大脳における意外な情報処理

㈲数理解析研究所（IMA）　一色 浩

第22回SAPT測位技術振興会

2022.01.21

Fristonの自由エネルギ最小の原理により大脳の働きを統一的に解釈できるようになりつつある。

大脳の研究自体も画期的な発展段階を迎えつつある。

人は時々刻々変わる環境の中で生き抜く。

そのためには、環境の中で最適な推論、意思決定、学習という知的行動が必須である。

ひとは与えられた環境を受動的に受け取るだけでない。

環境の中での最良の知的行動がもたらされるように、能動的に環境を観察することが求められる。

脳はランダムなニューロンの集まりであるが，学習により自己組織化する。

自己組織化（じこそしきか、英: self-organization）とは、物質や個体が、系全体を俯瞰する能力を持たないのに関わらず、個々の自律的な振る舞いの結果として、秩序を持つ大きな構造を作り出す現象

例：雪の結晶，生殖

https://ja.wikipedia.org/wiki/%E8%87%AA%E5%B7%B1%E7%B5%84%E7%B9%94%E5%8C%96#:~:text=%E8%87%AA%E5%B7%B1%E7%B5%84%E7%B9%94%E5%8C%96%EF%BC%88%E3%81%98%E3%81%93%E3%81%9D,%E7%9A%84%E7%A7%A9%E5%BA%8F%E5%BD%A2%E6%88%90%E3%81%A8%E3%82%82%E8%A8%80%E3%81%86%E3%80%82

生物とは何か

生物の最大の目的：生き続けること（自己保存）

生きていること：平衡状態の維持→熱力学的には自由エネルギ最小

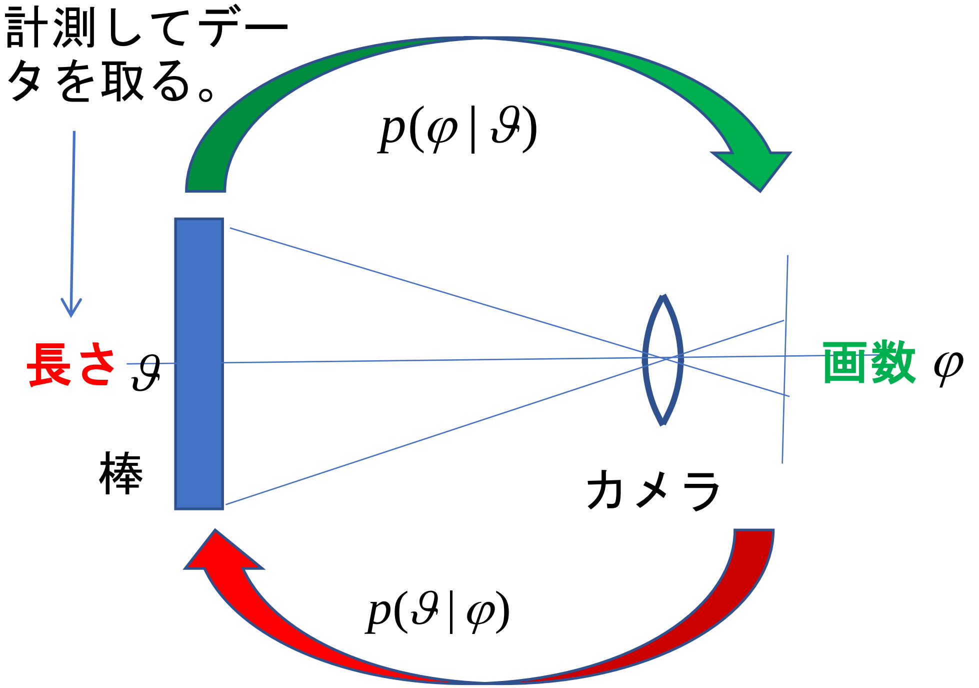
環境変化→自由エネルギ最小の原理により新しい平衡状態に移行

生物と機械は根本的に異なる。その違いは生命のあるなしである。生命の本質は未だ未知と言えよう。生命の本質が解明されない段階では、機械が生物を追い越すと考えるのは早計と言わざるを得ない。

