの第1項BS'は、式(8)にあるように観測データの期待値  $p(Y)$  をかけてやると、観測データ  $Y$  と潜在変数  $Z$  の間の相互情報量MIになっている。このことから、Bayesian surpriseが情報獲得に関わっているということがわかるかと思う。

$$\begin{aligned}
 \mathbb{E}_{p(Y)} \text{BS}' &= \mathbb{E}_{p(Y)} D_{KL}[p(Z|Y)||p(Z)] \\
 &= MI(p(Z|Y), p(Z)) \quad (1-8)
 \end{aligned}$$

なお式(8)は後ほど出てくる期待自由エネルギーEFE  $G$  でのepistemic valueと同じ形になっている(ただしepistemic valueでは  $p$  ではなくて  $q$  だが)。これについては再訪する。

## [1-8. 実例で変分自由エネルギーの最小化を試みる]

それでは、図1-1に出した例を使って、推測  $q$  を変化させながら変分自由エネルギー  $F$  の最小化をシミュレーションしてみよう。まず、先に正解をカンニングしてしまうと、図1-1の生成モデル  $p$  から true posterior を直接計算して、

$$p(Z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} | Y = [y_1 \quad y_2]) = \begin{bmatrix} 0.529 & 0.030 \\ 0.470 & 0.969 \end{bmatrix} \quad (1-9)$$

となる。よって、 $Y = y_1$  (窓ガラスが濡れている)のとき、 $Z = z_1$  (雨が降っている)の確率は  $p(Z = z_1 | Y = y_1) = 0.529$  とわかった。事前分布  $p(Z = z_1) = 0.2$  であったことを考えると、窓ガラスが濡れている、という観察はinformativeであったことがわかる。

では  $q(Z)$  を変化させてみよう。ここで  $q(Z)$  の形を変えと言ったが、実のところ、 $q(Z = z_1)$  が決まれば、 $q(Z = z_2) = 1 - q(Z = z_1)$  も決まる。そこでパラメーター  $\phi = q(Z = z_1)$  を  $0 - 1$  の範囲で動かして  $VFE, KLD, Surprisal$  を計算してプロットしてやる。すると図1-3Aのとおりになった。

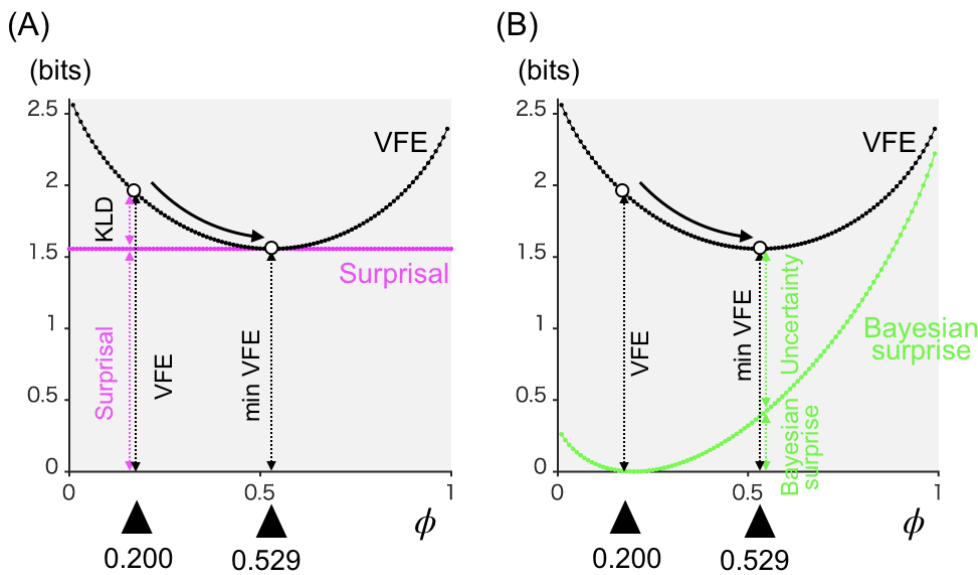


図1-3: 変分自由エネルギーVFEの最小化

$q(Z)$  の初期状態は事前分布  $p(Z) = \begin{bmatrix} 0.2 \\ 0.8 \end{bmatrix}$  と等しいと考えてやれば、 $\phi = 0.2$  からスタートするのが妥当だろう。このとき、

$$VFE = 1.8879$$

$$KLD = 0.3315$$

$$Surprisal = 1.5564 \text{ (bits)}$$

となっている。ここからVFEが小さくなる方に $\phi$ を動かしていくと、 $\phi = 0.529$ で

$$VFE = 1.5564$$

$$KLD = 0$$

$$Surprisal = 1.5564 \text{ (bits)}$$

となる。 $KLD = 0$ でVFEが最小化された状態では、 $VFE = Surprisal$ となっている。この $\phi = 0.529$ は直接true posteriorを計算した結果の $p(Z = z_1 | Y = y_1) = 0.529$ と同じになっていることが確認できた。

そういうわけで、変分自由エネルギーを最小化することで、true posteriorを近似できるということが確認できた。自由エネルギー原理では、脳はこのようなやり方で観察データ(感覚入力)から潜在変数(外界の状態)を推定していると考えられる。以降のセクションでは、より知覚や行動に似せた状況での生成モデルを構築して、どのようにVFEが計算されるかを見てゆく。

なお、図1-3から想像がつくように、もしVFEの曲線がlocal minimumを持っていたら、そこで停まってしまって、正しいtrue posteriorが計算できないであろうことも想像がつく。そして、もっと複雑な生成モデルではそのようなことはいかにも起こりそうだ。

## [1-9. 変分自由エネルギーの最小化をBayesian surpriseからしてみる]

図1-3Aでは式(5)  $F = KLD + Surprisal$  に基づいて、 $\phi$  を0.2から0.529に動かして、そのときに  $KLD = 0$  となり、 $F$  が最小化されるのを見た。では同じものを式(6)  $F = Bayesian\ surprise + Uncertainty$  に基づいて表示してみよう(図1-3B)。

Bayesian surpriseは  $q(Z)$  の初期状態  $\phi = 0.2$  では0になっていることがわかる。これは  $q(Z)$  の初期状態 =  $Z$  の事前分布  $q(Z) = p(Z)$  を採用したのが理由だ。まだこの段階では観察データ  $y_1$  から獲得した情報がゼロということ。

$$VFE = 1.8879$$

$$Bayesian\ surprise = 0$$

$$Uncertainty = 1.8879(bits)$$

$\phi$  は正負どちらに動かそうが、Bayesian surpriseは増える。現状よりもちがうというだけで情報獲得なので。ということで、Bayesian surpriseだけ見ても、VFEを下げることはできないということがわかる。

そして、 $\phi$  を0.2から0.529に動かしてVFE を最小化したとき、Bayesian surpriseが増えていることが確認できる。

$$VFE = 1.5564$$

$$Bayesian\ surprise = 0.3845$$

$$Uncertainty = 1.1719(bits)$$

図1-3Bのプロットを見ると、 $\phi = 0.529$  はBayesian surpriseを最大化する場所ではないこともわかる。

## [1-10. 変分自由エネルギーを微分して解析的に解いてみる]

最後に変分自由エネルギーVFE  $F$  を  $\phi$  で微分して、自由エネルギーの最小化について解析的に解いてみることにしよう。ここで必要な数学的知識は高校数学IIIまででいける。 $f'(z)$  は  $z$  の関数  $f(z)$  を  $z$  で微分したものを意味するとしただけで、 $(f(z) * g(z))' = f'(z) * g(z) + f(z) * g'(z)$  と  $(\ln z)' = \frac{1}{z}$  の二つだけ知っていれば解ける。

式(5)に  $Y = y_1, q(Z = z_1) = \phi, q(Z = z_2) = 1 - \phi$  を代入する。生成モデルから  $p(Y = y_1, Z = z_1) = 0.18, p(Y = y_1, Z = z_2) = 0.16$  も入れておく。



$$\begin{aligned}
F &= \mathbb{E}_{q(Z)}[\ln q(Z) - \ln p(Y, Z)] \\
&= \sum_Z q(Z) * [\ln q(Z) - \ln p(Y, Z)] \\
&= q(Z = z_1) * [\ln q(Z = z_1) - \ln p(Y = y_1, Z = z_1)] \\
&\quad + q(Z = z_2) * [\ln q(Z = z_2) - \ln p(Y = y_1, Z = z_2)] \\
&= \phi * [\ln \phi - \ln 0.18] + (1 - \phi) * [\ln(1 - \phi) - \ln 0.16] \\
&= \phi \ln \phi - \phi \ln 0.18 + \ln(1 - \phi) - \phi \ln(1 - \phi) \\
&\quad - \ln 0.16 + \phi \ln 0.16 \\
&= \phi \ln \phi + \ln(1 - \phi) - \phi \ln(1 - \phi) \\
&\quad + \phi \ln \frac{0.16}{0.18} - \ln 0.16 \\
&= \phi(\ln \phi + \ln \frac{0.16}{0.18}) + (1 - \phi) \ln(1 - \phi) - \ln 0.16 \tag{1-10}
\end{aligned}$$

これを  $\phi$  で微分する。

$$\begin{aligned}
F' &= [\phi(\ln \phi + \ln \frac{0.16}{0.18}) + (1 - \phi) \ln(1 - \phi) - \ln 0.16]' \\
&= \phi'(\ln \phi + \ln \frac{0.16}{0.18}) + \phi(\ln \phi + \ln \frac{0.16}{0.18})' \\
&\quad + (1 - \phi)' \ln(1 - \phi) + (1 - \phi) \ln(1 - \phi)' \\
&= \ln \phi + \ln \frac{0.16}{0.18} + \phi(\frac{1}{\phi}) \\
&\quad - \ln(1 - \phi) - (1 - \phi) \frac{1}{1 - \phi} \\
&= \ln \phi + \ln \frac{0.16}{0.18} - \ln(1 - \phi) \\
&= \ln \frac{0.16}{0.18} - \ln \frac{1 - \phi}{\phi} \tag{1-11}
\end{aligned}$$

$F' = 0$  と置くと  $\phi$  が求まった。

$$\begin{aligned}
0 &= \ln \frac{0.16}{0.18} - \ln \frac{1 - \phi}{\phi} \\
\ln \frac{1 - \phi}{\phi} &= \ln \frac{0.16}{0.18} \\
\frac{1 - \phi}{\phi} &= \frac{0.16}{0.18} \\
\frac{1}{\phi} &= 1 + \frac{0.16}{0.18} \\
\phi &= \frac{0.18}{0.18 + 0.16} \\
\phi &= 0.529 \tag{1-12}
\end{aligned}$$

たしかに、 $\phi = 0.529$  で  $F$  が最小化するのを確認できた。

これでなにがわかったかというと、われわれ生物が  $F$  を下げてゆくためには、

$$F \leftarrow F + \alpha * F' \quad (1-13)$$

という更新ルール( $\alpha$  は更新のスピードを決めるパラメーター)と  $F'$  を持っていればいい。そうすると  $F'$  が計算可能でないといけないわけだけど、確認のため、式(11)に元に記号を入れて戻してやると、

$$\begin{aligned} F' &= \ln \frac{0.16}{0.18} - \ln \frac{1-\phi}{\phi} \\ &= \ln \frac{p(Y = y_1, Z = z_2)}{p(Y = y_1, Z = z_1)} - \ln \frac{q(Z = z_2)}{q(Z = z_1)} \\ &= \ln \frac{p(Y = y_1 | Z = z_2)p(Z = z_2)}{p(Y = y_1 | Z = z_1)p(Z = z_1)} - \ln \frac{q(Z = z_2)}{q(Z = z_1)} \end{aligned} \quad (1-14)$$

となるので、第1項は生成モデルの  $A$  と  $C$  (図1-1C参照)さえあれば計算できる。True posteriorを計算する必要はない。そしてこの第1項は、二つの対立仮説( $Z = z_1$  vs.  $Z = z_2$ )での観察データ  $y_1$  についての対数尤度比になっている。これと第2項の現在の推測  $q$  による対数尤度比との差が  $F'$  の正体だった。

われわれ生物は(すくなくとも)このような単純な環境では、尤度比検定をすることによって認識・推論の過程を最適化している、というのが自由エネルギー原理の世界観だということになる。

## [1-11. このセクションで言いたかったこと]

長々と書いてきたが、以降のセクションを読むために理解していほしいことはたった二つだ。(a) 変分推定を使うためには、問題となっている状況の生成モデルがどういう構造になっているかを因果グラフ、因子グラフを用いてきっちり決めてやる必要がある。(b) そのうえで、変分自由エネルギー  $F$  を定義するためには、なにが観測データで、なにが潜在変数を理解して、あとはそれは式(5)に代入すれば一意に決まる。

ではこのセクションのおさらいを兼ねて、新しい生成モデルでこの二つの作業をやってみよう。図1-3Aではさきほどの図1-1の状況にひとつだけ要素が加わっている。図1-1Aでは、観測データ  $Y$  「窓ガラスが濡れているか」、潜在変数  $Z$  「雨が降っているか」の二つの要素だけがあった。図1-3Aではこれに加えて、潜在変数  $X$  として「スプリンクラーが作動しているかどうか」がある。この図ではスプリンクラーの動作は雨が降るかどうかとは無関係にしてある。(もし、雨が降ったときはスプリンクラーは動かさない、という因果関係を設定した場合には、 $Z$  から  $X$  への矢印も必要になる。)

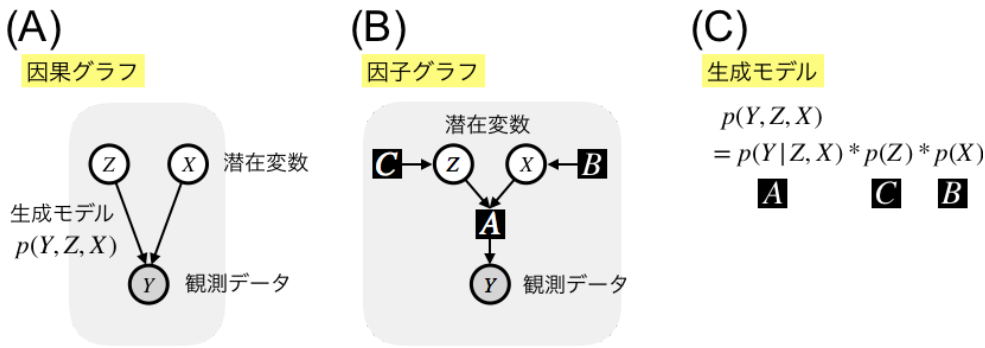


図1-4: 窓ガラスYから降雨Zとスプリンクラーの動作Yを推測する

このような関係を因子グラフにすると図1-3Bになる。観察モデルAは二つの入力を受けて一つの出力を出すように変わった。ゆえに、雨が降っているかないか  $\{z_1, z_2\}$  とスプリンクラーが動作しているかないか  $\{x_1, x_2\}$  の4通りの組み合わせ  $\{(z_1, x_1), (z_1, x_2), (z_2, x_1), (z_2, x_2)\}$  について、窓ガラスが濡れているかどうか  $\{y_1, y_2\}$  の確率を知っている、これが観察モデルAだ(図1-3C)。

するとこのときのVFE  $F$  を計算するためには、観察データ  $Y$  と潜在変数  $Z, X$  を式(5)の中に入れてやればいい。

$$\begin{aligned}
 F &= \mathbb{E}_{q(Z, X)} [\ln q(Z, X) - \ln p(Y, Z, X)] \\
 &= \underbrace{D_{KL}[q(Z, X) \| p(Z, X|Y)]}_{\text{KLD}} + \underbrace{-\ln p(Y)}_{\text{Surprisal}} \\
 &= \underbrace{D_{KL}[q(Z, X) \| p(Z, X)]}_{\text{Bayesian surprise}} + \underbrace{-\mathbb{E}_{q(Z, X)} \ln p(Y|Z, X)}_{\text{Uncertainty}} \quad (1-15)
 \end{aligned}$$

その結果が式(1-15)だ。つまり、式(1-5)で  $Z$  だったところが、 $Z, X$  になっているだけ。つまり、複数の変数は同時確率として扱われるので、カンマつけて並べておけばいい。

## Sec.2 自由エネルギー原理の前提、世界設定

### [2-1. 前置き]

さてこれからがいよいよ本番、自由エネルギー原理(Free energy principle, FEP)についての説明を行ってゆく。FEPは2005年のPhilos Trans R Soc Lond B Biol Sci論文から始まって、現在まで理論的にも進歩を続けている。このうちの最新バージョンがFriston et. al. Cogn Neurosci. 2015で提案されるようになったものだ。

この最新バージョンでは、自由エネルギー原理FEPはこれまでに用いられてきた連続的な時間での表現とは別に、部分観測マルコフ決定過程POMDPにもとづいた離散的な時間で表現

されるようになった。これによってFEPによれば、われわれagentは「現在の外界の状態を推測する際の(情報理論的)変分自由エネルギー $F$  (variational free energy VFE)を最小化するように知覚する」だけでなく「未来の外界の状態を推測する際の(情報理論的)期待自由エネルギー $G$  (expected free energy EFE)を最小化する行動を選択する」というように、我々の知覚と行動が定式化されるようになった。

この文書の以下のパートでは、この期待自由エネルギー $G$ をどのように導いているかについて、以下の論文を参考にまとめてみた。Fristonの論文では省略されているような、式の導出における前提について、なるだけ詳しくまとめてある。これだけ詳しいものはないのではないかという自負はあるが、それゆえに独自解釈になっている可能性もある。いったんこの文書を理解したならば次は原著論文の方に当たっていただきたい。

## References

- Friston et. al. [Cogn Neurosci. 2015](#) EFEの初出。
- Friston et. al. [Neural Comput. 2017](#) EFEについてのスタンダードな記述。
- Schwöbel et. al. [Neural Comput. 2018](#) Predicted free energyの概念の提案。
- Parr and Friston [bioRxiv 2018](#) Generalized free energyの概念の提案。
- Gershman [arXiv 2019](#) VFEとEFEについての的確な整理。
- Katahira et. al. [PsyArXiv 2019](#) 単純な例によるEFEの計算とretrospective surpriseの概念の提案。

## [2-2. 自由エネルギー原理における世界設定]

この説明では、単純化した世界設定を用いるが、式変形においてはなるだけ近似は使わない(近似の導入をなるだけ遅らせる)という方針で行く。記号は私の以前のスライドにあるものを踏襲する。(マルコフ意思決定プロセスで使われるobservation  $o$  とstate  $s$  や機械学習で使われる observed variable  $x$  とhidden variable  $z$  ではないので、置き換えて読んでほしい。)

本説明ではどのように  $F$  および  $G$  を計算するかだけを取り扱う。じっさいに生物が  $F$  および  $G$  を下げてゆく過程(微分方程式が必要な部分)についてはsec.1の最後で少々言及したが、以降は扱わない。

まず基本的な設定から説明しよう。

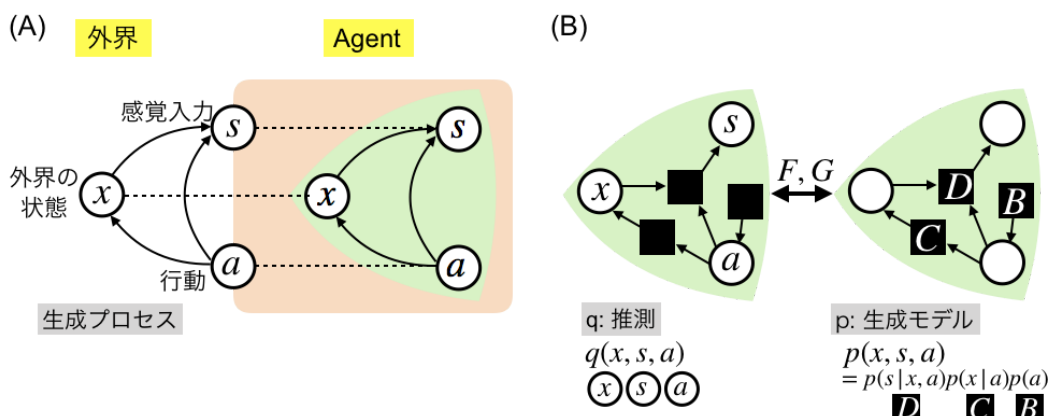


図2-1: 基本的な設定

われわれagentは外界の中に存在していて、感覚入力 $s$ と行動 $a$ を通して外界とつながっている(図2-1A)。現在観察している感覚入力 $s$ は世界の状態 $x$ から因果的に引き起こされる。Agentは世界の状態 $x$ には直接アクセスすることができない。Agentは行動 $a$ によってこの $s$ を変えることができるし、 $x$ を変えて、その変化を $s$ への変化として観測することもできる。このようにして、外界のうち、agentとの相互作用に関わっている部分が「生成プロセス」だ。

### [2-3. 生成モデル $p$ と推測 $q$ ]

Agentはこの生成プロセスを持っているのではなくて、それを経験から学習した「生成モデル」 $p(x, s, a)$ を持っている。生成モデルは外界にある生成プロセスと同様に $x, s, a$ の間での因果関係のグラフ構造として表現される。このグラフ構造は

$$p(x, s, a) = p(s|x, a)p(x|a)p(a) \quad (2-1)$$

と条件付き確率の掛け算として書くことができる。図2-1Aにある矢印の関係ををより明確にするために、図2-1Bでは因子グラフ(factor graph)を使って黒四角で表示している。つまり生成モデルの本体とはこのような変換器 $B, C, D$ であり、中に入っている変数 $x, s, a$ はまだ特定されていない。

(このことを明示するためには $p(X, S, A) = p(S|X, A)p(X|A)p(A)$ と大文字 $X = \{x_1, x_2, \dots\}$ で表記して、特定された値と区別することもできる。ここでは式(1)の表記を使う。)

( $p()$ は生成モデルのことを指し、一般的な確率 $Prob()$ はべつものとして表記する。)

さらにagentは、時々刻々変化する $x, s, a$ についてその時その時の推測 $q(x, s, a)$ を同時確率として持っている。つまり点推定ではなく、不確定性込みで確率的な表現をしている。推測 $q$ はただの同時確率であり、生成モデル $p$ のような因果的グラフ構造は持っていない。

$$q(x, s, a) \quad (2-2)$$

この表現は外界にある $x, s, a$ とは別にagentが $x, s, a$ についての推測を持っているということを表現している。実際にはSec.1でも書いたように、推測 $q$ とはtrue posteriorを推定するapproximate posteriorであるので、 $q$ の中身は潜在変数が入る。よってじっさいの $q$ の変数は、生成モデルのうちなにが潜在変数であり、なにが観測データであるかによって決まる。

(POMDPなどの強化学習の文脈では、 $q$ を信念beliefと呼ぶけれど、ここでは知覚も含めるので「推定」と呼ぶことにする。)



(話を少々先取りしておくとして、のちほど3-2-1で示すように、agentは外界の状態  $x$  についての推測をアップデートしてゆくが、それは実際には  $q(x)$  という確率分布の形を時々刻々アップデートすることであり、agentの内部状態を表すパラメータ  $\phi$  によって、 $q(x|\phi)$  と表す。また、4-4-2'に詳しく記したが、agentが  $s$  についての推定を持つとは、 $q(x, s) = q(s|x)q(x) = p(s|x)q(x)$  というように、生成モデルのうち、観察モデル  $p(s|x)$  を使って  $x$  についての推測  $q(x)$  からの計算をする、というのが本当のところ。よって正確に言うならば「agentは  $q(x, s, a)$  を推測している」というよりは「agentは  $q(x)$  を推測している」もしくは「内部状態として  $x$  についての推測のみを持っている」というほうが正しい。この点について将来的には、予測符号化のニューラルネットワークの単純なモデルとその説明のサブセクションを作成して、そこで詳しく説明する予定。)

Agentはイマココの潜在変数( $x, s, a$  のどれか)を推測するように その同時分布(=推測  $q$ ) を時々刻々アップデートさせてゆく。

このようにして、脳の中で行っていること(図2-1A)を生成モデル  $p$  と推測  $q$  とに分けたうえで(図2-1B)、その間で定義される変分自由エネルギー  $F$  と期待自由エネルギー  $G$  を最小化するように  $q(x, s, a)$  の分布の形をアップデートさせることが知覚( $x$  の推定)であり、行動選択( $a$  の推定)である、というのが自由エネルギー原理だ。ではどうやって  $F, G$  を定義するか。これがこの文書のメインピックなので、これから順を追って説明してゆく。

なお、推測  $q$  だけでなく生成モデル  $p$  を変えることによって自由エネルギーを下げることはできるが、ここでは考慮しない。つまり生成モデル  $p$  は時間によっては変化しないものとする。本当は学習などによって変化しうるけど、話を簡単にしておきたいので。

## [2-4. 観測データ、潜在変数、パラメーターの区別]

変分自由エネルギーVFEを計算する際には、観測データ、潜在変数、パラメーターを区別することが重要になる。観測データ、潜在変数についてはすでにsec.1で言及したので、パラメーターについて説明を加えておく。

以下の図2-2Aに、Sec.1の図1-1Bで示した因子グラフを再掲した。ここで生成モデルのうち  $Z$  の事前分布を  $C(=p(Z))$ 、観測モデルを  $A(=p(Y|Z))$  と分けて(因子化して)、黒四角で示したのが因子グラフだった。 $C, A$  の実体は行列だった。この行列が生成モデルを特定する「パラメーター」となっている。このパラメーターの部分を外に追い出して白四角で表現したのが図2-2Bだ。

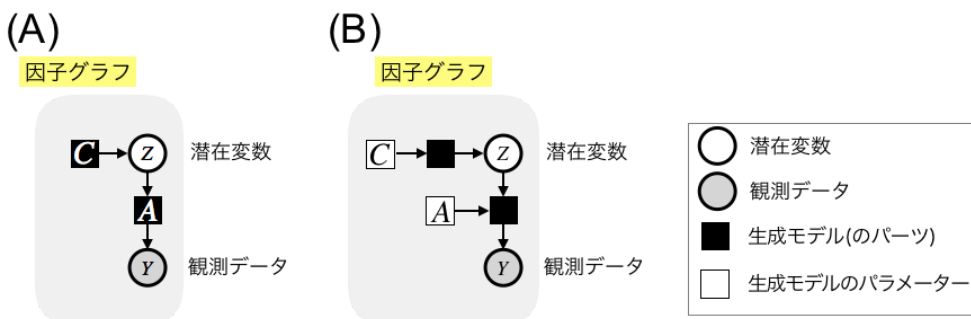


図2-2: 時間を考慮した生成モデル

こうしてやると、生成モデルのパーツである黒四角とは、変数やパラメーターが入力されるまではなんの具体性もない、無時間的な箱であり、変換器であるということがわかる。これらのパラメーターは現在の説明では固定としている。つまり生成モデルは学習済みであり、変化しないものとしている。

しかし、学習などのモデルを作るに当たっては、これらの行列  $A, C$  を変化しうるものとする事も可能だ。この場合、行列  $A, C$  は潜在変数として扱われる。たとえば  $C$  を潜在変数として扱われるようになると、生成モデルは  $p(Y, Z, C)$  となり、可能な範囲の  $C$  をすべて考慮する必要が出てくる。つまり、 $p(Y, Z) = \sum_C p(Y, Z, C)$  のような周辺化ができるようになる。いっぽうで、元々の設定では  $C$  はパラメーターだったわけだが、この場合は可能な範囲の  $C$  などというものはないし、現在設定されている  $C$  に対してただひとつの生成モデルがあるだけだ。このことを明示するためには、生成モデルを  $p_C(Y, Z)$  のように表示する(後で出てくる)。

パラメーターであるか、潜在変数であるかの違いは、のちほど行動  $a$  の扱いのところで重要な役割を果たす。というか、私はこの違いが理解できなくて、期待自由エネルギーの導出に2週間ほどドハマリした。(島崎さんの助けでそこから抜け出すことができました。Thanks!) そういうわけで、読者の皆様にも注意を喚起しておきたい。

脱線、というかある意味こちらが本題なのだが、いま書いたことの含意についてコメントしておきたい。

図2-1Aでは、外界が因果グラフであるかのように表示したが、正確を期すならば、「生成プロセス」がこのような因果グラフであるわけではない。あくまでagentが認識によって外界を切り分けて、因果で繋いだものを「生成モデル」(not生成プロセス)として持っているだけだ。外界はもっと多様であり、その中でagentが生成モデルとして切り分けることができたものを外界に投射し直したものが生成プロセスである、というほうが正しいだろう。人間には人間の、カエルにはカエルの生成モデルがある。

そういう目で見ると、推測  $q$  と生成モデル  $p$  というのは、外界を  $x, s, a$  といった「表象もしくはstate」を切り分けてきたのが推測  $q$  であり、 $x, s, a$  を繋ぐ「変換もしくはprocess」を切り分けてきたのが生成モデル  $p$  である、という整理ができる(図2-1B)。

つまり、推測  $q$  には因果グラフの矢印はないし、生成モデル  $p$  は逆に矢印だけがあって、推測  $q$  が決まるまでは空っぽだ。とくに図2-2Bのように生成モデルのパラメーターも追い出してしまえば、生成モデルの実体とは、どのstateを入力して、どのようにその入力を組み合わせた演算をして、どのstateを出力するか、という矢印の結節点という「関係」だけが残る。生成モデル  $p$  とは、無時間的な空っぽの箱であることがよく分かる。

言い方を変えるならば、生成モデル  $p$  が関数であり、推測  $q$  がそこに代入される変数だと言える。たとえば生成モデルとして決定論的な1次関数  $y = ax + b$  を考えた場合、推測  $q$  とは  $x$  や  $y$  のことだ。そして生成モデルのパラメーターは  $a$  や  $b$  のことだ。

世界をモデル化、形式化する際には、このようなstateとprocessの対が出てくる。この対とは、以前私が書いたブログ記事での、ベイトソン(さらにデビッド・マー)がいう「表象」と「プロセス」そのものであることがわかる。表象だけでも不十分だし、プロセスだけでも不十分で、両者が組み合わさってはじめてこの計算は実現する。

以前私は神経回路学会誌の「自由エネルギー原理と視覚的意識」において、agentが内部に外界を構成する(弱い意味での表象)をする際には、推測  $q$  と生成モデル  $p$  の両方が、リアルタイムで必要だ、という議論をした。いまの書いたこともこの議論を補強するだろう。つまり、推測  $q$  とはイマココのstateについての推測で

あり、生成モデル  $p$  とは(無時間的な)processであるが、両者が組み合わさることによってはじめて、イマココでの知覚が実現するのだ。

以上。それでは話を本題に戻す。

## [2-5. この文章の最終目標: 現在、過去、未来を統一的に捉えた説明]

図2-1に示した生成モデルには時間が入っていなかった。生物の知覚と行動をモデル化するためには時間を考慮した生成モデルが必要だろう。そこで図2-2には現在、過去、未来を入れた最小限の生成モデルを考えてみる。(それでも複雑なので、ここでは因子グラフではなくて、因果グラフで表示してある。)

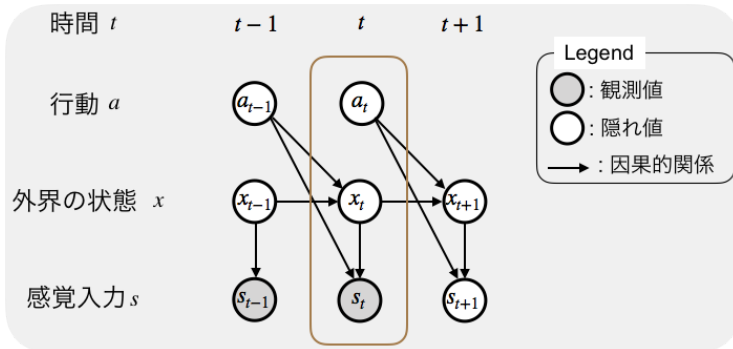


図2-3: 時間を考慮した生成モデル

話を簡単にするために、現在 $t$ 、過去 $t - 1$ 、未来 $t + 1$ の3つの時点だけを考える。部分観測マルコフ決定過程POMDPを仮定しているので、離散的な時間で、直前の状態のみから現在の状態が決まる。そして現在及び過去の感覚入力 $s_t, s_{t-1}$ だけが観察データで(灰色丸で表示)、あとはみな潜在変数(白丸で表示)として推定する必要がある。

現在観察している感覚入力 $s_t$ に基づいてその原因となる外界の状態 $x_t$ は直前の状態 $x_{t-1}$ と以前の行動 $a_{t-1}$ から決まる。つまり自分のstateというものを持っている。それに対して感覚入力 $s_t$ はそのときの外界の状態 $x_t$ と直前の行動 $a_{t-1}$ によってそのつど因果的に決まる。つまり感覚入力 $s$ は自分のstateを持っていない。(なんだか仏教の知覚感を模しているかのようだ。)

この文書の最終目的は、このような現在、過去、未来を統一的に扱う自由エネルギー原理の式を導出することだ。たとえばそれはSchwöbel et. al.の[Neural Comput. 2018](#)で扱われている。でもいきなりでは難しいので、次のセクション3では、この生成モデルのうち、部分部分を切り取ったうえで、それぞれどのようにモデル化してゆくかを順番に見ていくことにする。

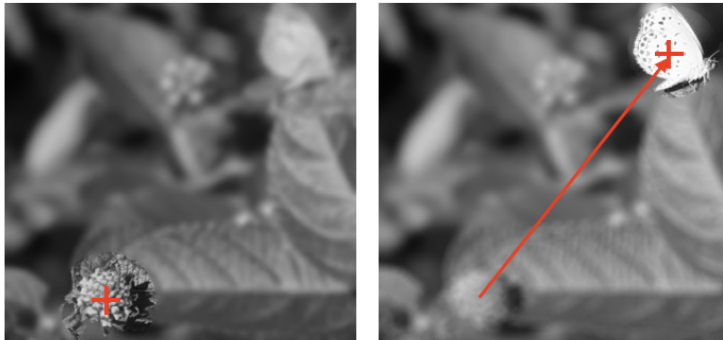
## Sec.3 実例を元にvariational FEの定式化について考える

### [3-1. 蝶と蛾の認識の例：世界設定]

ふたたび実例を用いてイメージしやすい形を説明をしてゆくことにしよう。私の以前のスライドで使った、蝶と蛾の認識の例を使って説明してゆく。こういう例だ。左下を見ていたら右上になにかあるので目を向けてみたら蝶がいた(図3-1A)。認識とは他の可能性を排除することだ。この場合、目を向けたら蛾がいる可能性もあったけど、そうではなかった。

この世界は図3-1Bのように単純で、感覚入力  $s$  は3通り、これらの感覚入力をサンプルする行動  $a$  は2通り、そこから推定される外界の状態  $x$  は2通りだけになっている。

(A)



(B)

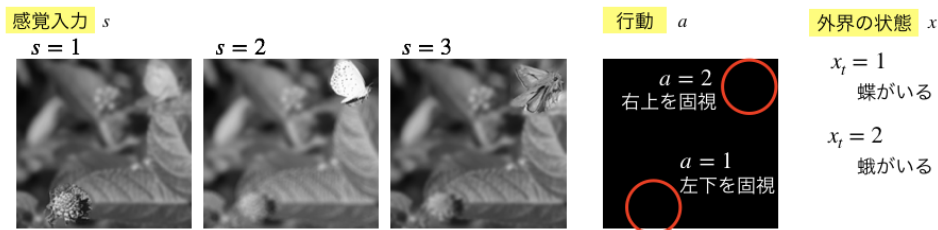


図3-1: 蝶と蛾の認識の例：世界設定

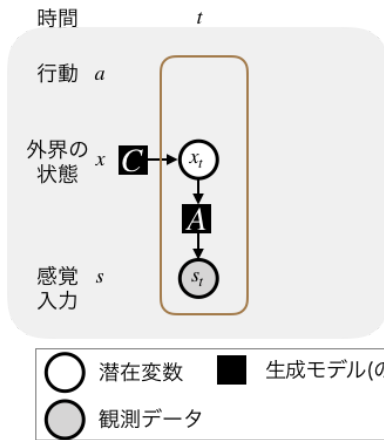
### [3-2. 行動なし、現在の状態 $x_t$ の推定]

#### [3-2-1. 生成モデル $p$ と推測 $q$ の設定]

では図2-2のうち、行動  $a$  なし、時間は現在だけ、という生成モデル  $p$  を考えてみよう。このときの生成モデルは因子グラフを用いて表現すると以下の通りになる(図3-2A)。

(A)

因子グラフ



(B)

生成モデル  $p(x_t, s_t)$   
 $= p(s_t | x_t) p(x_t)$

$\blacksquare$   $\blacksquare$

生成モデル  $p$  を特定するパラメーター

$\blacksquare$   $p(x_t) = [0.52 \quad 0.48]$   
 事前分布  $x_t = 1 \quad 2$

$\blacksquare$   $p(s_t | x_t) = \begin{bmatrix} 0.23 & 0.69 & 0.07 \\ 0.16 & 0.08 & 0.75 \end{bmatrix}$   $x_t = 1$   
 観察モデル  $s_t = 1 \quad 2 \quad 3$   $x_t = 2$

図3-2: 行動なし、時間は現在だけ

このとき生成モデル  $p$  のグラフ構造を考慮した式は以下の通り。

$$p(x_t, s_t) = p(s_t | x_t) p(x_t) \tag{3-1}$$

図3-2Bに現在の設定での生成モデルのパラメーターを表示した。事前分布  $C$  からすると、視野の中には蝶がいる確率 ( $p(x = 1) = 0.52$ ) が蛾がいる確率 ( $p(x = 2) = 0.48$ ) よりもちょっと高くなっている。また、観察モデル  $A$  によると、蝶がいる場合は  $s = 2$  の感覚入力をサンプルしている可能性が一番大きい。蛾がいる場合は  $s = 3$  の感覚入力をサンプルしている可能性が一番大きい。感覚入力  $s = 1$  のときは蝶なのか蛾なのか曖昧っぽい。でもこれについては true posterior  $p(x | s = 1)$  を計算してみないとわからないが。

つぎに推測  $q$  について考える。潜在変数は  $x_t$  のみだから、推測は  $q(x_t)$  となる。ここから推測  $q(x)$  の形を変えることで VFE  $F$  を下げることが定式化してゆくんだけど、 $q(x)$  の形を変えるというのは汎函数(functional)の概念が入ってきてややこしいので、 $q(x_t)$  という確率分布の形を決めるパラメータ  $\phi_x$  を使って表記することにする。

$$q(x_t) = q(x_t | \phi_x) \tag{3-2}$$

つまり、 $q(x_t = 1) = \phi_x$  としておけばもう片方の  $q(x_t = 2) = 1 - \phi_x$  も決まるので、 $q(x_t)$  の分布の形  $\phi_x$  ひとつで表現できる。この  $\phi_x$  を 0-1 の範囲で動かしてみる。

### [3-2-2. 変分自由エネルギー $F$ の設定]

では図3-2の設定の場合の  $F$  の式はどう書くか。これは sec.1 で用いたものと同じ構造だ。Sec.1 の式(5)に、観測データ=現在の感覚入力  $s_t$  と潜在変数=現在の外界の状態  $x_t$  を当てはめてみれば、 $F$  は以下のように定義できる。



$$\begin{aligned}
F(\phi_{x_t}) &= \mathbb{E}_{q(x_t|\phi_{x_t})} [\ln q(x_t|\phi_{x_t}) - \ln p(x_t, s_t)] \\
&= \underbrace{D_{KL}(q(x_t|\phi_{x_t})||p(x_t|s_t))}_{\text{KLD}(\phi_{x_t})} + \underbrace{-\ln p(s_t)}_{\text{Surprisal}} \\
&= \underbrace{D_{KL}(q(x_t|\phi_{x_t})||p(x_t))}_{\text{Bayesian surprise}(\phi_{x_t})} + \underbrace{-\mathbb{E}_{q(x_t|\phi_{x_t})} \ln p(x_t|s_t)}_{\text{Uncertainty}(\phi_{x_t})} \quad (3-3)
\end{aligned}$$

今の段階では行動 $a$ のことを考えていないので、 $s$ を変えることはできない。だから $F$ は $\phi_{x_t}$ のみの関数で、Surprisalは定数だ。

### [3-2-3. $F$ の最小化]

Agentは $\phi_{x_t}$ を変えることによって $F(\phi_{x_t})$ を下げるができる。今の例では $\phi_{x_t}$ を0-1まで変化させることによって、すべての可能な $q(x_t|\phi_{x_t})$ を網羅できる。それではSec.1の図1-3と同じ要領で、VFE, KLD, Surprisal, Bayesian surpriseをプロットしてみよう。

(Sec.1で使ったmatlabのコードを多少改変すればこの図を作ることができる。練習問題としてやってみるとよい。)

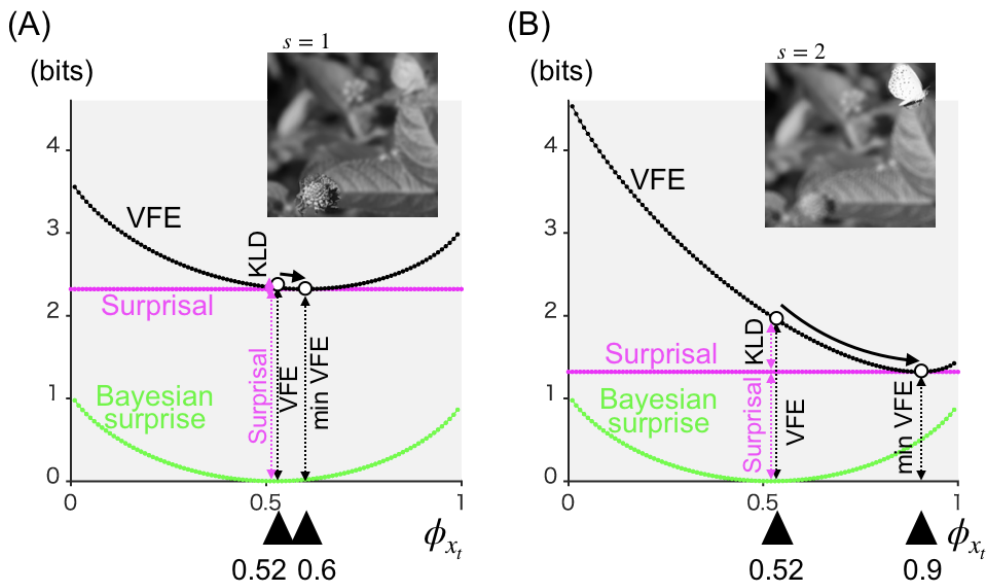


図3-3: 行動なし、時間は現在だけの場合のVFEの最小化

図3-3Aは感覚入力(観察データ)が $s = 1$  だった場合、図3-3Bは感覚入力(観察データ)が $s = 2$  だった場合を示す。 $(s = 3$  は省略するが、どういう形になるかは想像がつくだらう。)

図3-3A,Bどちらともに、 $q(x_t|\phi_{x_t})$ の初期状態として $p(x_t = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}) = \begin{bmatrix} 0.52 \\ 0.48 \end{bmatrix}$ を用いている。そこからそれぞれ $KLD = 0$ になる値へと向かっている。感覚入力 $s_1$ の場合、 $\phi_{x_t} = 0.6$ で $KLD=0$ となり、VFEが最小化する(2.32 bits)。感覚入力 $s_2$ の場合、

$\phi_{x_t} = 0.9$  で  $KLD=0$  となり、VFEが最小化する(1.32 bits)。これらはそれぞれtrue posterior  $p(x_t | s_t)$  を正しく推定できている。

直接計算によるtrue posteriorはそれぞれ以下の通り。

$$p(x_t = \begin{bmatrix} 1 \\ 2 \end{bmatrix} | s_t = [1 \ 2 \ 3]) = \begin{bmatrix} 0.6 & 0.9 & 0.1 \\ 0.4 & 0.1 & 0.9 \end{bmatrix} \quad (3-4)$$

### [3-2-4. $F$ の最小化ができたときの近似]

式(3-3)において  $KLD = 0$  となるような  $\phi_{x_t} = \phi_{x_t}^*$  を見つけることができたとき、

$$F(\phi_{x_t}^*) = -\ln p(s_t) \quad (3-5)$$

が成り立つ。繰り返すが、 $s_t$  は観察データなのでこれを変えることはできない。将来的に行動によって、サンプルされる感覚入力  $s_{t+1}$  を変えることはできるが、そのときは  $KLD(\phi_x)$  もゼロでなくなってしまう。よってこの式だけから行動について考えることはできない。ではどうすればいいかというと、そこで期待自由エネルギー  $G$  が出てくる。その話はまた後ほど。

変分推定はふつう近似をするために用いられるものだが、今は原理的な話をするために、推測  $q()$  を好きに変えることができることにした。つまり得られるものはtrue posteriorを直接計算するのと同じだが、それを  $F$  を下げるという回り道をして計算していることになる (Gershman 2019)。

---

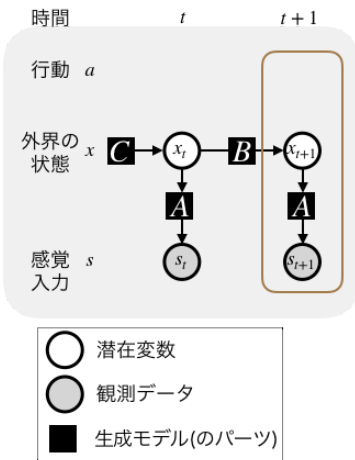
## [3-3. 行動なし、現在と過去の状態 $x_{t-1}, x_t$ の推定]

### [3-3-1. 生成モデル $p$ と推測 $q$ の設定]

つぎはこの行動なしのモデルで、時間を1ステップ進めてみよう。現在は  $t+1$  になった。このときの生成モデルの因子グラフは図3-4Aのようになる。

(A)