$$\boxed{\text{解}:p(\vartheta\,|\,\varphi)}$$

カメラと物差しを使って大きさを測る

目標：ベイズの推論により、事後分布 $p(\vartheta\,|\,\varphi)$ を求める。

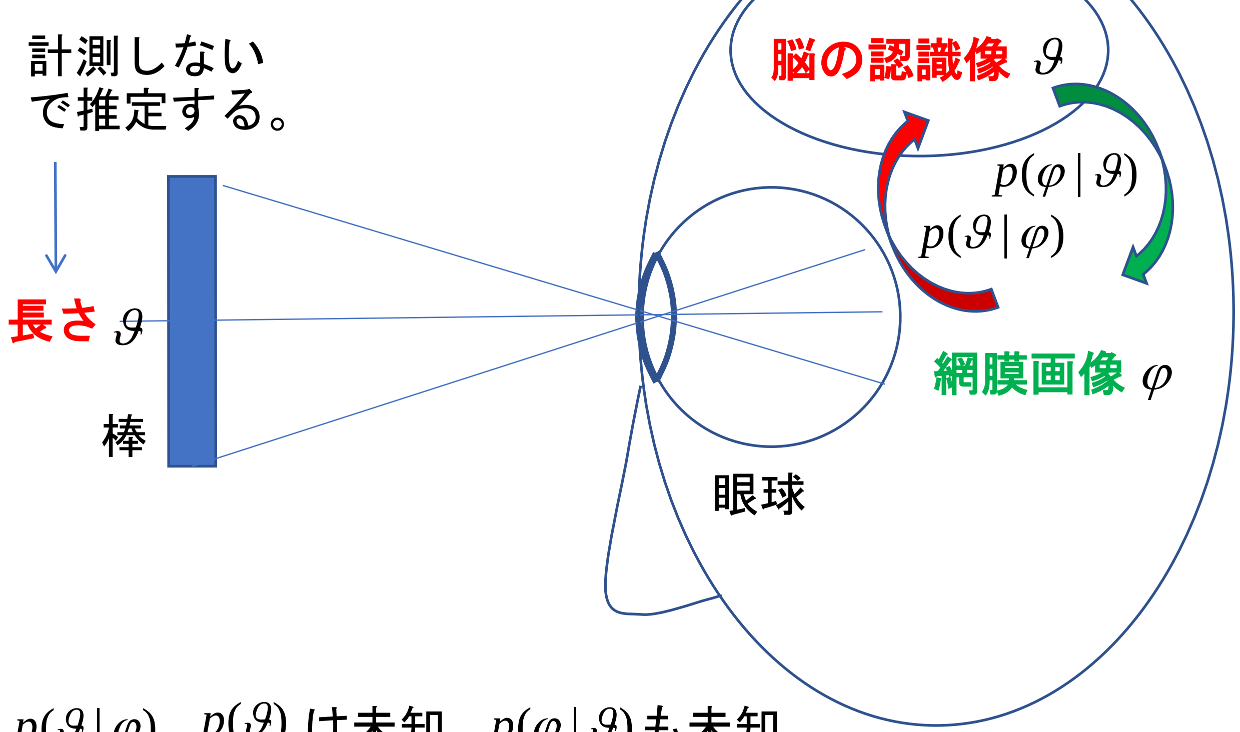


$$\textit{Bayes Theorem}:\ p(\vartheta\,|\,\varphi)=\frac{p(\varphi\,|\,\vartheta)\,p(\vartheta)}{p(\varphi)}$$