因子グラフ



(B)

生成モデル  $p(x_t, x_{t+1}, s_t, s_{t+1})$   
 $= p(s_{t+1} | x_{t+1}) p(s_t | x_t) p(x_{t+1} | x_t) p(x_t)$   
**A** **A** **B** **C**

生成モデル  $p$  を特定するパラメーター

**C** 事前分布  $p(x_t) = [0.52 \ 0.48]$   
 $x_t = 1 \ 2$

**B** 遷移モデル  $p(x_{t+1} | x_t) = \begin{bmatrix} 0.99 & 0.01 \\ 0.01 & 0.99 \end{bmatrix}$   $x_t = 1$   
 $x_{t+1} = 1 \ 2$

**A** 観測モデル  $p(s_t | x_t) = \begin{bmatrix} 0.23 & 0.69 & 0.07 \\ 0.16 & 0.08 & 0.75 \end{bmatrix}$   $x_t = 1$   
 $s_t = 1 \ 2 \ 3$

**A** 観測モデル  $p(s_{t+1} | x_{t+1}) = \begin{bmatrix} 0.23 & 0.69 & 0.07 \\ 0.16 & 0.08 & 0.75 \end{bmatrix}$   $x_{t+1} = 1$   
 $s_{t+1} = 1 \ 2 \ 3$

図3-4: 行動なし、時間は  $t$  から  $t+1$  に進んだ

新たに観測値  $s_{t+1}$  が追加された(灰色丸で表示)。外界の状態  $x_{t+1}$  は直前の状態  $x_t$  から影響を受ける。これは生成モデルのうち、 $B$  の遷移モデルとして記述されている(図3-4B)。これはもし  $x_t = 1$  なら、 $x_{t+1} = 1$  だし、もし  $x_t = 2$  なら、 $x_{t+1} = 2$  だという、外界の状態の恒常性のモデルになっている。

(完全に単位行列にしていないのは、計算上0が入るとKL距離が無限大になったりしてややこしいからだ。本当はKL距離においては、分母が0なら分子が0になっていないといけないのだが、それは現状では守られていないので、場当たりの処置をしている。自然は真空を嫌うらしいが、KL距離は0を嫌う。)

時点  $t+1$  における観測モデル  $A = p(s_{t+1} | x_{t+1})$  は時点  $t$  における観測モデル  $A = p(s_t | x_t)$  と同じにしてある。同じ過程については時間に依存しないのが自然だろう。

よって、生成モデルは式(3-6)のとおりとなる。

$$p(x_t, x_{t+1}, s_t, s_{t+1}) = p(s_{t+1} | x_{t+1}) p(s_t | x_t) p(x_{t+1} | x_t) p(x_t) \quad (3-6)$$

推測  $q$  はあらたに  $x_{t+1}$  も推測する必要があるので、潜在変数二つを推測することになる。

$$q(x_t, x_{t+1} | \phi_{x_t, x_{t+1}}) \quad (3-7)$$

となる。ここで  $q(x_{t+1} | \phi_{x_{t+1}})$  ではないことに注意。あらたな観測値  $s_{t+1}$  によって過去の状態  $x_t$  もアップデートすること(post-diction)を想定しているからだ。

あと、推測  $q$  のパラメーター  $\phi_{x_t, x_{t+1}}$  は以下のように3つのパラメーターからなっている。

$$\begin{aligned}\phi_{x_t, x_{t+1}} &= \{\phi_1 = q(x_t = 1, x_{t+1} = 1), \\ &\quad \phi_2 = q(x_t = 1, x_{t+1} = 2), \\ &\quad \phi_3 = q(x_t = 2, x_{t+1} = 1)\}\end{aligned}\tag{3-8}$$

あり得る  $(x_t, x_{t+1})$  の全ての組み合わせ4通りのうち、3つのパラメーターを決めれば、合計で1になるので、残りの一つは自動的に決まる。

$\phi_{x_t, x_{t+1}} \neq \{\phi_{x_t}, \phi_{x_{t+1}}\}$  であることにも注意。 $\{\phi_{x_t}, \phi_{x_{t+1}}\}$  という表記は  $\phi_{x_t}$  と  $\phi_{x_{t+1}}$  が独立して動かせることになってしまう。それは平均場近似  $q(x_t, x_{t+1} | \phi_{x_t, x_{t+1}}) = q(x_t | \phi_{x_t})q(x_{t+1} | \phi_{x_{t+1}})$  を仮定していることになるのでここではやらない。いまの説明ではなるべくexactな方法を行う。われわれが扱っているのは気体じゃないんだから、平均場近似が使える保証はなんてない。最近ではFristonのグループからもグラフィカルモデルの因子間での相関を扱えるベータ近似を使う論文が出てきている。

### [3-3-2. 変分自由エネルギー $F$ の設定]

それではこのときの  $F$  を定義してみよう。観測データ=過去と現在の感覚入力  $s_t, s_{t+1}$  と潜在変数=過去と現在の世界の状態  $x_t, x_{t+1}$  を当てはめてみれば、以下の通りとなる。

$$\begin{aligned}F(\phi_{x_t, x_{t+1}}) &= \mathbb{E}_{q(x_t, x_{t+1} | \phi_{x_t, x_{t+1}})} [\ln q(x_t, x_{t+1} | \phi_{x_t, x_{t+1}}) \\ &\quad - \ln p(x_t, x_{t+1}, s_t, s_{t+1})] \\ &= \underbrace{D_{KL}[q(x_t, x_{t+1} | \phi_{x_t, x_{t+1}}) || p(x_t, x_{t+1} | s_t, s_{t+1})]}_{KLD} \\ &\quad + \underbrace{-\ln p(s_t, s_{t+1})}_{\text{Surprisal}}\end{aligned}\tag{3-9}$$

うわあ、なんだか凄いいことになっちゃったぞ。今回はKL距離の変数部分が  $x_t, x_{t+1}$  の同時確率になっているので、 $\{x_t, x_{t+1}\}$  がとりうるすべての組み合わせ(4通り)について和を取っている。(  $x_t$  で和をとってから  $x_{t+1}$  で和をとったりするのも、平均場近似の仮定を使っていることになるのでそれはやらない。)

### [3-3-3. $F$ の最小化]

それでは蝶と蛾の例で、 $t$  に感覚入力  $s = 1$  をサンプルして、 $t + 1$  に感覚入力  $s = 2$  をサンプルした場合について考えてみよう。本当はこの  $t$  から  $t + 1$  の時間間にサッカーボールをして、外界の状態の恒常性を保ったまま、違った感覚入力をサンプルしたことになるのだけど、行動については後ほど行う。

まず比較のために、[3-2]でモデル化したように、時点  $t$  に感覚入力  $s = 1$  をサンプルして外界の状態  $x_t$  を推定したもの、そしてそれとは独立に時点  $t + 1$  に感覚入力  $s = 2$  をサンプルして外界の状態  $x_{t+1}$  を推定したものを図示してみる(図3-5)。

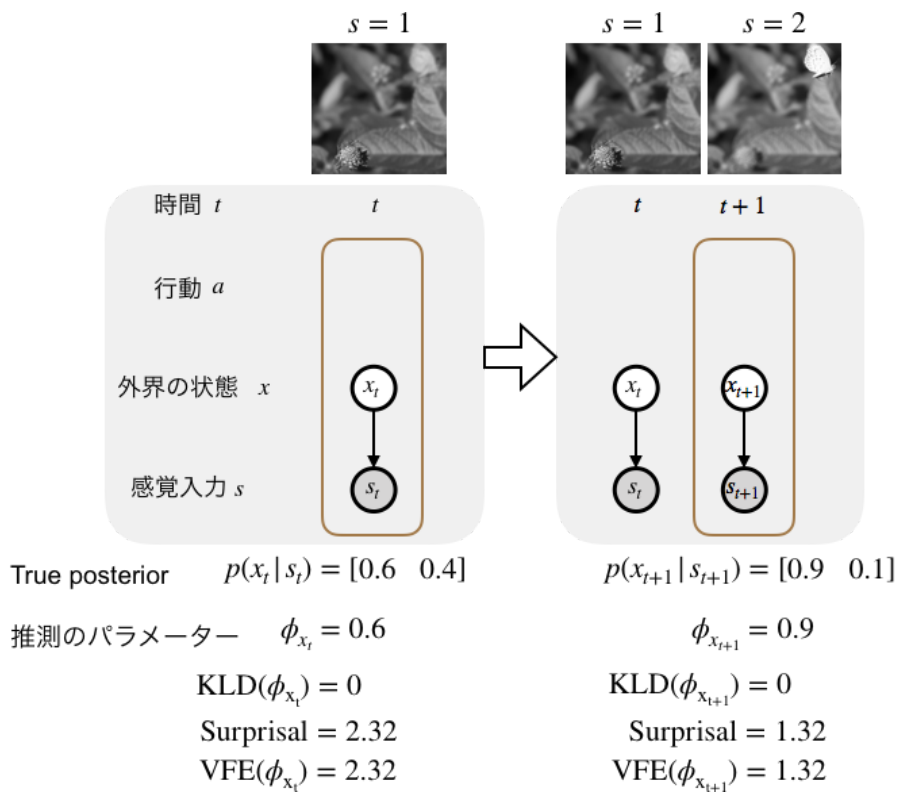
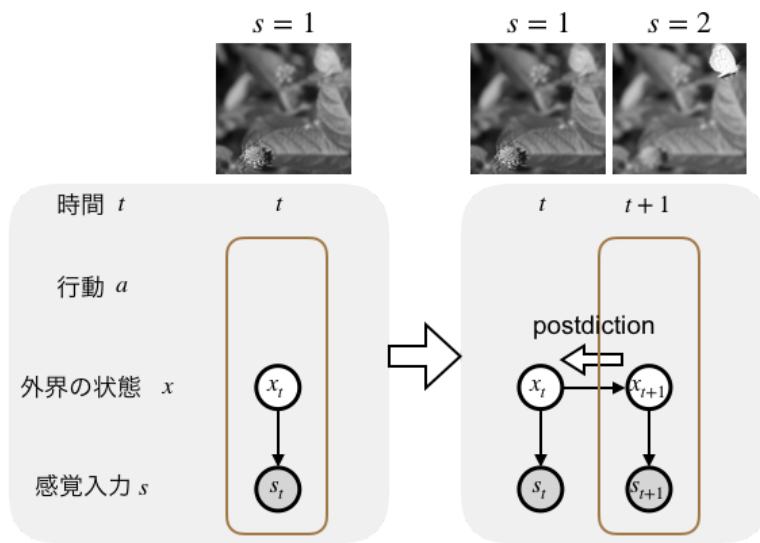


図3-5: 時間 $t$ と $t+1$ で別々に推定

時点  $t$  での「蝶がいる」  $x_t = 1$  の確率(true posterior  $p(x_t = 1 | s_t = 1)$ )は0.6 だったが、時点  $t + 1$  では「蝶がいる」  $x_{t+1} = 1$  の確率(true posterior  $p(x_{t+1} = 1 | s_{t+1} = 2)$ )は0.9 になった。 $t + 1$  でよりinformativeな感覚入力  $s = 2$  をサンプルしたことによって推定がアップデートされたわけだ。

では、今回のモデル(図3-4)で現在と過去の状態を推定するものではどうなるだろうか(図3-6)。





True posterior	$p(x_t   s_t) = [0.6 \quad 0.4]$	$p(x_t, x_{t+1}   s_t, s_{t+1}) = \begin{bmatrix} 0.92 & 0.00 \\ 0.00 & 0.07 \end{bmatrix}$
推測のパラメーター	$\phi_{x_t} = 0.6$	$\phi_{x_t, x_{t+1}} = [0.92 \quad 0.00 \quad 0.00]$
	$\text{KLD}(\phi_{x_t}) = 0$	$\text{KLD}(\phi_{x_t, x_{t+1}}) = 0$
	$\text{Surprisal} = 2.32$	$\text{Surprisal} = 3.50$
	$\text{VFE}(\phi_{x_t}) = 2.32$	$\text{VFE}(\phi_{x_t, x_{t+1}}) = 3.50$

図3-6: 時間tとt+1を繋げて推定

時点  $t$  での「蝶がいる」 $x_t = x_1$  の確率(true posterior  $p(x_t = 1 | s_t = 1)$ )が0.6であるのに対して、時点  $t + 1$  では「蝶がいる」 $x_{t+1} = 1$  の確率(true posterior  $p(x_{t+1} = 1 | s_t = 1, s_{t+1} = 2)$ )は0.92 になった。さっきよりも少々大きくなっている。情報の蓄積の効果があるということだ。

でもそれだけではない。 $t + 1$  で感覚入力  $s_2$  をサンプルしたことによって時点  $t$  での「蝶がいる」 $x_t = 1$  の確率(true posterior  $p(x_t = 1 | s_t = 1, s_{t+1} = 2)$ )も0.92 になった。つまり、過去にさかのぼって、当時の認識をアップデートさせたのだ。しかもその当時の推定は0.6だったのに。つまり、postdictionをモデル化することができているというわけだ。

この説明でわかるように、これはFEPの手柄ではなくて、ベイズのグラフィカルモデルを採用したことによる。グラフィカルモデルでtrue posteriorをexactに計算する方法において、このように因果グラフを  $x_{t+1}$  から  $x_t$  へと逆向きに情報が伝搬してゆくことを計算できる。これが確率伝搬belief propagationというものだ。

確率伝搬の過程は因子グラフで計算されたそれぞれの変数の確率が互いに影響を与えながら収束してゆく過程であり、この生成モデルの同時確率が収束する過程は自由エネルギーの最小化として表すことができる。あーそーゆーことね 完全に理解した(<-わかってない)

(追記)調べてみた。確率伝搬での自由エネルギーはベータ自由エネルギーといって、変分自由エネルギーとはべつものだった。Fristonグループの論文ではde Vries and Friston 2017がactive inferenceの因子グラフについて議論している。そこでbelief propagationについてのベータ自由エネルギーの最小化を、active

inferenceとは別の自由エネルギーの定式化として言及している。両者の数学的な関係については何も言っていない。ベータ自由エネルギーのほうは  $F = U - H$  と書いたうえで、 $U$ が内部エネルギーとして意味のある形になっているので、物理的な自由エネルギーとの対応という意味では、こちらのほうが大事っぽい。

### [3-3-4. $F$ の最小化ができたときの近似]

こうしてagentは  $\phi_{x_t, x_{t+1}}$  をアップデートすることによって  $F(\phi_{x_t, x_{t+1}})$  を最小化することができる。  $KLD = 0$  となるような  $\phi_{x_t, x_{t+1}} = \phi_{x_t, x_{t+1}}^*$  が見つかるとき、

$$F(\phi_{x_t, x_{t+1}}^*) = -\ln p(s_t, s_{t+1}) \quad (3-10)$$

が成り立つ。これを時刻 $t$ のときの値と比較してみよう。

$$F(\phi_{x_t}^*) = -\ln p(s_t) \quad (3-11)$$

式(c3)を変形して式(6)を代入すると、

$$\begin{aligned} F(\phi_{x_t, x_{t+1}}^*) &= -\ln p(s_t, s_{t+1}) \\ &= -\ln(p(s_{t+1}|s_t)p(s_t)) \\ &= -\ln p(s_{t+1}|s_t) - \ln p(s_t) \\ &= F(\phi_{x_t}^*) - \ln p(s_{t+1}|s_t) \end{aligned} \quad (3-12)$$

となる。つまり、 $p(s_{t+1}|s_t) = 1$  でないかぎり、 $F(\phi_{x_t, x_{t+1}}^*) > F(\phi_{x_t}^*)$  が成り立つ。

図3-6の例を見直してみれば、たしかに、surprisalは時点 $t$  から $t + 1$  で2.32 から3.50 に増えていた。単純な例ではあるけど「FEPに従うとFはどんどん小さくなる、そして最終的には完全平衡状態、つまり死に至るのだ」なんてことは無いということがわかる。時間経過に伴って観察データの数が増えてゆくの、過去の推測も加味した場合の  $F$  は増えてゆく一方だ。むしろ、最適なperceptionをすることによって  $F$  の増加を食い止めている、というのが実情だといえよう。なお、いまの単純なモデルでは忘却を考慮に入れてないので surprisalの項は増えてゆく一方だが、完全な記憶が無い場合には、surprisalの項は時間的になんらか定常状態になることが想像できる。

### [3-4. 行動のモデルの一手手前：推測 $q$ の代わりに感覚入力 $s$ を変える]

サブセクション3-3では、時刻 $t$  に感覚入力 $s = 1$ 、時刻 $t + 1$  に感覚入力 $s = 2$  が与えられるという状況だったが、これがすでに行動のモデルの一手手前まで来ている。図3-5で扱った、時刻 $t$  と時刻 $t + 1$  を独立した扱うもの(post-dictionなし版)についてももう一度考えてみよう。

時刻 $t$ では感覚入力 $s = 1$  (右上になにかあるっぽい画像)が与えられて、そのとき外界の状態の推定( $q(x_t)$ )は「蝶がいる」 $q(x_t = 1) = 0.6$ で「蛾がいる」 $q(x_t = 2) = 0.4$ だった(図3-3A)。次に時刻 $t + 1$ では感覚入力 $s = 2$  (右上に蝶がくっきり見える画像)が与えられて、そのとき外界の状態の推定( $q(x_{t+1})$ )は「蝶がいる」 $q(x_{t+1} = 1) = 0.9$ で「蛾がいる」 $q(x_{t+1} = 2) = 0.1$ だった(図3-3B)。このシークエンスを時刻 $t$ から時刻 $t + 1$ で視線を右上に動かした結果として捉えて直してみる。

式(3-3)の $F$ の式は $\phi_{x_t}$ のみの関数だった。これは感覚入力 $s_t = 1$ がすでに与えられた観察データであり定数であるからだった。でも未来の感覚入力 $s_{t+1}$ なら行動(視線を右上に動かす)によって変えることができる。そのように考えて $F$ の式を未来の感覚入力 $s_{t+1}$ の関数と捉えてみると式(3-13)のように書ける。

$$\begin{aligned}
 F(s_{t+1}, \phi_{x_t}) &= \underbrace{D_{KL}(q(x_{t+1}|\phi_{x_{t+1}})||p(x_{t+1}|s_{t+1}))}_{KLD(s_{t+1}, \phi_{x_{t+1}})} + \underbrace{-\ln p(s_{t+1})}_{Surprisal(s_{t+1})} \\
 &= \underbrace{D_{KL}(q(x_{t+1}|\phi_{x_{t+1}})||p(x_{t+1}))}_{Bayesian\ surprise(\phi_{x_{t+1}})} \\
 &\quad + \underbrace{-\mathbb{E}_{q(x_{t+1}|\phi_{x_{t+1}})} \ln p(s_{t+1}|x_{t+1})}_{Uncertainty(s_{t+1}, \phi_{x_{t+1}})} \tag{3-13}
 \end{aligned}$$

これで $F$ を下げるためには $\phi_{x_t}$ を変えるだけでなく、行動によって未来の $s_{t+1}$ を変える方策もできた。図3-7Aにあるように、時刻 $t$ では感覚入力として $s_t = 1$ が与えられている状態で、 $\phi_{x_t} = q(x_t = 1)$ を0.52から0.60に動かすことで $DKL = 0$ になって $F$ を最小化した(1)。そこからさらに $F$ を下げるためには $\phi_{x_t}$ を固定して、 $s_{t+1}$ を変える方策がある。もし $s = 2$ をサンプルした場合には(2)のように $F$ を下げる事ができて、さらに(3)のように $\phi_{x_t}$ を0.9まで動かすことによって $F$ を下げる事ができる。

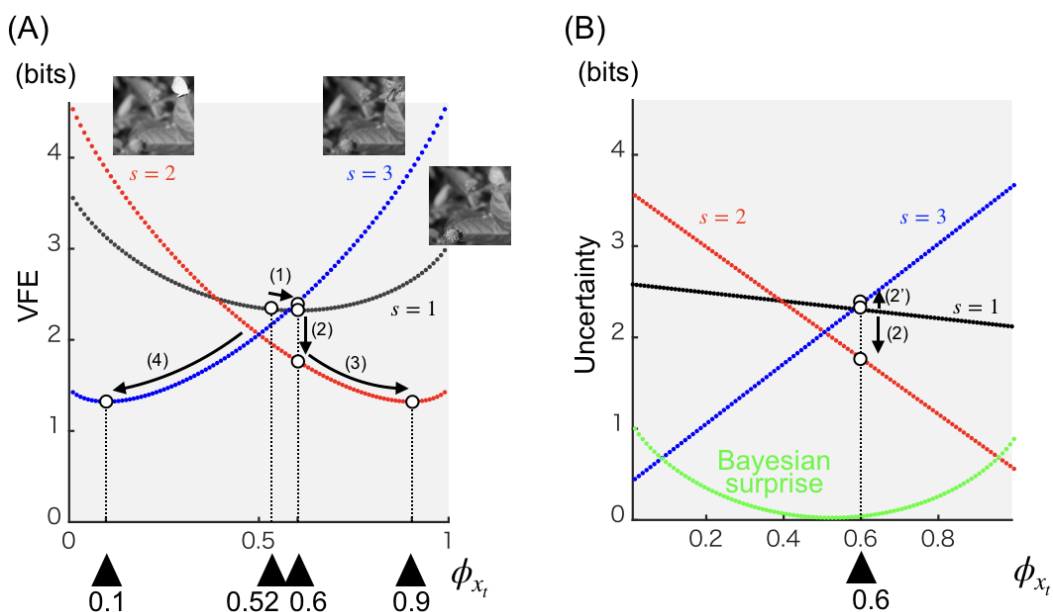


図3-7: 感覚入力 $s$ を変えることで $F$ を下げる

しかしagentができることは正確には  $s_{t+1}$  を変えることではなくて現在の行動選択  $a_t = 1$  (左下を固視)または  $a_t = 2$  (右上を固視)なので、 $s_{t+1} = 2, 3$  の両方の可能性がある。もし  $s = 3$  をサンプルした場合にはわずかではあるが  $F$  は大きくなってしまう。(最終的には(4)によって  $F$  を下げることはできるのだが。)

(2)でなにをやっているかもう少し説明してみよう。 $KLD = 0$  になったからsurprisalを下げるような  $s_{t+1}$  を選択すれば良いと思うかもしれないが、そうはいかない。なぜなら式(3-13)のとおり、 $KLD$  も  $s_{t+1}$  の関数なので、 $KLD = 0$  を保ったままSurprisalを下げることはできないから。

そこで代わりに下のBayesian surprise + Uncertaintyの式を使う。Bayesian surpriseは  $s_{t+1}$  に依存しないので、Uncertaintyが下がるように  $s_{t+1}$  を変えてやれば、現在の  $F$  をさらに低くすることができる。

このことは、 $s_{t+1} = 1, 2, 3$  についてUncertaintyをプロットしてみるとより明確になる(図3-7B)。 $\phi_{x_t} = 0.6$  を固定して、 $s_{t+1}$  を変えると、 $s_{t+1} = 2$  ではUncertaintyが下がるが、 $s_{t+1} = 3$  では上がる。Bayesian surprise(緑)は  $s_{t+1}$  に依存しないので、 $F$  の違いはUncertaintyの違いに等しい。