この場合、$p(\vartheta)$と $p(\varphi\,|\,\vartheta)$も既知

視覚像から物の大きさを推定する

目標：*VFE*=min により事後分布 $q(\vartheta)\sim p(\vartheta\,|\,\varphi)$を求める。アクションも動員。



$p(\vartheta\,|\,\varphi)$、$p(\vartheta)$ は未知。$p(\varphi\,|\,\vartheta)$も未知

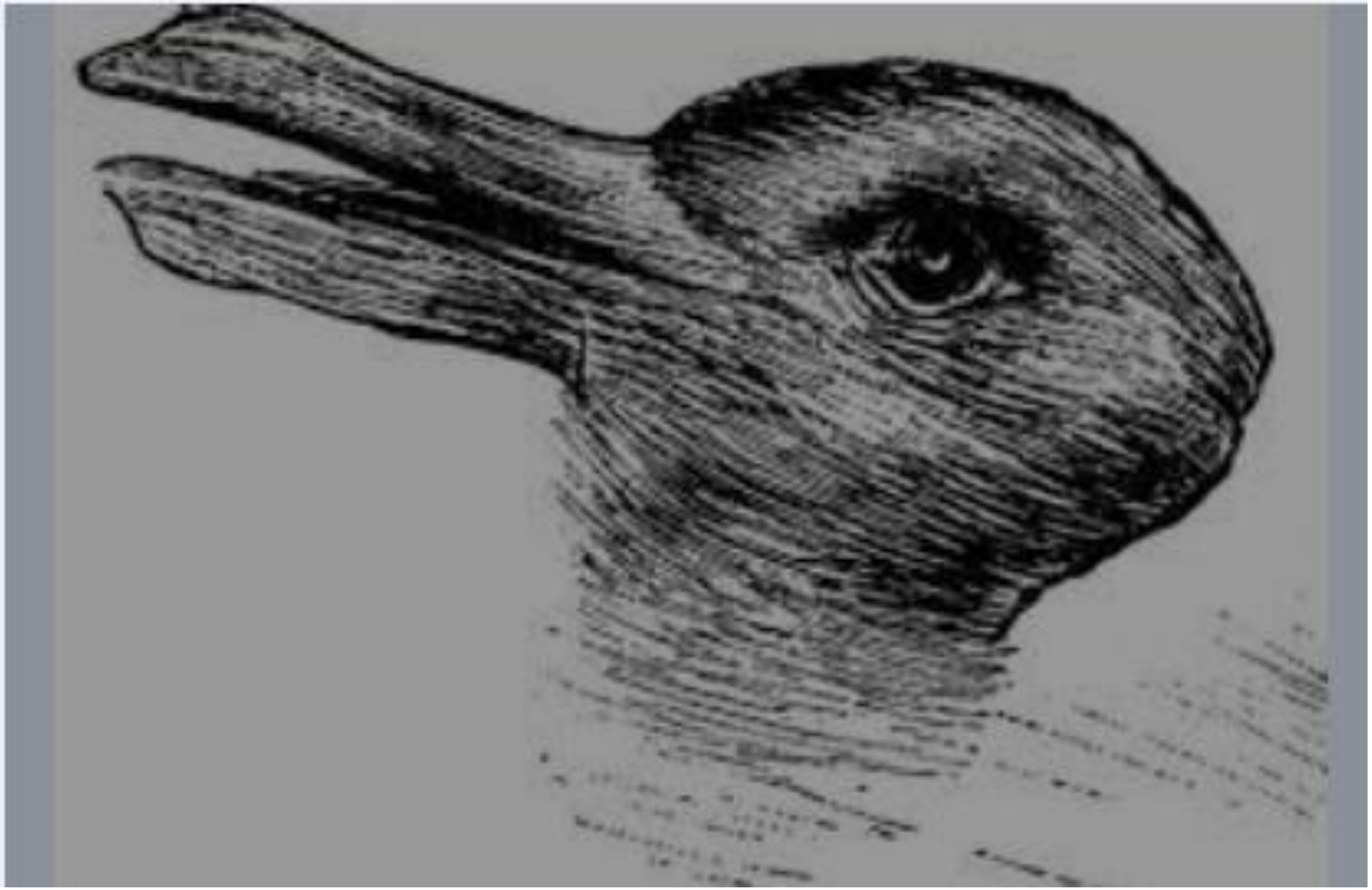
$$\textit{Variational Free Energy}:\ F\equiv KL(q(\vartheta)\,\|\,p(\vartheta,\varphi))$$

$$=\int q(\vartheta)\log\frac{q(\vartheta)}{p(\vartheta,\varphi)}dv=KL(q(\vartheta)\,\|\,p(\vartheta\,|\,\varphi))-\log p(\varphi)$$

ヘルムホルツ：ものが見えるのは、
無意識的推論である。

ヘルムホルツ： 1821年8月31日 - 1894年9月8日）は、ドイツ出身の生理学者、物理学者。
19世紀ドイツを代表する科学者。
熱力学におけるヘルムホルツの自由エネルギー （等温等積条件の下で仕事として取り出し可能なエネルギー：減少する方向が化学反応の方向）
https://ja.wikipedia.org/wiki/%E3%83%98%E3%83%AB%E3%83%9E%E3%83%B3%E3%83%BB%E3%83%95%E3%82%A9%E3%83%B3%E3%83%BB%E3%83%98%E3%83%AB%E3%83%A0%E3%83%9B%E3%83%AB%E3%83%84