以上の考察から、VFEの枠組みでは、行動選択は未来のUncertaintyの値の評価によって決まるらしいということがわかる。このことがのちほど4-4-10で議論する期待自由エネルギー  $G$  の見直しという点で関わってくる。

このサブセクションでは、行動  $a$  をモデルに組み込む一歩手前の簡易的な説明を行った。行動とは感覚入力を変えることであるという直感から、VFEの式でも行動の説明ができそうだということを示した。行動についてのより正確な取り扱いは次のSec.4で行う。

---

## [4. 生成モデルの中に行動 $a$ を含める]

### [4-1-1. 生成モデルの中に行動 $a$ を含める]

さてここから生成モデルの中に行動  $a$  を含めることを考える。まずは図2-3で使った因果グラフを再掲する(図4-1A)。

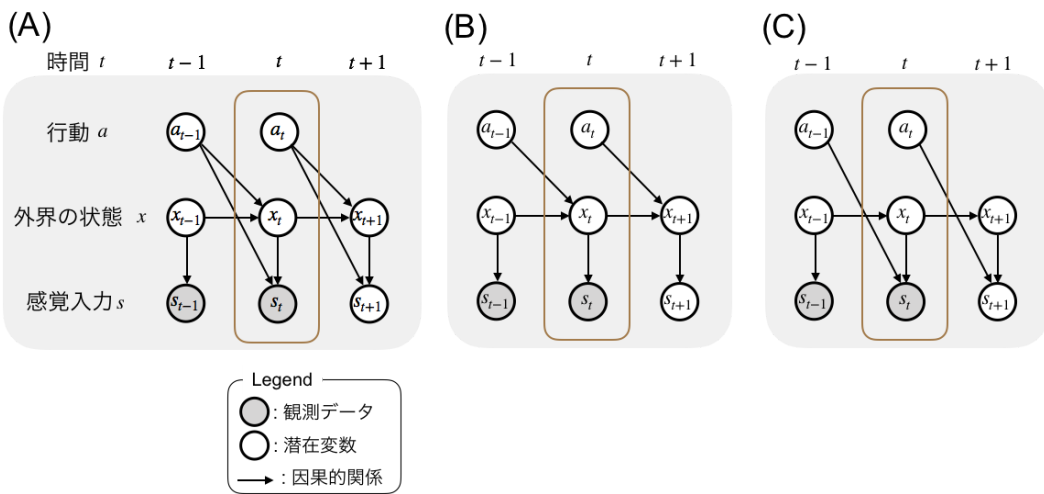


図4-1: 行動を付加した生成モデル

これはいわばフルバージョンであるが、じっさいにNeural Comput. 2017でのFristonの期待自由エネルギーの式が前提としている生成モデルは図4-1Bのとおりだ。つまり、行動  $a$  が外界の状態  $x$  のみに作用しており、感覚入力  $s$  への直接的な因果はない。一方で、ここまで用いていた蝶と蛾の例では、行動  $a$  とは眼球運動だったので、目を動かしても世界は変わらず、視野だけが変わる。つまり、外界の状態  $x$  には作用せず、感覚入力  $s$  のみに対して直接的な因果がある(図4-1C)。(これはKatahira et al 2018での病気  $x$  と検査法の選択  $a$  と検査結果  $s$  においても同じ。)

#### [4-1-2. このセクションで用いる例: スイッチで照明をつけたら明るい]

そこでこのセクションでは、Neural Comput. 2017での生成モデル(図4-1B)に基づいた実例として、図0-2で用いた照明の例「スイッチで照明をつけたら明るい」を使うことにする(図4-2)。

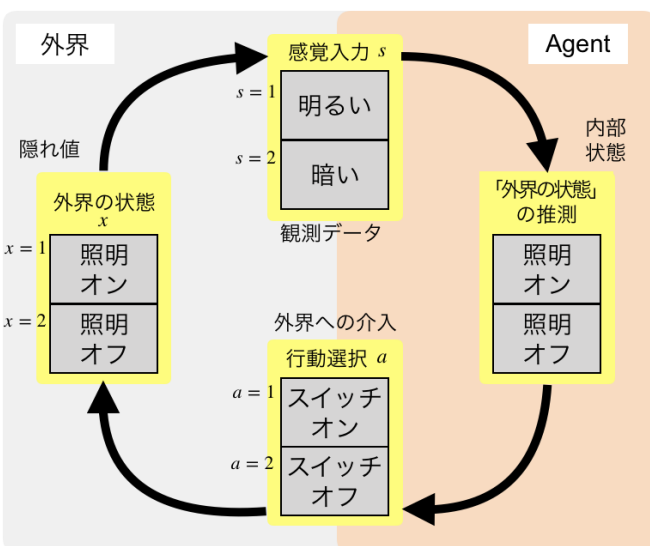


図4-2: 照明の例「スイッチで照明をつけたら明るい」



行動 $a$ 、外界の状態 $x$ 、感覚入力 $s$  はそれぞれ二つの可能な状態を持っている。 $a$  から $x$ 、 $x$  から $s$  の関係には不確実性があり、スイッチオンでも照明がオンにならないときもあれば、照明がオンでも暗いと感じるときもある。これらの関係は以下のサブセクションごとに生成モデルによって規定される。

それではこの例に沿って、expected free energy EFE, epistemic value, extrinsic value の式の導出までやってしまうことにする。

ToDo: 蝶と蛾の例はサッケードのモデルとして重要なのでまたあとで戻ってくる予定。

## [4-2. 行動あり、現在の状態 $x_t$ と過去の行動 $a_{t-1}$ の推定(1)]

### [4-2-1. 生成モデル $p$ と推測 $q$ の設定]

ではまず図3-2の現在の知覚の生成モデルに過去の行動 $a_{t-1}$  を付加することを考える。過去の行動 $a_{t-1}$  を潜在変数として捉えるならば、生成モデルのグラフィカルモデルによる表現は以下の通りになる(図4-3A)。

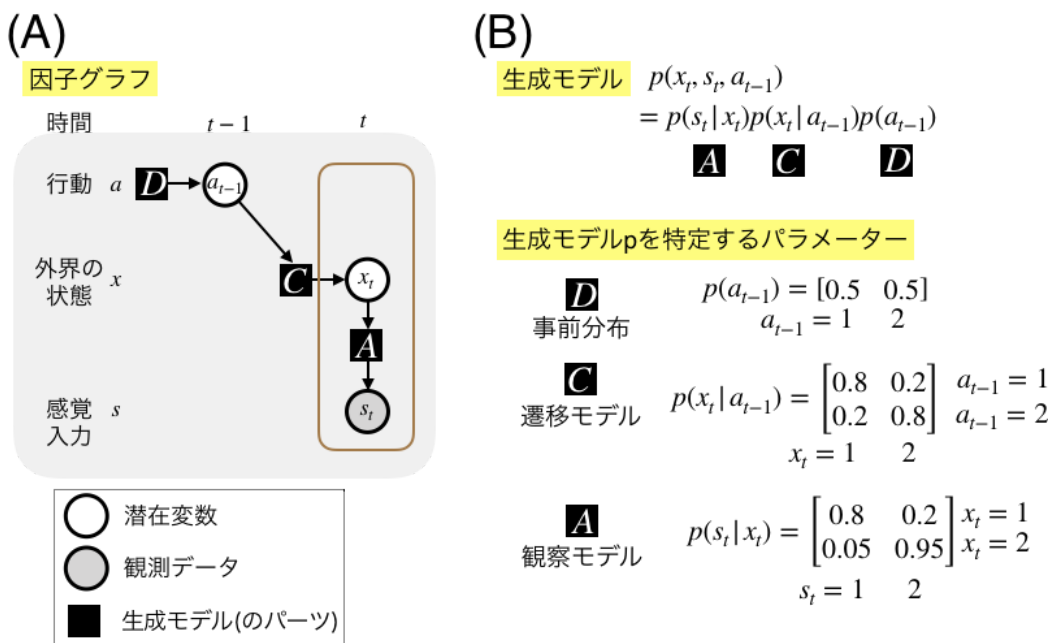


図4-3: 過去の行動を付加した生成モデル(1)

このとき生成モデルは以下のように書くことができる。

$$p(x_t, s_t, a_{t-1}) = p(s_t | x_t) p(x_t | a_{t-1}) p(a_{t-1}) \quad (4-1)$$

生成モデルは3つの因子に分解することができる(図4-3B)。  $A$  は観察データを生成する「観察モデル」 ( $p(s_t|x_t)$ )、  $C$  は状態の推移を決める「遷移モデル」  $p(x_t|a_{t-1})$ 、  $D$  は行動  $a$  の「事前分布」  $p(a_{t-1})$ 、 と呼ぶ。 行動  $a$  の「事前分布」とは、結果となる  $x$  や  $s$  に依存せずにどちらを選択するかのバイアスを表すから「preference」とも呼ぶ。この呼び方は後で出てくる。

$A$  と  $C$  のそれぞれの値をみてもらうと、不確定性のあるシステムであることがわかる。スイッチがオンでも照明がつかない確率が20%あるし ( $p(x = 2|a = 1) = 0.2$ )、照明がオンでも暗く感じる確率が20%ある ( $p(s = 2|x = 1) = 0.2$ )。照明がオフでも明るく感じる確率は5%なので ( $p(s = 1|x = 2) = 0.05$ )、ここに非対称性が入ってる。

また、推測  $q$  のほうは以下のように書く。直前の行動  $a_{t-1}$  は隠れ値なので  $x_t$  と同じく推定してやる必要がある。

$$q(x_t, a_{t-1} | \phi_{x_t, a_{t-1}}) \quad (4-2)$$

式(3-8)と同様、  $\phi_{x_t}$  と  $\phi_{a_{t-1}}$  は独立して動かさずに、同時に推定するため、パラメータとしては3変数あることになる。

## [4-2-2. 変分自由エネルギー $F$ の設定]

この場合の観測データは  $s_t$ 、潜在変数は  $x_t, a_{t-1}$  となるので、VFEの式は以下の通りになる。

$$\begin{aligned} F(\phi_{x_t, a_{t-1}}) &= \mathbb{E}_{\tilde{q}} [\ln \tilde{q} - \ln p(x_t, s_t, a_{t-1})] \\ &= \underbrace{D_{KL}(q(x_t, a_{t-1} | \phi_{x_t, a_{t-1}}) || p(x_t, a_{t-1} | s_t))}_{\text{KLD}(\phi_{x_t, a_{t-1}})} + \underbrace{-\ln p(s_t)}_{\text{Surprisal}} \quad (4-3) \end{aligned}$$

となる。Surprisalの中身は  $s_t$  となっていて、図3-2の生成モデル(現在のみ、運動なし)と変わらないことがわかる。KL距離を計算するときに  $x_t, a_{t-1}$  の両方で総和をとるので、KL距離は  $\phi_{x_t, a_{t-1}}$  のみの関数となる。

(さてあとはposteriorを計算して、 $F$  をプロットして、と行くところなのだが、ここでは省略する。)

ところで図4-3において過去の行動  $a_{t-1}$  を潜在変数として捉えるのは妥当だったろうか？もし生成モデルが、行動  $a_{t-1}$  が事前分布  $p(a)$  から確率的に選ばれて、その結果として外界の状態  $x_t$  と感覚入力  $s_t$  が引き起こされた、という場合なら合っている。しかし、expected free energyで想定されているのはそうではない。

## [4-3. 行動あり、現在の状態 $x_t$ と過去の行動 $a_{t-1}$ の推定(2)]

### [4-3-1. 生成モデル $p$ と推測 $q$ の設定]

図4-3の生成モデルはexpected free energyで用いられているものとは異なっている。Expected free energyにおいては、行動  $a_{t-1}$  はその時点での期待自由エネルギー  $G_{t-1}$  を計算した結果として確定する。つまり、行動  $a_{t-1}$  は推定すべき潜在変数ではなくて、生成モデルの一部である事前分布  $C$  を特定するパラメーターと捉えるべきなのだ。(「パラメーターとして扱うこと」の意味については、Sec.2-4を見直してほしい。)

そこで図4-4Aで表示した因子グラフは、行動  $a_{t-1}$  を生成モデルのパラメーター(白四角)として表現している。

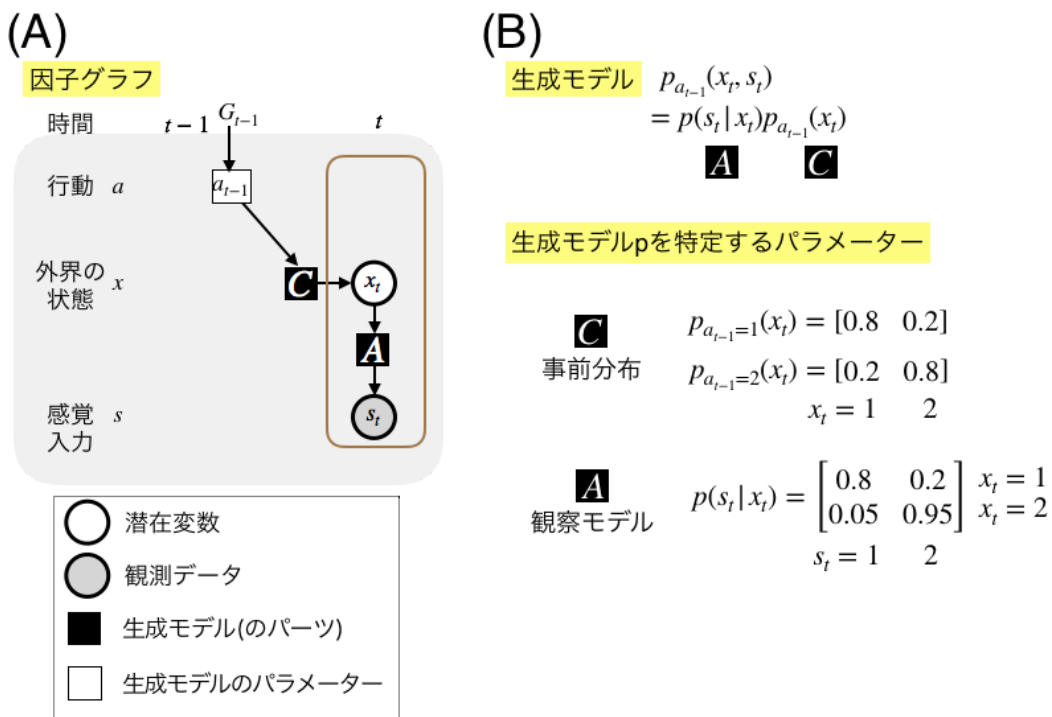


図4-4: 過去の行動を付加した生成モデル(2)

つまり生成モデルは以下のように書くことができる(図4-4B)。

$$p_{a_{t-1}}(x_t, s_t) = p(s_t | x_t) p_{a_{t-1}}(x_t) \quad (4-4)$$

図4-2についての生成モデルの式(4-1)と比べて、生成モデルの変数に  $a_{t-1}$  が入っていないことがわかる。その代わりに、 $a_{t-1}$  がパラメーターであることを明示するために  $p_{a_{t-1}}$  という表現をしている。

また、今回の生成モデルでは、 $a_{t-1} = 1, = 2$  のそれぞれについてべつべつの生成モデルがある。つまり、事前分布  $C$  の部分を必要に応じて差し替える必要がある(図4-4B)。この点でも図4-3Bの生成モデルとは異なっている。あちらでは  $C$  の部分に潜在変数  $a_{t-1}$  とそ

の事前分布  $D$  が入力するので、ひとつの生成モデルの中で、 $a_{t-1} = 1, = 2$  の両方を計算する必要がある。

生成モデル  $p$  を特定するパラメーター(図4-4B)自体はさきほどの図4-3Bと同じものを用いているが、上記の違いによって、計算される  $F$  などは異なってくる。

また、推測  $q$  のほうは以下のように書く。

$$q(x_t | \phi_{x_t}) \quad (4-5)$$

形式上は図3-2とまったく同じになった。あとは  $\phi_{x_t}$  を0-1で動かして推定してやればよい。

### [4-3-2. 変分自由エネルギー $F$ の設定]

この場合の観測データは  $s_t$ 、潜在変数は  $x_t$  となるので、VFEの式は以下の通りになる。

$$\begin{aligned} F(\phi_{x_t}, a_{t-1}) &= \mathbb{E}_{q(x_t | \phi_{x_t})} [\ln q(x_t | \phi_{x_t}) - \ln p_{a_{t-1}}(x_t, s_t)] \\ &= \underbrace{D_{KL}(q(x_t | \phi_{x_t}) || p(x_t | s_t))}_{\text{KLD}(\phi_{x_t})} + \underbrace{-\ln p_{a_{t-1}}(s_t)}_{\text{Surprisal}(a_{t-1})} \end{aligned} \quad (4-6)$$

となる。KLDの中身の  $p(s_t | x_t)$  は図4-3Bおよび式(4-4)からわかるように、 $a_{t-1}$  には依存しない。

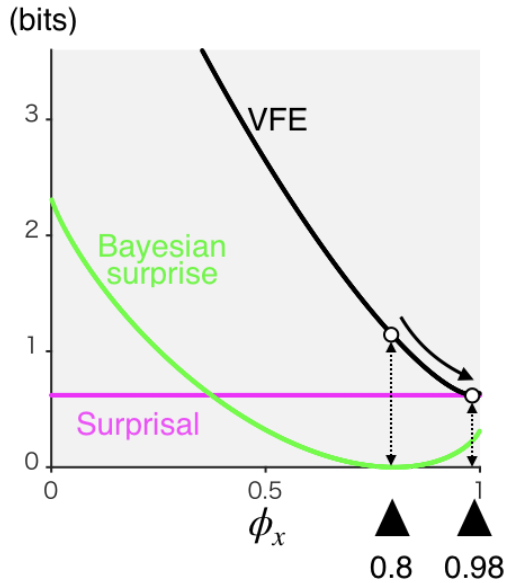
Surprisalの中身は  $s_t$  となっていて、式(3-3)および式(4-3)とは変わらない。しかしこれを計算するためには、 $p_{a_{t-1}}(s_t) = \sum_{x_t} p_{a_{t-1}}(x_t, s_t)$  というように同時確率の周辺化が必要になるから  $a_{t-1} = 1, = 2$  についてそれぞれ別に作る必要がある。よって、式(4-3)での surprisalとは違う結果になる。つまり、行動  $a_{t-1}$  を潜在変数として捉えるか、パラメーターとして捉えるかは、実際のVFEの計算自体に影響を及ぼす。

$F, \text{Surprisal}$ にとって  $a_{t-1}$  がパラメーターになっている、つまり別々の  $a_{t-1}$  に対して別々の値が出てくることを示すために、それぞれの ( ) の中に表記している。こちらは生成モデル  $p$  の中身の場合とは違って確率ではないので、この表現で正しい。

### [4-3-3. 変分自由エネルギー $F$ の最小化]

Agentは  $\phi_{x_t}$  を変えることによって  $F(\phi_{x_t})$  を下げることができる。今の例では  $\phi_{x_t}$  を0-1まで変化させることによって、すべての可能な  $q(x_t | \phi_{x_t})$  を網羅できる。それではSec.3-2-3の図3-3と同じ要領で、スイッチと照明の例について、VFE, Surprisal, Bayesian surpriseをプロットしてみよう(図4-5,図4-6)。

(A)  $s = 1$  (「明るい」)



(B)  $s = 2$  (「暗い」)

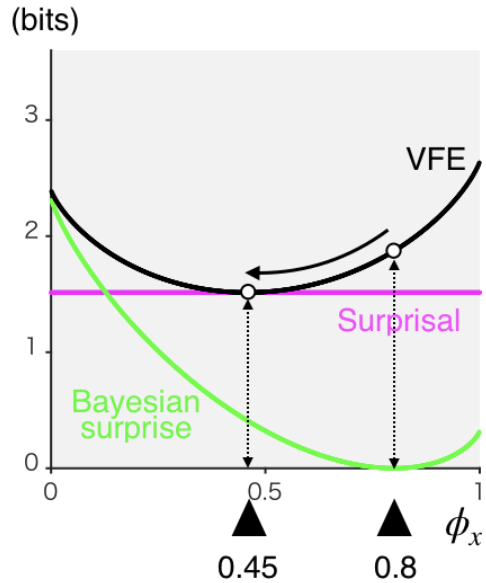
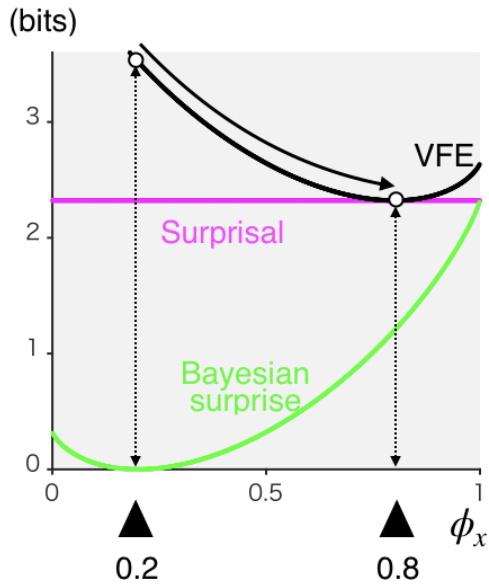


図4-5: VFEの最小化 ( $a=1$ : スイッチオン)

(A)  $s = 1$  (「明るい」)



(B)  $s = 2$  (「暗い」)

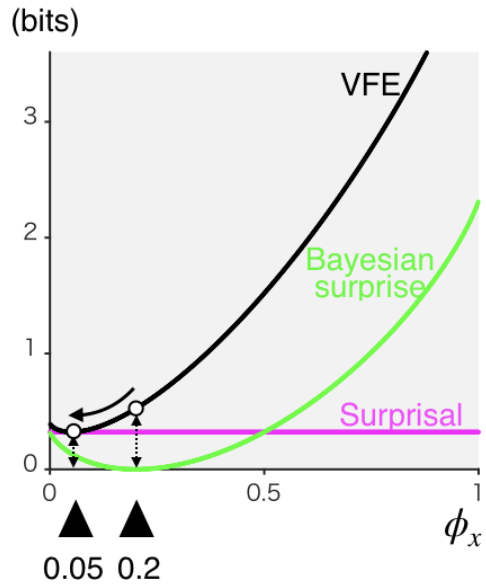


図4-6: VFEの最小化 ( $a=2$ : スイッチオフ)

まず図4-5が $a = 1$ 、つまりスイッチオンにしたときのVFEの最小化。観測データとして $s = 1$  (明るい)、 $s = 2$  (暗い)のそれぞれでVFEを最小化している。どちらも $\phi_{x_t}$ の初期状態は事前分布 $p(x_t)$ で決まるから、 $a = 1$ では0.8からスタートしている。