だまし絵：人間が観て認識するものは脳が推論したもの



右を見るとウサギ、左を見るとアヒル

同じものを見ているのに、見ているものが二つある？！

https://media.thisisgallery.com/20220629

ルビンの壺 https://biz.trans-suite.jp

黒を注視すると壺、
白を注視すると二人の顔

立体画像：
右目と左目の2枚の網膜像から3次元像を推論

Jmc.or.jp

雲がゴジラに見える。ゴジラを知らない人
ならただの雲？

https://yuki-wan.at.webry.info/201702/article_23.html

顔の自動認識機能を使って見つけたらしい？

https://gigazine.net/news/20140918-cloud-face/

ラバーハンド：
ラバーハンドを自分のハンドと錯覚

錯覚を作り出す。
フォークで突き刺す。
実ハンドを引っ込める。

Is Reality a Controlled Hallucination? - with Anil Seth ... 42:56
https://www.youtube.com/watch?v=qXcH26M7PQM&list=PLbnrZHfNEDZwT_sW6joezEWrkVYX8dPa3&index=2

脳の問題を議論する上での根本的な認識

問題を解くということは、事後確率 $p(\vartheta|\varphi)$ を求めること。

脳で事後確率 $p(\vartheta|\varphi)$ を求めるには？ 脳は**知識**(同時確率) $p(\vartheta,\varphi)$ を、**認識** (尤度関数) $p(\varphi|\vartheta)$ と**信念**(事前確率) $p(\vartheta)$ として持っている。

ベイズ推論を、**自由エネルギ最小の原理**に置き換え、ラプラス近似と最急降下法を用いると、脳で実現可能な計算法が導ける。

解 $p(\vartheta|\varphi)$の近似として $q(\vartheta)$ を考える。以下では $q(\vartheta)$ に関する最小原理を考えるので、自由エネルギは**変分自由エネルギ**になる。最小値に到達すると $q(\vartheta)=p(\vartheta|\varphi)$ になる。

変分自由エネルギ（Variational Free Energy: VFM）

Variational Free Energy (VFE)

$$F \equiv KL(q(\vartheta) \,\|\, p(\vartheta, \varphi)) \equiv \int q(\vartheta) \ln \frac{q(\vartheta)}{p(\vartheta, \varphi)} d\vartheta$$

$$= KL(q(\vartheta) \,\|\, p(\vartheta \mid \varphi)) - \ln p(\varphi) = \min \quad \text{のとき}$$

$$KL(q(\vartheta) \,\|\, p(\vartheta \mid \varphi)) = \min, \quad -\ln p(\varphi) = \min$$

Fristonの自由エネルギ最小の原理とアクション

感覚 φ は環境との向き合い方で変わる。例えば視覚像の場合には眼球の方向を変えれば見え方が変わる。

自由エネルギ最小の原理は、単に $\varphi = \mathrm{given}$ での事後確率 $p(\vartheta \mid \varphi)$ を求めるのではなくて、サプライズ $-\ln p(\varphi)$ 最小の条件で事後確率を求める事を可能にする。

すなわち、環境への働きかけ“アクション”により精度が著しく向上する…能動的推論

脳情報空間モデル…静的な場合

環境と人間との情報の流れを現象論的に数理化する。

認識 ϑ $\leftrightarrow$ 信念 $\bar{\mu}$

認識感覚モデル

$\varphi = \vartheta + \varepsilon_z$

$\vartheta = \bar{\mu} + \varepsilon_w$

信念認識モデル

環境 S $\leftrightarrow$ 感覚 φ

$\varphi = S + \gamma$ 感覚環境モデル

静的な場合：

脳と環境の情報の流れ

この問題の解法は、最小自乗法でも、ベイズ推論でも、変分自由エネルギ最小でも考えられる。

解析的な考察

$$E(\mu,\varphi)=\arg\min_{\mu}\frac{1}{2}\left(\frac{(\varphi-\mu)^2}{\sigma_z}+\frac{(\mu-\vartheta_d)^2}{\sigma_w}+\log(\sigma_z\sigma_w)\right)$$