図4-5Aのときは、スイッチオンしたら明るかったわけで、照明がオンであるという推測は初期状態0.8からより高くなって0.98になった。いっぽうで図4-5Bでは、スイッチオンしたのに暗かったわけで、照明がオンであるという推測は初期状態0.8から0.45へとかなり下がってる。その気持ちわかるなー。



図4-6が $a = 2$ 、つまりスイッチオフにしたときのVFEの最小化。大まかには 図4-5のときと左右ひっくり返した感じになっているが、生成モデルを確認してもらえばわかるように、完全に左右対称ではない。

以上をまとめてテーブルにすると以下の通りになる。

$a = 1$	$s = 1$	$s = 1$	$s = 2$	$s = 2$
	初期状態	VFE最小化	初期状態	VFE最小化
$\phi_{x_t}$	0.8	0.985	0.8	0.457
VFE	1.1219	0.6215	1.8723	1.5146
Bayesian surprise	0	0.2396	0	0.4133

$a = 2$	$s = 1$	$s = 1$	$s = 2$	$s = 2$
	初期状態	VFE最小化	初期状態	VFE最小化
$\phi_{x_t}$	0.2	0.800	0.2	0.050
VFE	3.5219	2.3219	0.5236	0.3219
Bayesian surprise	0	1.2000	0	0.1355

テーブル1: VFE最小化のまとめ

そういうわけで、やってることは図3-3とほとんど同じであることがわかる。でもこれが次のサブセクション4-4でexpected free energyを計算するときの題材になるので、この図に再訪することになる。

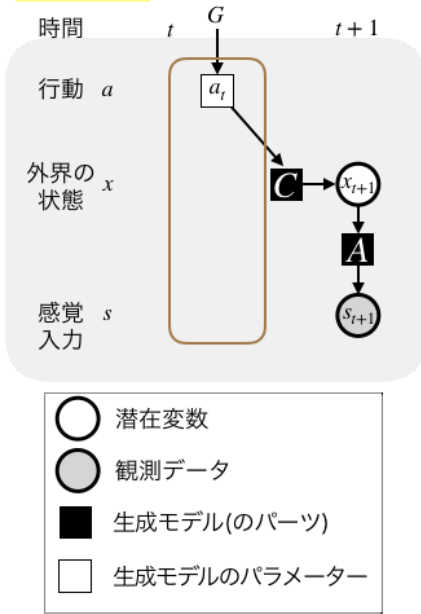
## [4-4. 行動あり、未来の状態 $x_{t+1}$ と未来の感覚入力 $s_{t+1}$ の推定(1)]

### [4-4-1. 生成モデル $p$ と推測 $q$ の設定]

ここまでくれば、現在の行動 $a_t$  とその直後の外界の状態 $x_{t+1}$  および感覚入力  $s_{t+1}$  についての生成モデルを作るときには、はどのようになるか想像がつくだろう。図4-4Aから、時間をずらしたものを作ってみた(図4-6A)。

(A)

因子グラフ



(B)

生成モデル

$$p_{a_t}(x_{t+1}, s_{t+1}) = p(s_{t+1} | x_{t+1}) p_{a_t}(x_{t+1})$$

A C

生成モデルpを特定するパラメーター

C  $p_{a_t=1}(x_{t+1}) = [0.8 \ 0.2]$

事前分布  $p_{a_t=2}(x_{t+1}) = [0.2 \ 0.8]$   
 $x_{t+1} = 1 \ 2$

A 観察モデル  $p(s_{t+1} | x_{t+1}) = \begin{bmatrix} 0.8 & 0.2 \\ 0.05 & 0.95 \end{bmatrix}$   
 $s_{t+1} = 1 \ 2$

図4-7: 現在の行動計画の因子モデルと生成モデル

これでいいかということ、ひとつ考慮しなければならないことがある。未来の感覚入力  $s_{t+1}$  は観測データでいいのだろうか？  $s_{t+1}$  はまだ観測されていないのだから。

でもここで考えてほしいのだけど、もし未来の感覚入力  $s_{t+1}$  が潜在変数だとしたらすべてが潜在変数になってしまう。観測データがないのだから、推定のしようがない。これじゃあ部分観測マルコフ決定過程POMDPではなくて、無観測マルコフ決定過程NOMDPだよ。エスパーか！

ここで二つの方法がありうる。一つは観測データとして現在および過去の感覚入力  $s_t, s_{t-1}, \dots$  を使うというものだ。これはのちほど紹介するSchwöbel et. al. 2018論文およびParr and Friston 2018論文でとった方法だ。もう一つがFriston et. al. 2017論文で使われたように未来の感覚入力  $s_{t+1}$  を観測データとして捉える、ただしcounter-factualな観測データとして捉える、というものだ。いまは後者について説明する。

概念整理のために図4-8のように考えてみる。潜在変数と観測データに加えて、factualかcounter-factualかという軸を考慮する。未来の感覚入力  $s_{t+1}$  はcounter-factualな観測データだ。よってこれまでに出てきたものと扱いが異なる。いっぽうで、未来の外界の状態  $x_{t+1}$  のほうはcounter-factualな潜在変数だが、潜在変数についてはfactualだろうがcounter-factualだろうが推定するだけなので、扱いを変える必要がない。

潜在変数	$x_t$	$x_{t+1}$
観測データ	$s_t$	$s_{t+1}$
	Factual	Counter-factual

図4-8: [潜在変数/観測データ]と[factual/counter-factual]の軸

それではあらためて、図4-7Aの因子グラフについて生成モデルを作ってみると以下のようになる(図4-7B)。

$$p_{a_t}(x_{t+1}, s_{t+1}) = p(s_{t+1}|x_{t+1})p_{a_t}(x_{t+1}) \quad (4-7)$$

さきほども書いたように、 $a_t$  は事前分布  $C$  の形を決めるパラメーターだ。よって  $C$  は counter-factualな行動選択  $a_t$  ごとに、 $p_{a_t=1}(x_{t+1})$  と  $p_{a_t=2}(x_{t+1})$  と別々に定義する。観察モデル  $A$  のほうは  $a_t$  に依存しない。

また、推測  $q$  のほうは以下のように書く。潜在変数は  $x_{t+1}$  ひとつだけだから、形式的には図3-2の現在の外界の状態を推定するのと同じ形になる。

$$q(x_{t+1}|\phi_{x_{t+1}}) \quad (4-8)$$

#### [4-4-2. 変分自由エネルギー $F$ の設定]

この場合の観測データは  $s_{t+1}$ 、潜在変数は  $x_{t+1}$  となるので、VFEの式は以下の通りになる。

$$\begin{aligned} F(\phi_{x_{t+1}}, a_t) &= \mathbb{E}_{q(x_{t+1}|\phi_{x_{t+1}})}[\ln q(x_{t+1}|\phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \\ &= \underbrace{D_{KL}(q(x_{t+1}|\phi_{x_{t+1}})||p(x_{t+1}|s_{t+1}))}_{\text{KLD}(\phi_{x_{t+1}})} + \underbrace{-\ln p_{a_t}(s_t)}_{\text{Surprisal}(a_t)} \end{aligned} \quad (4-9)$$

式(4-6)と同様に、KLDの中身は  $a_t$  に依存しなくて、surprisalの中身は  $a_t$  に依存する。

ここまでは図3-2の現在の外界の状態を推定するのと同じ形だが、これでは終わらない。なぜならこの式に入っている  $s_{t+1}$  は counter-factualな観測データなので、実際には一つには決まってないからだ。そこでこのようにして定義された  $F(\phi_{x_t}, a_{x_t})$  をさらに  $s_{t+1}$  ごとの期待値で足し合わせてやる必要がある。これが期待自由エネルギーEFE  $G$  だ。

ではどうやって  $s_{t+1}$  ごとに足し合わせるかというと、二つのやり方がある。1つ目はこれまで計算した  $F$  に  $q(s_{t+1})$  を期待値として重み付け平均してやるものだ(式4-10)。

$$G(\phi_{x_{t+1}}, a_t) = \mathbb{E}_{q(s_{t+1})}F(\phi_{x_{t+1}}, a_t) \quad (4-10')$$

もうひとつは  $F$  に  $q(s_{t+1}|x_{t+1})$  を期待値としてかけてやるものだ(式4-10)。

$$\begin{aligned}
G(\phi_{x_{t+1}}, a_t) &= \mathbb{E}_{q(s_{t+1}|x_{t+1}, \phi_{x_{t+1}})} \mathbb{E}_{q(x_{t+1}|\phi_{x_{t+1}})} [\ln q(x_{t+1}|\phi_{x_{t+1}}) \\
&\quad - \ln p_{a_t}(x_{t+1}, s_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1}|x_{t+1}, \phi_{x_{t+1}}) * q(x_{t+1}|\phi_{x_{t+1}})} [\ln q(x_{t+1}|\phi_{x_{t+1}}) \\
&\quad - \ln p_{a_t}(x_{t+1}, s_{t+1})] \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1}|\phi_{x_{t+1}})} [\ln q(x_{t+1}|\phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \quad (4-10)
\end{aligned}$$

こちらは結果として、 $F$  の計算に使っていた  $q(x_{t+1})$  と併せて、 $q(x_{t+1}, s_{t+1})$  で重み付け平均していることになる。(なお、2行目から3行目のところでは、 $\mathbb{E}_{A*B} = \mathbb{E}_A * \mathbb{E}_B$  を使ってる。)

式(4-10)と式(4-10')では異なる結果になる(詳しくはサブセクション4-4-4を参照で)。どっちが良いかといえば、 $q(x_{t+1}, s_{t+1})$  を使うほうが  $x_{t+1}, s_{t+1}$  の相関の情報を失わないので良さそうだ。

Neural Comput. 2017のp.10で定義されているFristonの期待自由エネルギーの式では、期待値として式(4-10)のほうが採用されている。それは以上の理屈からは納得できると思う。

## [4-4-2'. 期待値の重みについて詳しく]

では実際にこの期待値  $q(x_{t+1}, s_{t+1}|\phi_{x_{t+1}})$  の部分を計算するにはどうすればよいか。Neural Comput. 2017のp.10によれば、このように書いてある。(表記は現在の文書に合わせてる。)

$$\begin{aligned}
q(x_{t+1}, s_{t+1}|\phi_{x_{t+1}}) &= q(s_{t+1}|x_{t+1}, \phi_{x_{t+1}})q(x_{t+1}|\phi_{x_{t+1}}) \\
&= p(s_{t+1}|x_{t+1})q(x_{t+1}|\phi_{x_{t+1}}) \quad (4-11)
\end{aligned}$$

つまり、まず未来の外界の状態  $x_{t+1}$  の推定  $q(x_{t+1}|\phi_{x_{t+1}})$  をしておいて、それに生成モデルのうち観察モデルの部分  $p(s_{t+1}|x_{t+1})$  を掛け算してやって  $s$  についての推定を作っている。

よって実際の計算では、期待値として掛けるべきものは  $p(s_{t+1}|x_{t+1})$  であるとわかった。式(4-10)を書き直すと、以下が期待自由エネルギー  $EFE = G$  の式となる。

$$\begin{aligned}
G(\phi_{x_{t+1}}, a_t) &= \mathbb{E}_{p(s_{t+1}|x_{t+1})} F(\phi_{x_{t+1}}, a_t) \\
&= \mathbb{E}_{p(s_{t+1}|x_{t+1})} [\mathbb{E}_{q(x_{t+1}|\phi_{x_{t+1}})} [\ln q(x_{t+1}|\phi_{x_{t+1}}) \\
&\quad - \ln p_{a_t}(x_{t+1}, s_{t+1})]] \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1}|\phi_{x_{t+1}})} [\ln q(x_{t+1}|\phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \quad (4-12)
\end{aligned}$$

じつはここでFriston et al 2017ではさらにもう一段階近似を行っている。

$$\begin{aligned}
q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}}) &= q(s_{t+1} | x_{t+1}, \phi_{x_{t+1}}) q(x_{t+1} | \phi_{x_{t+1}}) \\
&= p(s_{t+1} | x_{t+1}) q(x_{t+1} | \phi_{x_{t+1}}) \\
&\approx p_{a_t}(x_{t+1}, s_{t+1} | s_t)
\end{aligned} \tag{4-11'}$$

また、Friston et. al. Cognitive Neuroscience 2015のp.192では以下の通りに書かれている。

$$q(x_{t+1}, s_{t+1}) = \mathbb{E}_{q(x_t)} [p_{a_t}(x_{t+1}, s_{t+1} | x_t)] \tag{4-11''}$$

これらは何を意味しているかということ、式(4-11)の  $q(x_{t+1} | \phi_{x_{t+1}})$  の部分さえも現在の観測データ  $s_t$  から作っているということだ。現在の生成モデルおよび照明の例では過去のことは考えていない。ここでは話を簡単にするために、 $q(x_t) = p(x)$  に加えて、遷移モデル  $p(x_{t+1} | x_t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$  を仮定する。これで  $q(x_{t+1}) = q(x_t) = p(x)$  として計算することにする。これはFriston et al 2017やKatahira et al 2018で採用されている方法だ。

ただしこの文書の方針は、「単純な例は用いるが近似はなるべく用いない」というものだった。以下の照明の例では、 $q(x_{t+1}) = p(x)$  の近似を使わずに  $q(x_{t+1} | \phi_{x_{t+1}})$  は  $\phi_{x_{t+1}}$  によって動かした計算も行う。じっさいやってみてわかったのだが、 $\phi_{x_{t+1}}$  を動かすと、 $G(a_t)$  を  $a_t$  ごとに最小化することができる。

### [4-4-3. 照明の例を用いた期待自由エネルギー $G$ の計算]

では以上の結果を用いて、照明の例で期待自由エネルギーEFE  $G$  を計算してみよう。

実際の計算方法についてだが、式(4-12)1行目のように、VFEを計算してからそれに期待値をかけて計算することはできない。なぜなら内側の  $F$  を計算するとき、すべての  $x_{t+1}$  で足し合わせて  $x_{t+1}$  が消えているから、外側の期待値の  $x_{t+1}$  と対応させることができない。あくまでも式(4-12)3行目のように  $q(x_{t+1}, s_{t+1})$  で期待値をとる必要がある。

まずこの期待値の部分について計算しておこう。式(4-11)に、

$q(x_{t+1} = 1 | \phi_{x_{t+1}}) = \phi_{x_{t+1}}$  および  $q(x_{t+1} = 2 | \phi_{x_{t+1}}) = 1 - \phi_{x_{t+1}}$ 、それから観察モデルの値  $p(s_{t+1} | x_{t+1}) = \begin{bmatrix} 0.8 & 0.2 \\ 0.05 & 0.95 \end{bmatrix}$  を入れてやると、

$$\begin{aligned}
q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}}) &= p(s_{t+1} | x_{t+1}) q(x_{t+1} | \phi_{x_{t+1}}) \\
q(x_{t+1} = 1, s_{t+1} = 1 | \phi_{x_{t+1}}) &= p(s_{t+1} = 1 | x_{t+1} = 1) \phi_{x_{t+1}} \\
q(x_{t+1} = 1, s_{t+1} = 2 | \phi_{x_{t+1}}) &= p(s_{t+1} = 2 | x_{t+1} = 1) \phi_{x_{t+1}} \\
q(x_{t+1} = 2, s_{t+1} = 1 | \phi_{x_{t+1}}) &= p(s_{t+1} = 1 | x_{t+1} = 2) (1 - \phi_{x_{t+1}}) \\
q(x_{t+1} = 2, s_{t+1} = 2 | \phi_{x_{t+1}}) &= p(s_{t+1} = 2 | x_{t+1} = 2) (1 - \phi_{x_{t+1}})
\end{aligned} \tag{4-13}$$

これをプロットしてやると図4-9の通りとなる。



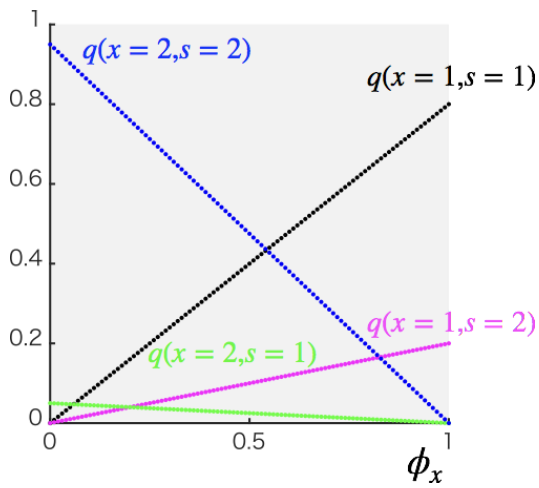


図4-9: 照明の例における $q(x,s)$

$\phi_{x_{t+1}}$  を0-1の間で動かすと、どこでも  $\sum_s \sum_x q(x, s) = 1$  になっていることが確認できる。あらためて、推測  $q(x, s)$  という同時確率が  $\phi_{x_{t+1}}$  ひとつの関数であること、つまり、agentがここで変化させているのは外界の状態の推定  $q(x)$  だけで、感覚入力  $s$  についての推定はそれによって一意に決まる、というモデルになっていることがわかる。(これ自体は自明なことではなく、あくまでFriston論文ではそういうモデルになっているというだけのこと。)

それでは式(4-12)に基づいて、 $G$  を計算してみよう。式(4-13)を使って、 $x = 1, 2$  および  $s = 1, 2$  のすべての変数の組み合わせでの和の形に変形してやる。あと、2行目から3行目の変換で、 $p_{a_t}(x_{t+1}, s_{t+1}) = p(s_{t+1}|x_{t+1})p_{a_t}(x_{t+1})$  を使っている。生成モデルの設定より、観察モデル  $p(s_{t+1}|x_{t+1})$  の部分は  $a$  に依存しないので、こちらはこの段階で値を代入できる。

$$\begin{aligned}
& G(\phi_{x_{t+1}}, a_t) \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | \phi_{x_{t+1}}) - \ln p(s_{t+1} | x_{t+1}) - \ln p_{a_t}(x_{t+1})] \\
&= q(x_{t+1} = 1, s_{t+1} = 1 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 1 | \phi_{x_{t+1}}) \\
&\quad - \ln p(s_{t+1} = 1 | x_{t+1} = 1) - \ln p_{a_t}(x_{t+1} = 1)] \\
&\quad + q(x_{t+1} = 1, s_{t+1} = 2 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 1 | \phi_{x_{t+1}}) \\
&\quad - \ln p(s_{t+1} = 2 | x_{t+1} = 1) - \ln p_{a_t}(x_{t+1} = 1)] \\
&\quad + q(x_{t+1} = 2, s_{t+1} = 1 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 2 | \phi_{x_{t+1}}) \\
&\quad - \ln p(s_{t+1} = 1 | x_{t+1} = 2) - \ln p_{a_t}(x_{t+1} = 2)] \\
&\quad + q(x_{t+1} = 2, s_{t+1} = 2 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 2 | \phi_{x_{t+1}}) \\
&\quad - \ln p(s_{t+1} = 2 | x_{t+1} = 2) - \ln p_{a_t}(x_{t+1} = 2)] \\
&= 0.8 * \phi_{x_{t+1}} [\ln \phi_{x_{t+1}} - \ln 0.8 - \ln p_{a_t}(x_{t+1} = 1)] \\
&\quad + 0.2 * \phi_{x_{t+1}} [\ln \phi_{x_{t+1}} - \ln 0.2 - \ln p_{a_t}(x_{t+1} = 1)] \\
&\quad + 0.05 * (1 - \phi_{x_{t+1}}) [\ln(1 - \phi_{x_{t+1}}) - \ln 0.05 - \ln p_{a_t}(x_{t+1} = 2)] \\
&\quad + 0.95 * (1 - \phi_{x_{t+1}}) [\ln(1 - \phi_{x_{t+1}}) - \ln 0.95 - \ln p_{a_t}(x_{t+1} = 2)] \quad (4-14)
\end{aligned}$$

あとは生成モデルのうち  $a$  に依存する事前分布  $p_{a_t}(x_{t+1})$  の値を入れれば、 $G$  は  $a$  ごとに決まる  $\phi_{x_{t+1}}$  の関数になる。

$$\begin{aligned}
& G(\phi_{x_{t+1}}, a_t = 1) \\
&= 0.8 * \phi_{x_{t+1}} [\ln \phi_{x_{t+1}} - \ln 0.8 - \ln 0.8] \\
&\quad + 0.2 * \phi_{x_{t+1}} [\ln \phi_{x_{t+1}} - \ln 0.2 - \ln 0.8] \\
&\quad + 0.05 * (1 - \phi_{x_{t+1}}) [\ln(1 - \phi_{x_{t+1}}) - \ln 0.05 - \ln 0.2] \\
&\quad + 0.95 * (1 - \phi_{x_{t+1}}) [\ln(1 - \phi_{x_{t+1}}) - \ln 0.95 - \ln 0.2] \\
& G(\phi_{x_{t+1}}, a_t = 2) \\
&= 0.8 * \phi_{x_{t+1}} [\ln \phi_{x_{t+1}} - \ln 0.8 - \ln 0.2] \\
&\quad + 0.2 * \phi_{x_{t+1}} [\ln \phi_{x_{t+1}} - \ln 0.2 - \ln 0.2] \\
&\quad + 0.05 * (1 - \phi_{x_{t+1}}) [\ln(1 - \phi_{x_{t+1}}) - \ln 0.05 - \ln 0.8] \\
&\quad + 0.95 * (1 - \phi_{x_{t+1}}) [\ln(1 - \phi_{x_{t+1}}) - \ln 0.95 - \ln 0.8] \quad (4-15)
\end{aligned}$$

こうして計算できた  $G(a, \phi_{x_t})$  をプロットしたのが図4-10。

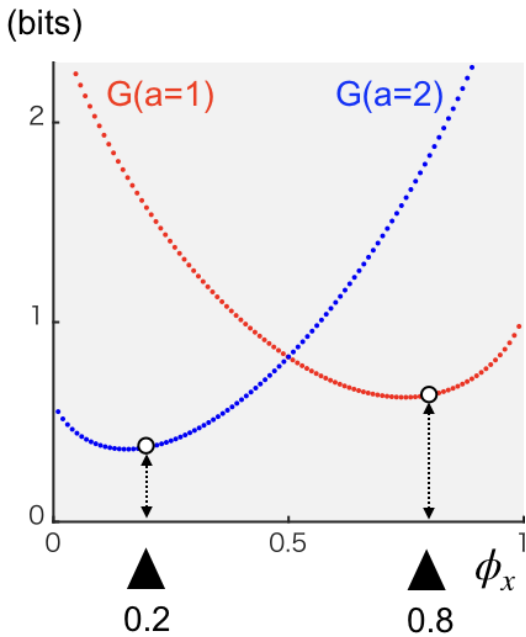


図4-10: 期待自由エネルギーGに基づく行動選択

図4-10では  $G(a = 1)$  および  $G(a = 2)$  を  $\phi_{x_t}$  を0-1の間で計算している。図4-5のときは  $s = 1, s = 2$  それぞれのときに  $\phi_{x_t}$  を動かして  $VFE$  を最小化することができたが、 $G$  の場合には  $s = 1, s = 2$  の両者を考慮しているのでそういうわけにはいかない。

そこでサブセクション4-4-2でも言及したように、 $\phi_{x_t}$  には、初期値である事前分布の  $p_{a_t=1}(x_{t+1} = 1) = 0.8$  および  $p_{a_t=2}(x_{t+1}) = 0.2$  を採用する近似が使われている。よくみるとじつは  $\phi_{x_t}$  を動かすと  $G$  をさらに最小化することができるが、この件については省略。

こうしてできた  $G(a)$  の最小値(0.6348 bits vs. 0.3735 bits)を比較して、より低い  $a = 2$  の方を行動選択として使う(図4-10)。正確にはsoftmax関数を使うことで、決定論的ではなくて確率論的により大きい方を行動選択する確率が増えるようにしてある。

これはスイッチをオフ  $a = 2$  にするほうが、照明オフ  $x = 2$  のときの観測モデル  $p(s_{t+1}|x_{t+1})$  の曖昧さによる不確定性を減らすことができるから、と説明することができるだろう。さらに詳しくは、後述のepistemic valueの計算が必要となってくる。

ところで現在使っている観測モデルは  $p(s_{t+1}|x_{t+1}) = \begin{bmatrix} 0.8 & 0.2 \\ 0.05 & 0.95 \end{bmatrix}$  と非対称性があった。これが行動選択の原因となっていた。もしこの観測モデルが  $\begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix}$  と対照的であった場合について  $G(a)$  をプロットすると図4-11の通りとなる。

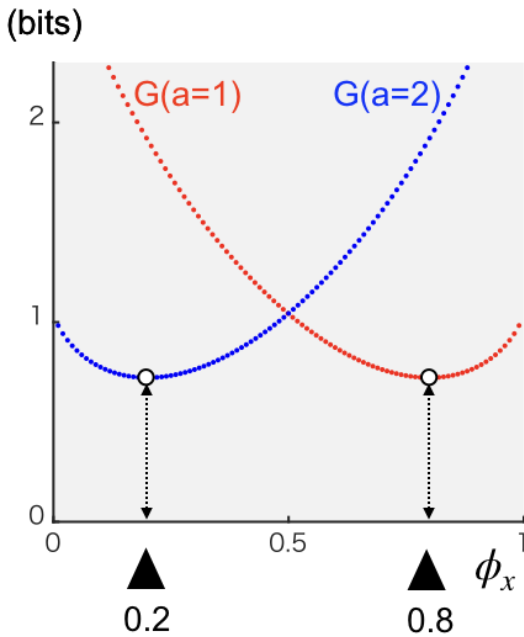


図4-11:  $G$ に基づく行動選択

$G(a)$  の最小値は  $a = 1, 2$  の間でまったく同じ  $0.7219$  bits になるので、どちらの行動を取りやすい、ということはない。これも予想通りだ。

#### [4-4-4. 期待自由エネルギー $EFE = G$ と変分自由エネルギー $VFE = F$ の関係]

サブセクション4-4-3の頭にも書いたように、 $G$  は二つの  $VFE$  の重み付き平均としては計算できない。このことを上記の照明の例において示しておく。

まず式(4-15)のように  $F(s_{t+1})$  を  $F'(x_{t+1}, s_{t+1}) = F'_{xs}$  という二つの項に分けておく。

$$\begin{aligned}
 F(\phi_{x_{t+1}}, s_{t+1} = 1, a_t) &= q(x_{t+1} = 1 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 1 | \phi_{x_{t+1}}) \\
 &\quad - \ln p_{a_t}(x_{t+1} = 1, s_{t+1} = 1)] \\
 &\quad + q(x_{t+1} = 2 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 2 | \phi_{x_{t+1}}) \\
 &\quad - \ln p_{a_t}(x_{t+1} = 2, s_{t+1} = 1)] \\
 &= F'_{11} + F'_{21} \\
 F(\phi_{x_{t+1}}, s_{t+1} = 2, a_t) &= q(x_{t+1} = 1 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 1 | \phi_{x_{t+1}}) \\
 &\quad - \ln p_{a_t}(x_{t+1} = 2, s_{t+1} = 1)] \\
 &\quad + q(x_{t+1} = 2 | \phi_{x_{t+1}}) [\ln q(x_{t+1} = 2 | \phi_{x_{t+1}}) \\
 &\quad - \ln p_{a_t}(x_{t+1} = 2, s_{t+1} = 2)] \\
 &= F'_{12} + F'_{22}
 \end{aligned} \tag{4-16}$$

と、 $G$  の式は以下のように表現できる。

$$\begin{aligned}
G(\phi_{x_{t+1}}, a_t) &= p(s_{t+1} = 1 | x_{t+1} = 1) F'_{11} \\
&\quad + p(s_{t+1} = 2 | x_{t+1} = 1) F'_{21} \\
&\quad + p(s_{t+1} = 1 | x_{t+1} = 2) F'_{12} \\
&\quad + p(s_{t+1} = 2 | x_{t+1} = 2) F'_{22}
\end{aligned} \tag{4-17}$$

よって、 $G$  は  $F$  の重み付き平均をただけではないことがわかる。つかもしそうだったら  $G$  には  $F$  に付加される情報がないので、行動選択のためにわざわざ  $G$  を定義する必要はなくなるだろうから、これは理にかなっている。

(検算しておく、式(4-17)からわかるように、観察モデルが完全に無情報なとき、つまり

$$p(s_{t+1} | x_{t+1}) = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \text{このときだけ } G \text{ は } F(s=1) \text{ と } F(s=2) \text{ の平均になっている。}$$

#### [4-4-5. epistemic value + extrinsic valueへの分解、の前に]

Neural Comput. 2017では、このEFE  $G$  の式を分解してepistemic valueと extrinsic valueの輪の形に変形している。式(4-12)から式変形してゆくが、ここで生成モデルの変形として  $p(x_{t+1}, s_{t+1}) = p(x_{t+1} | s_{t+1})p(s_{t+1})$  を使っている。これまで使っていた「観察モデル\*事前分布」による変形をしている。つまり  $F = DKL + Surprisal$  による変形だ。

(Epistemic valueが情報獲得を意味するならば、ここでの変形は  $p(x_{t+1}, s_{t+1}) = p(s_{t+1} | x_{t+1})p(x_{t+1})$  を用いた  $F = Bayesian\ surprise + Uncertainty$  になるはずなのだが。この件については後ほど再訪する。)

$$\begin{aligned}
G(\phi_{x_{t+1}}, a_t) &= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | \phi_{x_{t+1}}) \\
&\quad - \ln p_{a_t}(x_{t+1} | s_{t+1}) - \ln p(s_{t+1})]
\end{aligned} \tag{4-18}$$

ここで私が疑問なのが、式(4-18)の2行目でSurprisal  $\ln p(s_{t+1})$  は  $a_t$  に依存しないように書かれていることだ。しかし、

$$\begin{aligned}
p(s_{t+1}) &= \sum_{x_{t+1}} p_{a_t}(x_{t+1}, s_{t+1}) \\
&= \sum_{x_{t+1}} p(s_{t+1} | x_{t+1}) p_{a_t}(x_{t+1})
\end{aligned} \tag{4-19}$$

となっていることを考えると、 $p(s_{t+1})$  は  $p_{a_t}(s_{t+1})$  と書くべきではないか。

論より証拠、いまの照明の例で計算してみよう。

$$\begin{aligned}
p_{a_t=1}(s_{t+1} = [1 \ 2]) &= \sum_{x_{t+1}} p_{a_t=1}(x_{t+1}, s_{t+1}) \\
&= \sum_{x_{t+1}} p(s_{t+1}|x_{t+1})p_{a_t=1}(x_{t+1}) \\
&= p(s_{t+1}|x_{t+1} = 1)p_{a_t=1}(x_{t+1} = 1) \\
&\quad + p(s_{t+1}|x_{t+1} = 2)p_{a_t=1}(x_{t+1} = 2) \\
&= [0.8 \ 0.2] * 0.8 + [0.05 \ 0.95] * 0.2 \\
&= [0.65 \ 0.35] \\
p_{a_t=2}(s_{t+1} = [1 \ 2]) &= \sum_{x_{t+1}} p_{a_t=2}(x_{t+1}, s_{t+1}) \\
&= \sum_{x_{t+1}} p(s_{t+1}|x_{t+1})p_{a_t=2}(x_{t+1}) \\
&= p(s_{t+1}|x_{t+1} = 1)p_{a_t=2}(x_{t+1} = 1) \\
&\quad + p(s_{t+1}|x_{t+1} = 2)p_{a_t=2}(x_{t+1} = 2) \\
&= [0.8 \ 0.2] * 0.2 + [0.05 \ 0.95] * 0.8 \\
&= [0.20 \ 0.80] \tag{4-20}
\end{aligned}$$

というわけで、実例による計算から「常に  $a_t$  に依存してない」に対する反証は出せた。よってここは本来  $p_{a_t}(s_{t+1})$  として計算するべきであることが確認できた。

この点については、Gershman [arXiv 2019](#)およびKatahira et. al. [PsyArXiv 2019](#)は正しい計算をしている(Gershman論文では  $G$  の期待値に使っているものが違うけど)。これは些末な話ではなくて、この項をFristonのようにextrinsic valueに捉えるのか、それともKatahira et. al. 2019のようにpredicted surpriseとして捉えるのかで解釈にも大きな違いが生まれる。この点については以下のextrinsic valueの説明のところで再訪する。

なお、これは生成モデルの違いによっては説明できない。のちほど再訪する予定だけど、Katahira et. al. 2019での生成モデルはこの文書での蝶と蛾の例と同じ形で以下のようになるので、 $p_{a_t}(s_{t+1})$  が  $a_t$  に依存するという点については変わらない。

$$\begin{aligned}
p(s_{t+1}) &= \sum_{x_{t+1}} p_{a_t}(x_{t+1}, s_{t+1}) \\
&= \sum_{x_{t+1}} p_{a_t}(s_{t+1}|x_{t+1})p(x_{t+1}) \tag{4-21}
\end{aligned}$$

#### [4-4-6. Epistemic valueの導出]

ともあれ今は[Neural Comput. 2017](#)に準拠して、EFE  $G$  からepistemic value, extrinsic valueの導出までたどり着くことを目標にしよう。式(4-18)の2行目を再掲する。



$$\begin{aligned}
& G(\phi_{x_{t+1}}, \mathbf{a}_t) \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1} | s_{t+1}) \\
&\quad - \ln p(s_{t+1})] \tag{4-18}
\end{aligned}$$

まずは式(4-18)の第1項、第2項を式変形することでepistemic valueを導出する。 $\phi_{x_{t+1}}$  の表示は省略しておく。

$$\begin{aligned}
& \mathbb{E}_{q(x_{t+1}, s_{t+1})} [\ln q(x_{t+1}) - \ln p_{a_t}(x_{t+1} | s_{t+1})] \\
&\approx \mathbb{E}_{q(x_{t+1}, s_{t+1})} [\ln q(x_{t+1}) - \ln q(x_{t+1} | s_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1})} \mathbb{E}_{q(x_{t+1} | s_{t+1})} [\ln q(x_{t+1}) - \ln q(x_{t+1} | s_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1})} [\mathbb{E}_{q(x_{t+1} | s_{t+1})} [\ln q(x_{t+1}) - \ln q(x_{t+1} | s_{t+1})]] \\
&= -\mathbb{E}_{q(s_{t+1})} [\mathbb{E}_{q(x_{t+1} | s_{t+1})} [\ln q(x_{t+1} | s_{t+1}) - \ln q(x_{t+1})]] \\
&= -\mathbb{E}_{q(s_{t+1})} D_{KL}[q(x_{t+1} | s_{t+1}) || q(x_{t+1})] \\
&= -\text{Epistemic value}(\mathbf{a}_t) \tag{4-22}
\end{aligned}$$

1行目から2行目のところで推測  $q(x_{t+1} | s_{t+1})$  がtrue posterior  $p_{a_t}(x_{t+1} | s_{t+1})$  に充分近づけることができた、という近似を入れている。

このようにして、epistemic valueでは、 $q(x_{t+1} | s_{t+1})$  と  $q(x_{t+1})$  の間の距離を計算している。これは未来の感覚入力  $s_{t+1}$  によってどのくらい推測  $q$  が変化するか、つまり Bayesian surpriseを計算したうえで、それを全ての  $s_{t+1}$  で重み付け平均した期待値を計算している。これによって、行動 $\mathbf{a}_t$  が引き起こす未来の感覚入力  $s_{t+1}$  のばらつきを考慮した上でBayesian surpriseを計算していることになる。

ここで式(4-22)の一番最後の形を見てもらうと、これは  $q(x_{t+1} | s_{t+1})$  と  $q(x_{t+1})$  の間の相互情報量になっていることがわかる。

$$\begin{aligned}
& \text{Epistemic value}(\mathbf{a}_t) \\
&= \mathbb{E}_{q(s_{t+1})} [D_{KL}[q(x_{t+1} | s_{t+1}) || q(x_{t+1})]] \\
&= D_{KL}[q(x_{t+1} | s_{t+1}) || q(x_{t+1}) q(s_{t+1})] \\
&= MI(q(x_{t+1} | s_{t+1}); q(x_{t+1})) \tag{4-23}
\end{aligned}$$

つまり、epistemic valueとは、未来の感覚入力 $s_{t+1}$  を観測することによって推測  $q(x_{t+1})$  をどのくらいアップデートできるか、という情報獲得の大きさを計算していることになる。だから「epistemic valueが高い行動を選択する」とは「将来的に情報獲得の大きいことが見込まれる行動を選択する」という意味になる。こう考えると、epistemic valueが高い行動を選択するというのは理にかなっている。

(ここで使われている、「推測  $q(x_{t+1} | s_{t+1})$  がtrue posterior  $p_{a_t}(x_{t+1} | s_{t+1})$  に充分近づけることができた」という近似は、計算を可能にするための近似ではなくて、式(4-18)の第1,2項を情報獲得という解釈をするための読み替えに過ぎない。たとえばサブセクション4-4-2では、推測  $q(x_{t+1})$  を生成モデル(のうち事前

分布  $p(x)$  で近似したが、これは不可能な計算を可能にするための近似。でも今やってるのは逆。生成モデル  $p$  から作った true posterior を推測  $q$  で近似するというのだから。これは counter-factual な prediction においては、脳が  $q$  を変化させて true posterior と完全に一致させることができる、という仮定である。現在の例のように観察モデルに不確定な要素がある状況では(たぶん)これは成り立たない。) )

#### [4-4-7. Extrinsic valueの導出]

次に、式(4-18)の2行目の第3項を変形させることで extrinsic value を導出する。

$$\begin{aligned}
 & \mathbb{E}_{q(x_{t+1}, s_{t+1})} [-\ln p(s_{t+1})] \\
 &= \mathbb{E}_{q(s_{t+1})} [\mathbb{E}_{q(x_{t+1}|s_{t+1})} [-\ln p(s_{t+1})]] \\
 &= \mathbb{E}_{q(s_{t+1})} [\mathbb{E}_{q(x_{t+1}|s_{t+1})} [-\ln p(s_{t+1})]] \\
 &= -\mathbb{E}_{q(s_{t+1})} [\ln p(s_{t+1})] \\
 &= -\mathbb{E}_{q(s_{t+1})} [\ln \bar{p}(s_{t+1})] \\
 &= -\text{Extrinsic value}(a_t)
 \end{aligned} \tag{4-24}$$

ここで3行目から4行目のところで  $p(s_{t+1})$  が  $q(x_{t+1}|s_{t+1})$  には依存しない( $x_{t+1}$  ごとに足し合わせない)ということを使ってる。

4行目から5行目のところでは  $p(s_{t+1})$  を  $\bar{p}(s_{t+1})$  と違った書き方をしているが、これは Schwöbel et. al. 2018の表記を採用している。Schwöbel et. al. 2018のp.2537にあるように「将来得られるであろう感覚入力についての事前分布(=preference)  $\bar{p}(s_{t+1})$ 」と「将来得られるであろう感覚入力について周辺化した期待値  $p(s_{t+1}) = \sum_{x_{t+1}} p(x_{t+1}, s_{t+1})$ 」とはべつものである。Schwöbel et. al. 2018にしる、Friston et. al. 2017にしる、前者の preference の方を採用している。

しかしこれは恣意的な読み替えのように思える。というのも、前者の preference の意味であるなら、生成モデルの中に明示的にそのような事前分布を入れるべきだからだ。

ここで Parr and Friston [bioRxiv 2018](#)の図1に因子グラフがあるのでそれで確認してみると、期待自由エネルギーでは、期待自由エネルギー  $G$  の事前分布としてこの  $\bar{p}(s_{t+1})$  を入れている(bioRxiv 2018 図1A)。いっぽうで、この論文で提案されている Generalized free energy では、各時点の感覚入力(論文内では observation だが)  $s_{t-1}, s_t, s_{t+1}, \dots$  の全てに共通の事前分布としてこの  $\bar{p}(s_{t+1})$  が入っている(bioRxiv 2018 図1B)。つまり、Parr and Friston 2018では、期待自由エネルギー-EFEを計算する際に  $\bar{p}(s_{t+1})$  が生成モデルには入っていないこと、 $G$  を計算するためだけに使ったパラメーターであることをわかっている。

この文書で(そしてEFE関連の論文で)採用されている生成モデルでは、感覚入力  $s$  はそのつど  $x$  から作られるので、生成モデルの式の中に感覚入力についての prior  $p(s)$  というものはそもそも入ってない。よって、 $p(s_{t+1})$  は生成モデルから周辺化して作ったと考えるほうが自然で、事前分布=preferenceであると主張するには根拠が無いと思う。どうしてそこまでして  $\bar{p}(s_{t+1})$  を入れているのかといえば、素の  $G$  の式のままで extrinsic value 的なものを埋め込むことができないということだろう。でもまあ正直なところ、reward のような extrinsic value をこのような形で自由エネルギー原理の中に組み込むのは無理があるので、強化学習でやってるように明示的に reward を入れるべきじゃないの? と思った。

#### [4-4-8. 期待自由エネルギー EFEの式の最終形]

ともあれこれで期待自由エネルギー EFEの式が完成した。式(4-18),(4-23),(4-24)を組み合わせると、最終的な期待自由エネルギーの式になる。

$$\begin{aligned} G(a_t) &= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \\ &\approx \underbrace{-\mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} [D_{KL}[q(x_{t+1} | s_{t+1}) || q(x_{t+1})]}]}_{\text{Epistemic value}(a_t)} \\ &\quad - \underbrace{\mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} [\ln \bar{p}(s_{t+1})]}_{\text{Extrinsic value}(a_t)} \end{aligned} \tag{4-25}$$

つまり、epistemic valueが高いこと、extrinsic valueが高いこと、の両方が期待自由エネルギー EFE  $G$  を高くすることに貢献する。

#### [4-4-9. 照明の例を用いてepistemic valueとextrinsic valueを実感してみる]

図4-10、4-11で計算した $G(a = 1)$ 、 $G(a = 2)$ について、それぞれのepistemic valueとextrinsic valueを計算してみよう。

まずepistemic valueを計算してみる。前述したようにepistemic valueの近似は計算のための近似ではないので、近似前の式を使って計算する。なお、true posterior  $p_{a_t}(x_{t+1} | s_{t+1})$  は行動選択  $a_t$  に依存する。 $q(x_{t+1} | \phi_{x_{t+1}})$  のほうは近似(=  $p(x)$ )は使わずに  $\phi_{x_{t+1}}$  の関数の形のまま計算する。

$$\begin{aligned}
& \text{Epistemic value}(\mathbf{a}_t) \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln p_{a_t}(x_{t+1} | s_{t+1}) - \ln q(x_{t+1} | \phi_{x_{t+1}})] \\
&= q(x_{t+1} = 1, s_{t+1} = 1 | \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 1 | s_{t+1} = 1) \\
&\quad - \ln q(x_{t+1} = 1 | \phi_{x_{t+1}})] \\
&\quad + q(x_{t+1} = 1, s_{t+1} = 2 | \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 1 | s_{t+1} = 2) \\
&\quad - \ln q(x_{t+1} = 1 | \phi_{x_{t+1}})] \\
&\quad + q(x_{t+1} = 2, s_{t+1} = 1 | \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 2 | s_{t+1} = 1) \\
&\quad - \ln q(x_{t+1} = 2 | \phi_{x_{t+1}})] \\
&\quad + q(x_{t+1} = 2, s_{t+1} = 2 | \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 2 | s_{t+1} = 2) \\
&\quad - \ln q(x_{t+1} = 2 | \phi_{x_{t+1}})] \\
&= p(s_{t+1} = 1 | x_{t+1} = 1) \phi_{x_{t+1}} [\ln p_{a_t}(x_{t+1} = 1 | s_{t+1} = 1) - \ln \phi_{x_{t+1}}] \\
&\quad + p(s_{t+1} = 2 | x_{t+1} = 1) \phi_{x_{t+1}} [\ln p_{a_t}(x_{t+1} = 1 | s_{t+1} = 2) - \ln \phi_{x_{t+1}}] \\
&\quad + p(s_{t+1} = 1 | x_{t+1} = 2) (1 - \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 2 | s_{t+1} = 1) \\
&\quad - \ln(1 - \phi_{x_{t+1}})] \\
&\quad + p(s_{t+1} = 2 | x_{t+1} = 2) (1 - \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 2 | s_{t+1} = 2) \\
&\quad - \ln(1 - \phi_{x_{t+1}})] \\
&= 0.8 * \phi_{x_{t+1}} [\ln p_{a_t}(x_{t+1} = 1 | s_{t+1} = 1) - \ln \phi_{x_{t+1}}] \\
&\quad + 0.2 * \phi_{x_{t+1}} [\ln p_{a_t}(x_{t+1} = 1 | s_{t+1} = 2) - \ln \phi_{x_{t+1}}] \\
&\quad + 0.05 * (1 - \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 2 | s_{t+1} = 1) - \ln(1 - \phi_{x_{t+1}})] \\
&\quad + 0.95 * (1 - \phi_{x_{t+1}}) [\ln p_{a_t}(x_{t+1} = 2 | s_{t+1} = 2) - \ln(1 - \phi_{x_{t+1}})] \quad (4-26)
\end{aligned}$$

引き続き、個別の  $a_t$  について true posterior の値を入れて、Epistemic value を計算する。

$$\begin{aligned}
& \text{Epistemic value}(\mathbf{a}_t = 1) \\
&= 0.8 * \phi_{x_{t+1}} [\ln 0.9846 - \ln \phi_{x_{t+1}}] \\
&\quad + 0.2 * \phi_{x_{t+1}} [\ln 0.4571 - \ln \phi_{x_{t+1}}] \\
&\quad + 0.05 * (1 - \phi_{x_{t+1}}) [\ln 0.0153 - \ln(1 - \phi_{x_{t+1}})] \\
&\quad + 0.95 * (1 - \phi_{x_{t+1}}) [\ln 0.5428 - \ln(1 - \phi_{x_{t+1}})] \\
& \text{Epistemic value}(\mathbf{a}_t = 2) \\
&= 0.8 * \phi_{x_{t+1}} [\ln 0.8 - \ln \phi_{x_{t+1}}] \\
&\quad + 0.2 * \phi_{x_{t+1}} [\ln 0.05 - \ln \phi_{x_{t+1}}] \\
&\quad + 0.05 * (1 - \phi_{x_{t+1}}) [\ln 0.2 - \ln(1 - \phi_{x_{t+1}})] \\
&\quad + 0.95 * (1 - \phi_{x_{t+1}}) [\ln 0.95 - \ln(1 - \phi_{x_{t+1}})] \quad (4-27)
\end{aligned}$$