分散大（精度低）のデータは無視する

$$\frac{\partial E}{\partial\mu}=\frac{-(\varphi-\mu)}{\sigma_z}+\frac{(\mu-\vartheta_d)}{\sigma_w}=0$$

$$\mu=\frac{1}{\sigma_z+\sigma_w}(\sigma_w\varphi+\sigma_z\vartheta_d)\rightarrow\begin{cases}\varphi & \text{when}\quad \sigma_z\ll\sigma_w\\ \vartheta_d & \text{when}\quad \sigma_z\gg\sigma_w\end{cases}$$

感覚の精度が高い場合

信念が強い場合

認識が感覚に引きずられたり，信念に引きずられることが数式の上でも分かる → 経験を積むことにより脳が変化し精度が向上（脳の可塑性）

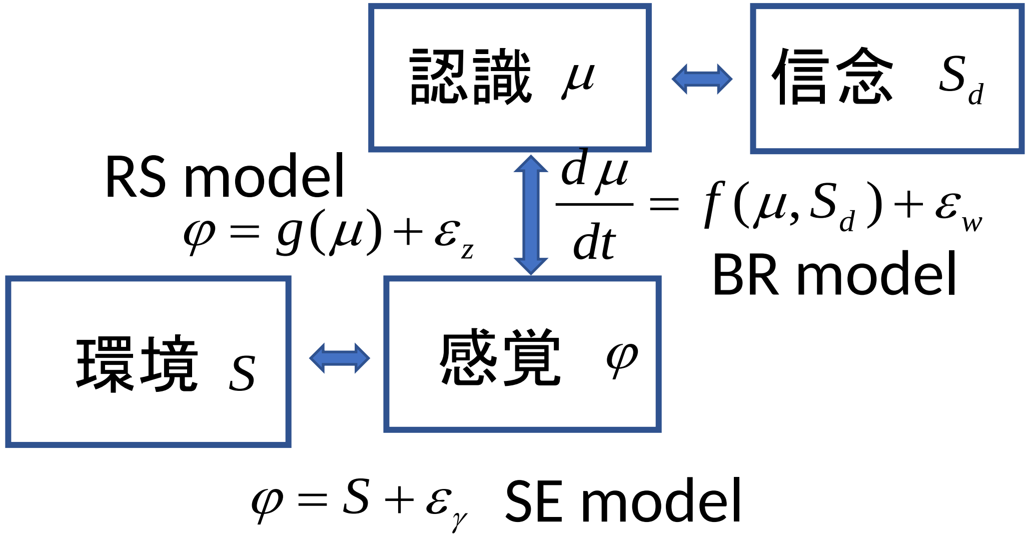
脳情報空間モデル…動的な場合

動的問題に関するFristonの考え方は複雑で分り難い． Buckleyと同様な問題を通して，Fristonと異なる動的問題の考え方を以下に示す．はるかに分かり易いと思う．

ストーブのような熱源がある．近づけば温度が上がるし，離れれば下がる．好みの温度の所に行きたい.1次元空間の時刻 t_i,座標 x_iにおける温度T_iの分布は

$$T_i = \frac{T_0}{1+x_i^2}, \quad \left(\frac{dT}{dx}\right)_i = -\frac{2x_iT_0}{(1+x_i^2)^2}$$

環境と人間との情報の流れを数理化する。



動的な場合：
脳と環境の情報の流れ

$$T_i = \frac{T_0}{1+x_i^{\,2}}, \quad \left(\frac{dT}{dx}\right)_i = -\frac{2x_i T_0}{(1+x_i^{\,2})^2}$$