Extrinsic value の計算には  $\bar{p}(s_{t+1})$  による読み替えを使わず、式(4-18)の第3項のまま、行動選択  $a_t$  に依存する exact な計算を試みよう。

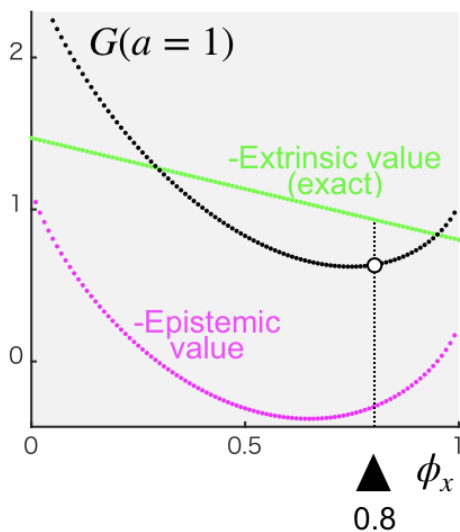
$$\begin{aligned}
&= \text{Extrinsic value}(\mathbf{a}_t) \\
&= \mathbb{E}_{q(s_{t+1}|\phi_{x_{t+1}})}[\ln p_{\mathbf{a}_t}(s_{t+1})] \\
&= \sum_{s_{t+1}} q(s_{t+1}|\phi_{x_{t+1}}) \ln p_{\mathbf{a}_t}(s_{t+1}) \\
&= q(s_{t+1} = 1|\phi_{x_{t+1}}) \ln p_{\mathbf{a}_t}(s_{t+1} = 1) \\
&\quad + q(s_{t+1} = 2|\phi_{x_{t+1}}) \ln p_{\mathbf{a}_t}(s_{t+1} = 2) \\
&= \phi_{x_{t+1}} \ln p_{\mathbf{a}_t}(s_{t+1} = 1) + (1 - \phi_{x_{t+1}}) \ln p_{\mathbf{a}_t}(s_{t+1} = 2) \tag{4-28}
\end{aligned}$$

引き続き、個別の  $\mathbf{a}_t$  について  $p_{\mathbf{a}_t}(s_{t+1})$  の値(式(4-20)で計算済み)を入れて、Extrinsic valueを計算する。

$$\begin{aligned}
&= \text{Extrinsic value}(\mathbf{a}_t = 1) \\
&= \phi_{x_{t+1}} \ln p_{\mathbf{a}_t=1}(s_{t+1} = 1) + (1 - \phi_{x_{t+1}}) \ln p_{\mathbf{a}_t=1}(s_{t+1} = 2) \\
&= \phi_{x_{t+1}} \ln 0.65 + (1 - \phi_{x_{t+1}}) \ln 0.35 \\
&= \text{Extrinsic value}(\mathbf{a}_t = 2) \\
&= \phi_{x_{t+1}} \ln p_{\mathbf{a}_t=2}(s_{t+1} = 1) + (1 - \phi_{x_{t+1}}) \ln p_{\mathbf{a}_t=2}(s_{t+1} = 2) \\
&= \phi_{x_{t+1}} \ln 0.20 + (1 - \phi_{x_{t+1}}) \ln 0.80 \tag{4-29}
\end{aligned}$$

これでEpistemic valueとexactなExtrinsic valueをプロットできる(図4-12)。Exactな値なので、 $G = -\text{Epistemic value} - \text{Extrinsic value}$ が成り立つ。図でも正負逆になっているので注意。つまり、それぞれ値が低いほうがvalueが高い。Epistemic valueについては前述のとおり、近似は計算結果を変えない。

(A)  $G(a = 1)$   
(bits)



(B)  $G(a = 2)$   
(bits)

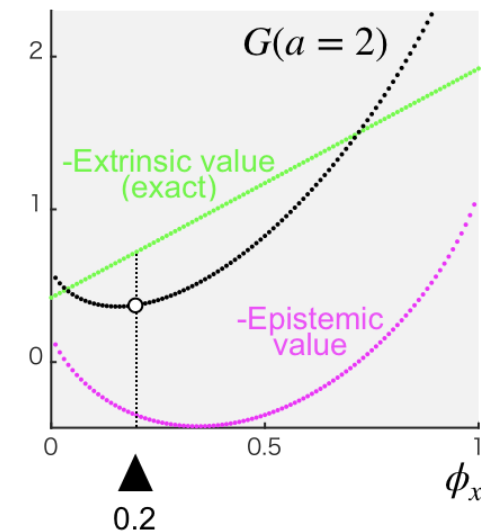


図4-12: Exactな計算によるEpistemic valueとExtrinsic value

Epistemic valueおよびextrinsic valueそれぞれが行動選択  $a_t$  に依存していることがわかる。Epistemic value ( $a_t = 1$ ) よりもEpistemic value ( $a_t = 2$ ) のほうが僅かに最小値が小さいことから、Epistemic valueがたんなる左右反転でないことがわかる。

$q(x_{t+1}) = p(x)$  の近似のもとでは、行動  $a_t = 1$  を選択したときには  $\phi_{x_{t+1}} = 0.8$ 、行動  $a_t = 2$  を選択したときには  $\phi_{x_{t+1}} = 0.2$  なので、以下のようにまとめることができる。

—	$a = 1$	$a = 2$
$\phi_{x_t}$	0.8	0.2
-Epistemic value	0.2992	0.3484
-Extrinsic value	0.9341	0.7219
G(a)	0.6348	0.3735

テーブル2: Exactな計算によるEpistemic valueとExtrinsic value

つまり、epistemic valueとextrinsic valueの効果は 行動  $a_t$  によって逆転しているのだけど、両方の和の結果として、期待自由エネルギー  $G$  は  $a = 2$  のときのほうが低くなるので、行動選択としては  $a = 2$  を選ぶことになる。

ではつぎに、Fristonの定義に基づいて preferenceとして読み替えたときのextrinsic valueを計算してみよう。ここでpreferenceを以下のように設定する。

$$\bar{p}(s = [1 \ 2]) = [0.75 \ 0.25]$$

つまり、いま「部屋を75%の確率で明るくあってほしい(100%明るいのも良くない)」というpreferenceをこれまでの生成モデルに設定として付け加えたということになる。これは行動  $a$  には依存しない。

(こう書いて気づいたけど、行動  $a$  に依存するpreferenceがあってもいいよね？つまり、「スイッチを付けたら明るくなってほしいし、スイッチを消したら暗くなってほしい」というpreference。そしてこれが、行動の結果の予測可能性の最大化であり、FEPが前提としているものだ。つまり、明示的なpreference=ゴールがないときにもFEPはpreferenceを持っている。)

つぎに  $q(s_{t+1})$  を計算してみよう。式(4-13)を  $x_{t+1}$  で足し合わせてしまえばいい。こちらでも行動  $a$  に依存しないことがわかる。あくまでも  $q(s_{t+1})$  は  $\phi_{x_{t+1}}$  の関数であって、行動  $a$  によって  $\phi_{x_{t+1}}$  が決まる。(行動  $a = 1$  のときは  $\phi_{x_{t+1}} = 0.8$  となるのだった。図4-10,4-11参照。)



$$\begin{aligned}
q(s_{t+1}|\phi_{x_{t+1}}) &= \sum_{x_{t+1}} q(x_{t+1}, s_{t+1}|\phi_{x_{t+1}}) \\
&= \sum_{x_{t+1}} p(s_{t+1}|x_{t+1})q(x_{t+1}|\phi_{x_{t+1}}) \\
q(s_{t+1} = 1|\phi_{x_{t+1}}) &= p(s_{t+1} = 1|x_{t+1} = 1)\phi_{x_{t+1}} \\
&\quad + p(s_{t+1} = 1|x_{t+1} = 2)(1 - \phi_{x_{t+1}}) \\
&= 0.8 * \phi_{x_{t+1}} + 0.05(1 - \phi_{x_{t+1}}) \\
&= 0.75 * \phi_{x_{t+1}} + 0.05 \\
q(s_{t+1} = 2|\phi_{x_{t+1}}) &= p(s_{t+1} = 2|x_{t+1} = 1)\phi_{x_{t+1}} \\
&\quad + p(s_{t+1} = 2|x_{t+1} = 2)(1 - \phi_{x_{t+1}}) \\
&= 0.2 * \phi_{x_{t+1}} + 0.95(1 - \phi_{x_{t+1}}) \\
&= -0.75 * \phi_{x_{t+1}} + 0.95
\end{aligned} \tag{4-30}$$

これでextrinsic valueを  $\phi_{x_{t+1}}$  の関数として書くことができる。

$$\begin{aligned}
\text{Extrinsic value}(\phi_{x_{t+1}}) &= \sum_{s_{t+1}} q(s_{t+1}|\phi_{x_{t+1}}) \ln \bar{p}(s_{t+1}) \\
&= q(s_{t+1} = 1|\phi_{x_{t+1}}) \ln \bar{p}(s_{t+1} = 1) \\
&\quad + q(s_{t+1} = 2|\phi_{x_{t+1}}) \ln \bar{p}(s_{t+1} = 2) \\
&= q(s_{t+1} = 1|\phi_{x_{t+1}}) \ln 0.75 \\
&\quad + q(s_{t+1} = 2|\phi_{x_{t+1}}) \ln 0.25 \\
&= (0.75 * \phi_{x_{t+1}} + 0.05) \ln 0.75 \\
&\quad + (-0.75 * \phi_{x_{t+1}} + 0.95) \ln 0.25
\end{aligned} \tag{4-31}$$

これのようにしてできた-extrinsic value(preferenceバージョン)から、Friston et al 2017での定義に基づいた  $G(a) = -\text{epistemic value} - \text{extrinsic value}$  をプロットしてみる(図4-13)。

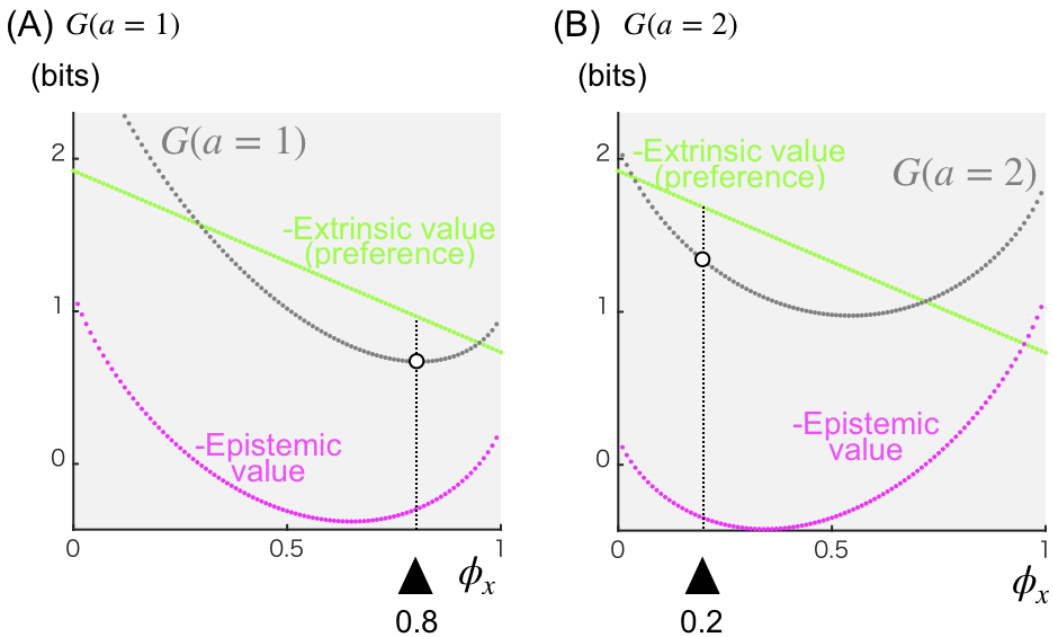


図4-13: FristonのEpistemic valueとExtrinsic value

前述したように、preferenceに基づいたextrinsic valueは  $a_t$  には依存していないので、図4-13A,Bともに同じ線になる("-Extrinsic value (preference)", 薄緑色)。そして  $q(x_{t+1}) = p(x)$  の近似のもとでは、行動  $a_t = 1$  を選択したときには  $\phi_{x_{t+1}} = 0.8$ 、行動  $a_t = 2$  を選択したときには  $\phi_{x_{t+1}} = 0.2$  なので、以下のようにまとめることができる。

—	$a = 1$	$a = 2$
$\phi_{x_t}$	0.8	0.2
-Epistemic value	0.2992	0.3484
-Extrinsic value	0.9698	1.6830
G(a)	0.6705	1.3346

テーブル3: FristonのEpistemic valueとExtrinsic value

つまり、行動  $a_t = 1$  (スイッチを付ける) のほうが-extrinsic valueは低くなる。これは preferenceに合致した未来の感覚入力  $s$  をサンプルするであろう行動  $a_1$  がよりextrinsic valueが高い、という理屈だ。

その結果として期待自由エネルギー  $G$  は  $a = 1$  のときのほうが低くなるので、行動選択としては  $a = 1$  を選ぶことになる。図4-12のときは  $a = 2$  を選んでいたのので、結果が逆

になった。Extrinsic valueをどう定義するかが行動選択に重大な違いを生むということがわかる。

図4-13を見ていて気になるのがなんでextrinsic valueは直線であって、曲線でないかということだ。式(4-31)を見ればわかるようにextrinsic valueは  $\phi_{x_{t+1}}$  の一次関数だ。この性質はpreferenceの値には依存してない。Extrinsic valueはその定義上、preference  $\bar{p}$  と  $q(s_{t+1})$  が一致するときに最大になっていてほしいのだけど、そうになってない。代わりにslopeの大きさがvalueの大きさ (utility) と比例すると解釈することになる。でもそうすると、このやり方だと複数のゴールを設定することはできない。

#### 4-4-10. Expected free energyにかんする疑問、 $G$ の解釈と計算についての代案

Expected free energyの説明は以上だが、これで納得行っただろうか？ 私はepistemic valueの導出を見ていて気づいたことがある。式(4-22)では、 $q(x_{t+1})$  が観測データ  $s_{t+1}$  無しでの推測を意味することになっている。しかしこれまでの例、例えば図3-2の行動無しの設定においては、観測データ  $s_{t+1}$  を得た上での推測=approximate posteriorは  $q(x_{t+1})$  と書いてきた。  $q(x_{t+1}|s_{t+1})$  ではなくて。そしてそもそも式(4-18)での  $G$  の式の第1項が  $\ln q(x_{t+1})$  になっているのはこれがapproximate posteriorだったからだ。事前分布で近似できる  $q(x_{t+1}) = q(x_t) = p(x)$  ではなくて。つまりFriston et al 2017では、知らず知らずのうちに  $q(x_{t+1})$  の読み替え(誤読)を行っていないだろうか？

EFEについて理解しようとするたびに毎回混乱するのだが、その理由はここではないかと私は思う。それならば、はじめからapproximate posteriorのことを  $q(x_{t+1}|s_{t+1})$  と書くことにして、式(4-18)も式(4-32)のように表示するのが筋だろう。

$$\begin{aligned} G(\phi_{x_{t+1}}, a_t) \\ = \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}, s_{t+1})] \end{aligned} \quad (4-32)$$

そしてさらにこれの生成モデルの部分を「観察モデル\*事前分布」とする変形  $p_{a_t}(x_{t+1}, s_{t+1}) = p_{a_t}(x_{t+1} | s_{t+1}) p_{a_t}(s_{t+1})$  を用いて変形してゆくと、

$$\begin{aligned}
& G(\phi_{x_{t+1}(s_{t+1})}, a_t) \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}(s_{t+1})})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})}) \\
&\quad - \ln p_{a_t}(x_{t+1} | s_{t+1}) - \ln p_{a_t}(s_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}(s_{t+1})})} \mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})}) \\
&\quad - \ln p_{a_t}(x_{t+1} | s_{t+1}) - \ln p_{a_t}(s_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}(s_{t+1})})} [\mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})}) \\
&\quad - \ln p_{a_t}(x_{t+1} | s_{t+1})] - \mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})})} [\ln p_{a_t}(s_{t+1})]] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}(s_{t+1})})} [D_{KL}[q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})}) || p_{a_t}(x_{t+1} | s_{t+1})] \\
&\quad - \ln p_{a_t}(s_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}(s_{t+1})})} \underbrace{D_{KL}[q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}(s_{t+1})}) || p_{a_t}(x_{t+1} | s_{t+1})]}_{KLD} \\
&\quad - \underbrace{\mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}(s_{t+1})})} \ln p_{a_t}(s_{t+1})}_{Surprisal} \tag{4-33}
\end{aligned}$$

元々のVFEの式でのKLDとsurprisalの足し算になっている。こちらのほうが変分自由エネルギーを最小化する過程と整合的になっていると思う。そもそも  $G$  の式の第1,2項は、もともとのVFEの式ではKLDに対応していたのだから、ここを変形したら情報獲得になった、というのがおかしい話だった。

なお、ここで  $\phi_{x_{t+1}}$  はそれぞれの  $s_{t+1}$  に合わせて決めてやる必要があるので、 $\phi_{x_{t+1}(s_{t+1})}$  と表記している。つまり、counter-factualな推測  $q$  としては、すべての可能な  $s_{t+1}$  についての  $x_{t+1}$  の推測をしている。

それならば、生成モデルの部分を「true posterior\*evideence」とする変形  $p_{a_t}(x_{t+1}, s_{t+1}) = p(s_{t+1} | x_{t+1}) p_{a_t}(x_{t+1})$  を用いて変形していけば、 $F = \text{Bayesian surprise} + \text{Uncertainty}$  に対応したものになるはずだ。やってみよう。

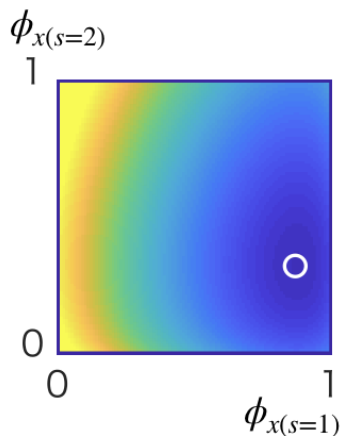
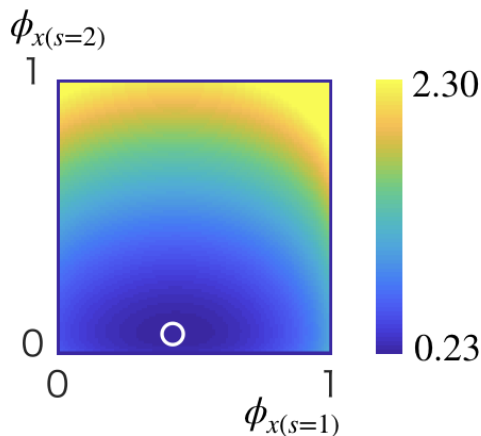
$$\begin{aligned}
& G(\phi_{x_{t+1}}, \mathbf{a}_t) \\
&= \mathbb{E}_{q(x_{t+1}, s_{t+1} | \phi_{x_{t+1}})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}) \\
&\quad - \ln p(s_{t+1} | x_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} \mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1}) \\
&\quad - \ln p(s_{t+1} | x_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} [\mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}})} [\ln q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}}) - \ln p_{a_t}(x_{t+1})] \\
&\quad - \mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}})} [\ln p(s_{t+1} | x_{t+1})]] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} [D_{KL}[q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}}) || p_{a_t}(x_{t+1})] \\
&\quad - \mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}})} \ln p(s_{t+1} | x_{t+1})] \\
&= \mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} \underbrace{D_{KL}[q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}}) || p_{a_t}(x_{t+1})]}_{\text{Bayesian Surprise}(a_t, s_{t+1}, \phi_{x_{t+1}})} \\
&\quad + \underbrace{\mathbb{E}_{q(s_{t+1} | \phi_{x_{t+1}})} [-\mathbb{E}_{q(x_{t+1} | s_{t+1}, \phi_{x_{t+1}})} \ln p(s_{t+1} | x_{t+1})]}_{\text{Uncertainty}(s_{t+1}, \phi_{x_{t+1}})} \tag{4-34}
\end{aligned}$$

こちらのほうが、bayesian surpriseを元にして情報獲得を計算しているという意味で理にかなっている。そして、そもそも元のVFEの式では、agentはbayesian surpriseを最大化していない。あくまでもKLDを最小化するように推測  $q = \text{approximate posterior}$  をアップデートするのが先で、その結果としてbayesian surpriseが  $> 0$  となるだけだ(サブセクション1-9を参照)。だから、 $G$  がepistemic value(とextrinsic valueの和)を最大化するように行動 $a$  を選択するというのは、agentが変分推定をするという元々の考え方からすればおかしな話なのだ。

あとは式(4-34)から言えることとして、Uncertaintyの項は行動選択  $a_t$  に依存しないということも重要。よってbayesian surpriseの項だけを比べれば、どちらの行動選択  $a_t$  のほうが  $G$  が低くできるかがわかる。Friston et al 2017でのepistemic valueは実質的にはこのbayesian surpriseの項と同じになっている( $q(x_t) = p(x)$  の近似をした場合)。ゆえにFriston et al 2017での期待自由エネルギー $G$  は式(4-34)からUncertaintyの項は捨てて、代わりに  $-\ln \bar{p}(s_{t+1})$  の項を足したものと解釈できる。だから、途中はおかしいけど結果としてはうまく回っているということではないだろうか？

(あと、Bayesian surpriseはその定義より、事前分布  $p_{a_t}(x_{t+1})$  からの情報獲得となっているので、 $t = 0$  からの総獲得情報量を表すことになる。その時点での情報獲得を計算するためには、 $t - 1, t, t + 1$  と時間ごとにBayesian surpriseを計算したうえで、その差分として計算されることになる。いまは生成モデルは固定なので、 $p_{a_{t-2}}(x_{t-1}) = p_{a_{t-1}}(x_t) = p_{a_t}(x_{t+1})$  であり、ただの定数。)

さてそれでは、このように解釈した  $G(a)$  をプロットしてみる(図4-14)。 $\phi_{x_{t+1}(s_{t+1})}$  は  $s = 1, 2$  についてそれぞれあるので、二次元のデータをカラー表示することにする。

(A)  $G(a = 1)$ (B)  $G(a = 2)$ 図4-14: 改良版期待自由エネルギー  $G$ 

白丸が行動  $a = 1$  および  $a = 2$  それぞれの期待自由エネルギー  $G$  が最小になる場所を表している。行動  $a = 1$  では  $\phi_{x_{t+1}(s_{t+1}=1)} = q(x_{t+1} | s_{t+1} = 1) = 0.88$  かつ

$\phi_{x_{t+1}(s_{t+1}=2)} = q(x_{t+1} | s_{t+1} = 2) = 0.33$  のときに最小となる。同様に行動  $a = 2$  では  $\phi_{x_{t+1}(s_{t+1}=1)} = q(x_{t+1} | s_{t+1} = 1) = 0.43$  かつ

$\phi_{x_{t+1}(s_{t+1}=2)} = q(x_{t+1} | s_{t+1} = 2) = 0.08$  のときに最小となる。

これらの値は過去の行動選択の例(テーブル1)のときと似ているがやや違う。同じにならないのはサブセクション4-4-4にあるように、まったく同じ計算をしているわけではないから。

期待自由エネルギー  $G$  は行動  $a = 1$  (スイッチオン)のときよりも行動  $a = 2$  (スイッチオフ)のときにより低くなる。よってagentは行動  $a = 2$  (スイッチオフ)を選択する、ということになる。

—	$a = 1$	$a = 2$
$\phi_{x_{t+1}(s_{t+1}=1)}$	0.88	0.43
$\phi_{x_{t+1}(s_{t+1}=2)}$	0.33	0.08
$G(a)$	0.3082	0.2140

テーブル4: 改良版期待自由エネルギー  $G$ 

同じデータについてKLD, surprisal, bayesian surprise, uncertaintyについて2次元表示をすることができる。



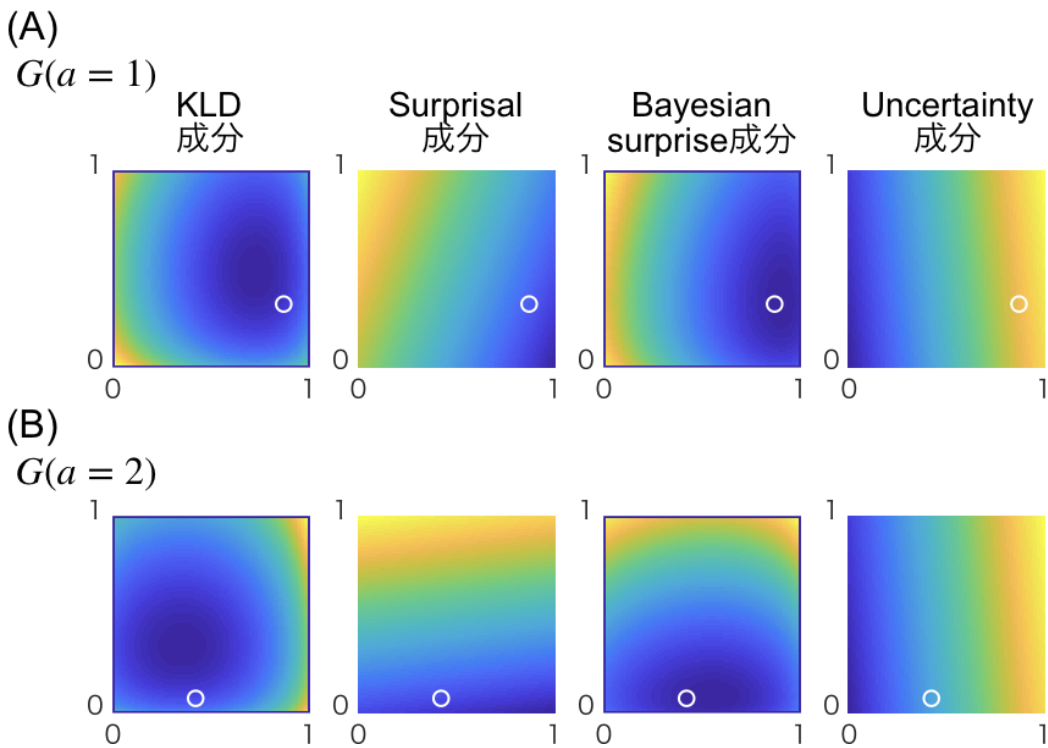


図4-15: 改良版期待自由エネルギー  $G$ での各要素

これを見ると、 $G$  が最小になるところでKLD成分が0になるわけではないことがわかる。一方で $G$  が最小になるところではBayesian surprise成分がほぼ最小になっている。数値的に確認してみると、完全に一致しているわけではないが、これは上述のとおり、Uncertainty成分が $a$  に依存しないため、Bayesian surprise成分の最小値によってどちらの $G$  の最小値がより低いかが決まるということを考えれば納得がいく。

そういうわけで、改良版期待自由エネルギー  $G$  は変分推定をするという元々の考え方と合致していて、ちゃんと行動選択にも使えそうだ。

結論としては、Friston et al 2017の期待自由エネルギー  $G$  ってぜんぜん変分推定してないよね、ってことになる。じゃあちゃんと変分推定しようよってのが、Schwöbel et al 2018とParr and Friston 2018が行っていることだと私は解釈している。そういうわけで、次はSchwöbel et al 2018の解説に行く。

## [5. 現在、過去、未来を統一的に捉えるモデル]

### [5-1. Schwöbelらのアプローチ]

Friston et al 2017の期待自由エネルギー  $G$  ってぜんぜん変分推定してないよね、ってのが前のサブセクションの総評だったわけだけど、じゃあちゃんと変分推定しようよってのが、Schwöbel et al 2018とParr and Friston 2018。ここではSchwöbel et al 2018の解説を行う。

## [5-2. 生成モデル $p$ と推測 $q$ の設定]

Schwöbel et. al. [Neural Comput. 2018](#)では、現在、過去、未来を統一的に捉える生成モデルを考えている。現在、過去、未来を統一的に捉えるというのはどういうことかという、時刻  $1 - T$  というある時間幅(これはなんらかの行動課題の1試行のようなひとかたまりの行動を想定している)のなかでagentがいま時刻  $t$  にいるという状況を考える。このときの因果グラフが図5-1だ。

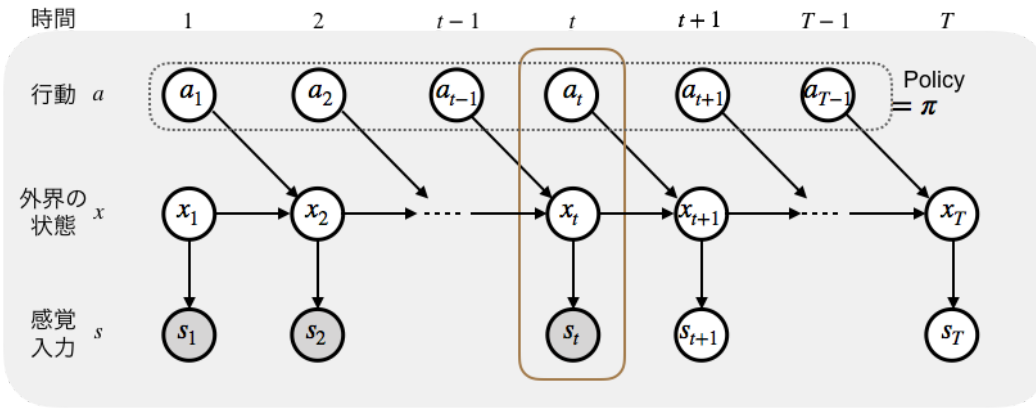


図5-1: 現在、過去、未来の生成モデル

これまでと同様に、POMDP過程を前提としているので、直前の行動のみが外界の状態に影響を及ぼしうる。なによりも最大の違いは、行動  $a$  がパラメーターではなく潜在変数になっているという点だ。変分自由エネルギー  $VFE$  を定義するために観測データと潜在変数をリストにしてみよう：

- 観測データ
  - 感覚入力(過去から現在まで)  $\underline{s} = \{s_1, s_2, \dots, s_{t-1}, s_t\}$
- 潜在変数
  - 外界の状態(過去から現在まで)  $\underline{x} = \{x_1, x_2, \dots, x_{t-1}, x_t\}$
  - 外界の状態(未来)  $\tilde{x} = \{x_{t+1}, \dots, x_{T-1}, x_T\}$
  - 感覚入力(未来)  $\tilde{s} = \{s_{t+1}, \dots, s_{T-1}, s_T\}$
  - 行動選択(過去)  $\underline{a} = \{a_1, a_2, \dots, a_{t-1}\}$
  - 行動選択(現在から未来まで)  $\tilde{a} = \{a_t, \dots, a_{T-1}, a_T\}$

よってこのときの生成モデル  $p$  はこれまで同様、機械的に当てはめれば以下のように書ける。

$$p(\underline{x}, \tilde{x}, \underline{s}, \tilde{s}, \underline{a}, \tilde{a}) \quad (5-1)$$

また、これを推定する  $q$  のほうは潜在変数だけが入る。  $q$  の形を決めるパラメーター  $\phi$  はこれまでの議論からわかるように  $x$  についての推定だけを決めればあとは生成モデルを使って計算してゆくので、今回の場合、  $\phi_{\underline{x}, \tilde{x}}$  となる。省略して  $\tilde{\phi}$  と表示する。

$$\begin{aligned}
& q(\underline{x}, \tilde{x}, \tilde{s}, \underline{a}, \tilde{a} | \phi_{\underline{x}, \tilde{x}}) \\
& = q(\underline{x}, \tilde{x}, \tilde{s}, \underline{a}, \tilde{a} | \tilde{\phi})
\end{aligned} \tag{5-2}$$

### [5-3. 変分自由エネルギー $F$ の設定]

あとはこの二つから現在、過去、未来全ての期間をひとまとめにして推定をするための変分自由エネルギー  $F$  が定義できる。 $q$  が長いので  $\tilde{q}$  で表記する。

$$\begin{aligned}
F(\tilde{\phi}) & = \mathbb{E}_{\tilde{q}} [\ln \tilde{q} - \ln p(\underline{x}, \tilde{x}, \underline{s}, \tilde{s}, \underline{a}, \tilde{a})] \\
& = \underbrace{D_{KL}(\tilde{q} \| p(\underline{x}, \tilde{x}, \underline{s}, \tilde{s}, \underline{a}, \tilde{a} | \underline{s}))}_{\text{KLD}(\tilde{\phi})} + \underbrace{-\ln p(\underline{s})}_{\text{Surprisal}}
\end{aligned} \tag{5-3}$$

長いけど、たんに観測値と隠れ値を分けてVFEの式に代入しているだけだから、式の定義は明確だ。あとこうして改めて式を見てみると、surprisalの中身は常に観測データであり、surprisalは  $F$  を変えようがない定数であるということがわかる。(だから、暗い部屋問題など無いわけ。)

じつは式(5-3)は正確でない。Schwöbel et. al. 2018では生成モデルの方に前述の、未来の感覚入力についての事前分布(=preference)  $\bar{p}(\tilde{s})$  が掛け算されている。私はどうにも納得いかないけど。ともあれ論文の中での  $F$  の式は以下の通りになる。

$$F(\tilde{\phi}) = \mathbb{E}_{\tilde{q}} [\ln \tilde{q} - \ln p(\underline{x}, \tilde{x}, \underline{s}, \tilde{s}, \underline{a}, \tilde{a}) - \ln \bar{p}(\tilde{s})] \tag{5-4}$$

この方法がFriston et al 2017での期待自由エネルギーと比べてなにが素敵かというと、VFEに期待値を掛けるとかそういうややこしいことを言わずに、ストレートにすべての潜在変数を推定するための変分下限としての変分自由エネルギー  $VFE$  を計算できている点。もし  $KLD = 0$  になるように  $q$  のパラメーター  $\tilde{\phi}$  をアップデートすることができれば、 $q$  はtrue posteriorを近似することができて、そのとき  $F = \text{Surprisal}$  となる、これまでとまったく同じ。

このため、期待自由エネルギーのときにでてきたややこしい概念、counter-factualな観測データというものがここでは見えなくなっている。

### [5-4. observed free energy $V$ とpredicted free energy $G$ への分解]

Schwöbel et. al. 2018ではさらにこの  $F$  の式を[現在+過去]と[現在の行動計画および未来]とに分けている。彼らは前者をobserved free energy  $V$ 、後者をpredicted free energy  $G$  と呼んでいる(p.2539-2540)。

$$V(\tilde{\phi}) = \mathbb{E}_{q(\underline{x}, \underline{a} | \tilde{\phi})} [\ln q(\underline{x}, \underline{a} | \tilde{\phi}) - \ln p(\underline{x}, \underline{s}, \underline{a})] \quad (5-5)$$

$$G(\tilde{\phi}) = \mathbb{E}_{q(\underline{x}, \tilde{s}, \tilde{a} | \tilde{\phi})} [\ln q(\tilde{x}, \tilde{s} | \underline{x}, \underline{a}, \tilde{\phi}) - \ln p(\tilde{x}, \tilde{s} | \underline{x}, \underline{a}, \tilde{a})] \\ - \mathbb{E}_{q(\tilde{s})} \ln \bar{p}(\tilde{s}) \quad (5-6)$$

式(5-4)からの変形でこれが導き出せるのだと思うのだけど、当の論文には何も説明がない。私もまだ確認できてない。ともあれこれによって、これまでの論文で出てきた変分自由エネルギーや期待自由エネルギーとの対応付けができるようになった。

Observed free energy  $V$  はこの文書でも、4-2-1において「過去の行動を付加した生成モデル」に対応したVFEとして式(4-3)ですでに定式化している。

後者のpredicted free energy  $G$  がこれまでの期待自由エネルギー  $G$  に対応するようになってる。だからこちらにはpreferenceの項を付け加えている。

### [5-5. おまけ：このモデルの含意]

おまけです。FEP入門的にはおまけだけど、ここからがワタシ的には本筋。個人的にこのモデルが面白いと思うのは、現在、過去、未来全ての期間をひとまとめにして取り扱うというのが、まさに変分原理的な視点であるという点だ。つまり、agentはいま時刻  $t$  に生死をかけあらゆる意思決定をなしているつもりなのだけれども、いったんそれを時間の外から(スピノザの言う「永遠の相」から)眺めてやれば、与えられた状況の中で  $VFE$  を下げる方向に進むという変分原理に従って、agentの内部状態(推測  $q$  の本体)が変動しているだけなのだ。(agent「が」内部状態を変動「させている」ではないことに注意。) あたかも光が屈折することで最短時間で進むルートを選んでいるかのように。

いっぽうで、このVFEの式をいったん過去と未来に分けてやると、とたんにcounter-factual predictionを考えないといけない内部の視点が出てくる。ここでは未来と過去は観測データの有無の違いという点で非対称性がある。このことについてこれまで使ってきた図を改変してイメージを膨らませてみよう(図5-2)。

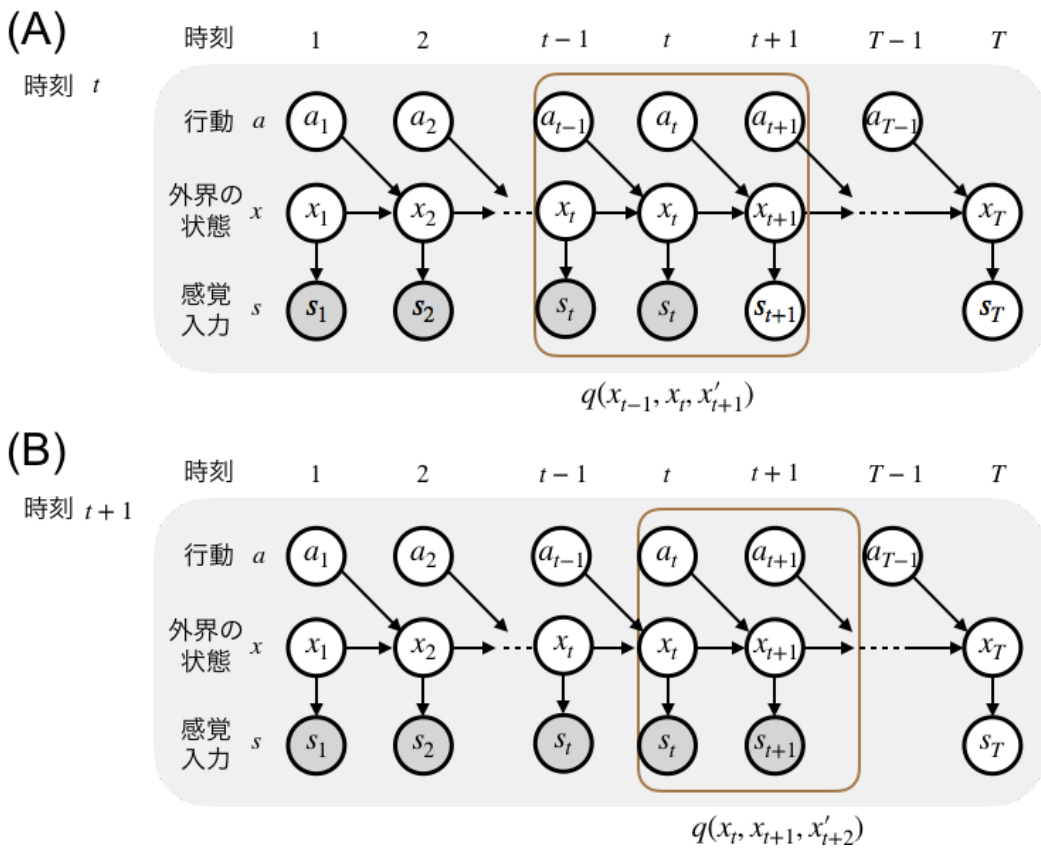


図5-2: 推測 $q$ の時間幅

図5-2Aでは時刻  $t$  における変分自由エネルギー  $F$  の計算からすべての潜在変数が推定される。時刻  $t+1$  になると(図5-2B)、時刻  $t+1$  での感覚入力  $s(t+1)$  を獲得して観測データが一つ増えたことによって、変分自由エネルギー  $F$  がアップデートされ、すべての潜在変数の推定もアップデートされる。これが外側から、「永遠の相」から見たときの描写だ。

いっぽうでagentから見れば、図5-2Aでは時刻  $t$  では未来の外界の状態  $x_{t+1}$  についてのcounter-factualな推定を持っている(counter-factualであることを  $x'_{t+1}$  と表記しておく)。さらにagentは過去の外界の状態  $x_{t-1}$  についてのfactualな推定を持っている。この推定は現在の新しい感覚入力  $s_t$  によって時間を遡ってアップデートされること(post-diction)については3-3-3で言及した。

そして時刻  $t+1$  になると、未来の外界の状態  $x_{t+2}$  についてのcounter-factualな推定を持つようになり、かつてcounter-factualだった外界の状態についての推定  $q(x'_{t+1})$  はfactualな推定  $q(x_{t+1})$  となった。このようなcounter-factualな推定とfactualな推定の違いは時刻 $t$  にいるagentにとってはそのつど行動 $a$  を決める際に初めて重要になるが、変分原理的な視点からは見えなくなっている(ように思える)。

[Friston 2018 \(Frontiers in Psychology\)](#)における意識の議論でFristonは、agentが変分推定をしてゆく過程でのcounter-factualなpredictionおよび過去に向けてのpostdictionの時間幅が長く深い(temporal thickness or counterfactual depth)ものに意識が宿るのだという言い方をしている。

式(5-2)でも書いたように、現在の設定でagentが推測  $q$  をアップデートさせるために動かしているものの実体とは、 $\phi_{x, \tilde{x}}$  つまり、過去から未来までの外界の状態  $x$  の推定  $q(x)$  だけだ。よってFristonの言うtemporal thicknessというのはどのくらい未来と過去について推測  $q(x)$  を持っているかということに言い換えることができる。推測  $q$  の時間推移についてフッサールの内的時間意識のスキームに乗っけて表示してみよう(図5-3)。

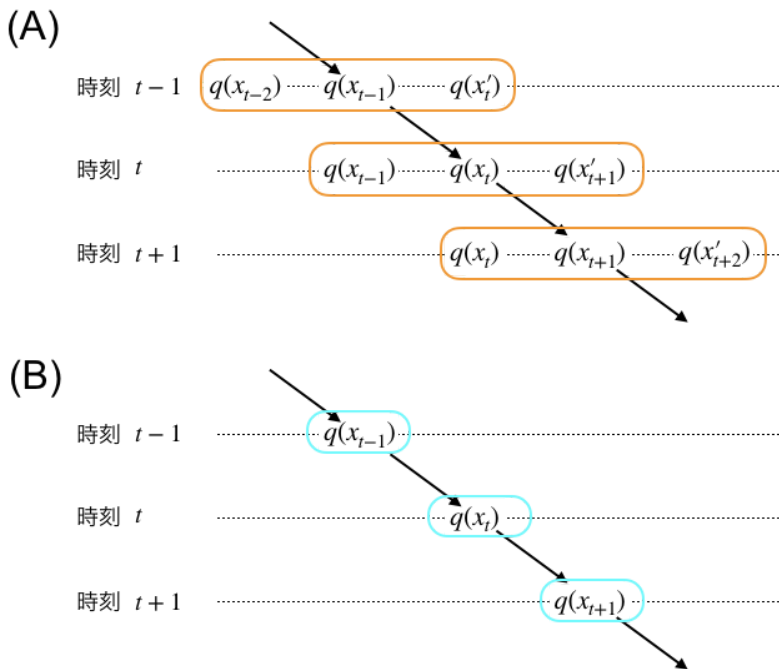


図5-3: 推測 $q$ の時間幅の違うagent

図5-3Aがこれまでの説明で使っていたものだが、図5-3Bはtemporal thicknessがまったくない「今を生きる」生物だ。Fristonによれば後者は意識を持たないという話になる。しかし、大腸菌の遺伝子発現ですらも予測的な振る舞いをする事が知られていること(私のブログに言及あり)を考慮に入れると、図5-3Bのように極端な生物はいそうにない。しかし予測的な行動をすることと、counter-factual predictionをすることはイコールではない。そしてFEPがまだ十分に因果推論的な形式になっていないことを考えると、Counter-factual predictionを捉えるのもっとうまい方法が必要になるのではないかと私は考えている。

以上でこの文書は終了です。ここまで読んでくれた方、ありがとうございます。

## [編集履歴]

- 2019/5/3 完成版作成
- 2019/5/4 式(4-34)、図4-15関連部分の修正。図1-3と関連部分の修正。図4-2と関連部分の修正。



- 2019/5/17 島崎秀昭さん(京大)のコメントを元に修正多数。(島崎さん、ありがとうございます！)

## [ToDo]

- Bogaczの二項分布バージョンを2-2あたりに入れる。これによって、脳が持っているものが事前分布と観察モデルだけであること、true posteriorは持っていないこと、 $q(s)$ や $q(a)$ は持っていないくて、たんに $q(x_t)$ ,  $q(x_{t+1})$ から生成モデルを使って変換しているだけだけであること、の説明が可能となる。
- $t=0$ での状況を作っておく。暗い部屋にいて、今照明がついてないと推測した。ではスイッチをオンするか。こうするともっと矛盾が明確になる。
- サッカーとサリエンシーのモデル