環境T_i〜感覚φ_i, 感覚φ_i〜認識μ_iおよび認識μ_i〜意志μ_dモデル

$$\varphi_i = T_i + \varepsilon_{\gamma i}, \qquad \varphi_i = g(\mu_i) + \varepsilon_{zi}, \qquad \mu_i = \mu_{i-1} - dt(\mu_{i-1} - \mu_d) + \varepsilon_{w\,i-1}$$

$$E_i = -\ln p_i(\mu_i, \varphi_i) = \frac{(\varphi_i - g(\mu_i))^2}{2\sigma_{zi}} + \frac{(\mu_i - [\mu_{i-1} - dt(\mu_{i-1} - \mu_d)])^2}{2\sigma_{wi}}$$

$$+\frac{1}{2}\ln(2\pi\sigma_{zi}) + \frac{1}{2}\ln(2\pi\sigma_{wi}) = \min$$

$$\frac{\partial E_i}{\partial \mu_i} = -\frac{\varphi_i - g(\mu_i)}{\sigma_{zi}} g'(\mu_i) + \frac{\mu_i - [\mu_{i-1} - dt(\mu_{i-1} - \mu_d)]}{\sigma_{wi}}, \qquad \frac{\partial E_i}{\partial x_i} = \frac{\partial E_i}{\partial \varphi_i}\frac{\partial \varphi_i}{\partial x_i} = -\frac{(\varphi_i - \mu_i)}{\sigma_{zi}}\left(\frac{dT}{dx}\right)_i$$

$$\mu_i = \mu_{i-1} - \kappa \frac{\partial E_i}{\partial \mu_i}, \quad x_i = x_{i-1} - \kappa \frac{\partial E_i}{\partial x_i}$$

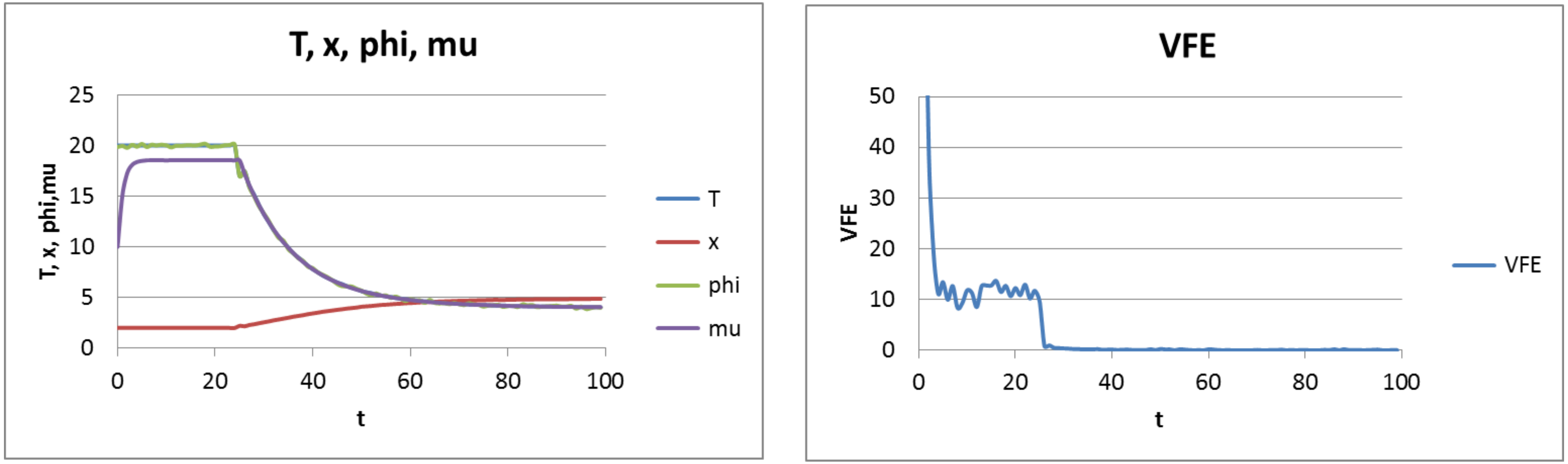


図　温度が設定値に収束する。アクションは t ＝25に開始する。

根源論的数理化：統計力学のランジュバン方程式、フォッカー・プランク方程式、ランダム系の自己組織化

(https://www.youtube.com/watch?v=H9I0PmXwhdo)

熱運動をする多数の小粒子と衝突する粒子

Langevin eq. $m\dot{v} = \underbrace{-\gamma v}_{\text{粘性力}} + \underbrace{R(t)}_{\text{ランダム力}}$

$$\langle R(t)\rangle = 0 \quad \langle R(t)R(t')\rangle = q\delta(t-t')$$

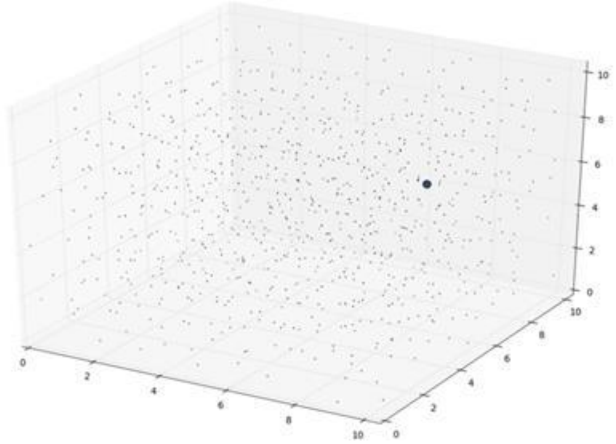
$$P\left(R(t)\right) \propto e^{-R^2/2\langle R^2\rangle}$$

Solution $v(t) = v_0 e^{-\gamma t/m} + \dfrac{1}{m}\int_0^t e^{-\gamma(t-t')/m}R(t')dt'$

$$\to \langle v(t)\rangle = v_0 e^{-\gamma t/m}$$

$$\langle v^2(t)\rangle = v_0 e^{-2\gamma t/m} + \frac{q}{2\gamma m}\left(1-e^{-2\gamma t/m}\right)$$

$$\langle v^2(t\to\infty)\rangle = \frac{q}{2\gamma m} \Rightarrow q = 2\gamma k_B T$$



$P(v,t)$: probability density function

Fokker-Planck eq.

$$\frac{\partial P(v,t)}{\partial t} = \frac{\gamma}{m}\frac{\partial}{\partial v}\left(vP(v,t)\right) + \frac{\gamma k_B T}{m^2}\frac{\partial^2 P(v,t)}{\partial v^2}$$

Stationary solution: $\dfrac{\partial P(v,t)}{\partial t} = 0$

$P(v) = Ce^{-mv^2/2k_BT}$: Maxwell distribution

結論に代えて般若心経

(https://www.nhk-sc.or.jp/han/pdf/hannyashingyou.pdf)

仏説摩訶般若波羅蜜多心経

唐三蔵法師玄奘 訳

観自在菩薩 行深般若波羅蜜多時
照見五蘊皆空 度一切苦厄
舎利子 色不異空 空不異色
色即是空 空即是色
受想行識 亦復如是
舎利子 是諸法空相
不生不滅 不垢不浄 不増不減
是故空中 無色無受想行識
無眼耳鼻舌身意 無色声香味触法
無眼界 乃至無意識界
無無明 亦無無明尽
乃至無老死 亦無老死尽
無苦集滅道 無智亦無得 以無所得故
菩提薩埵 依般若波羅蜜多故
心無罣礙 無罣礙故 無有恐怖
遠離一切顛倒夢想 究竟涅槃
三世諸仏 依般若波羅蜜多故
得阿耨多羅三藐三菩提
故知般若波羅蜜多
是大神呪 是大明呪
是無上呪 是無等等呪
能除一切苦 真実不虚
故説般若波羅蜜多呪
即説呪曰
羯帝 羯帝 波羅羯帝 波羅僧羯帝
菩提僧莎訶
般若心経

経典には優れた文学性が求められる。論理を超えた表現が使われる。

般若心経で空を強調するのは、文字通りの空ではないのでは？

我々が考える有は我々が考えるような堅固なものではなく、時とともに移り変わるもので、固執すべきものではない。

要するに絶対的な有が存在すると考えるのは幻である。

絶対的な有が存在しないということを象徴的に空と呼んでいるのではないか？

最後の二行は呪文である。いわばキリスト教の「アーメン」と同じものであろう。呪文は意味を求める対象でなく「合言葉」であるから、言語によらず共通であることに意味があると玄奘が考えたのであろう。

般若心経の教えを老いを生きる杖としたい。

そのために必要なことは、老いを直視することであろう。

年取ると脳細胞が減少し、知的活動に障害が出る。

深い考察ができなくなり、錯覚が多くなる。

無意識のうちに老害をまき散らすことになる。

これを避ける道は老いを自覚することしかないのでは？

ご清聴ありがとうございます！