

深層生成モデルを用いた意識の持つ機能の一部の実験その 2

Experiments of Several Functions in Consciousness Using Deep Generation Models Part 2

疋田 聡¹

Satoshi Hikida¹

¹ AGICRON 研究所株式会社

¹ AGICRON Research Institute, Ltd.

Abstract: In a previous presentation, we reported how to use deep generation models to improve the performance of AI using some of the functions of consciousness. However, the experiments were not completed in the last time, so I will report on the experiments this time.

1. 背景

前回の発表では、深層生成モデルを使用して、意識の機能の一部を用いて AI の性能を向上させる方法について報告した[30]。しかし、前回は時間の関係上実験が間に合わなかったので、今回は実験について報告する。

1.1. 「意識」という言葉の意味について

「意識」という言葉の意味については前回の発表でも説明したが[30]、誤解を減らすために再度記載しておく。

「意識」については、従来から様々な研究や提案が行われている

[1][9][10][11][12][14][15][16][17][22][23][25][28]。また、「意識」という言葉は、使う人によって意味が異なることが多いので、その意味をまとめておく。例えばドゥアンヌの[16]を参考にして「意識」という言葉の使われ方をまとめてみると、表 1 のようになる。

表 1 「意識」という言葉の使われ方

覚醒状態	眠っているか目覚めているかなど。
注意	特定の情報に集中すること。
コンシャスアクセス	注意を向けられた情報に気づくこと。
自己意識	自己が存在することに気づく能力。
メタ認知	自分自身の認知行動を再帰的に把握することができる能力
自由意思	自律的な意思決定を行える能力。
クオリア	経験の主観的な生の感覚。

このように、「意識」という言葉は場合によって様々な意味に用いられるので、どの意味で用いられているのかに注意する必要がある。

また、「意識」という言葉は意識が一つのもののように使われがちであるが、意識は複数の機能が相互作用して成り立っている可能性が高いので、その言葉が意識のどの一部分の機能を指して使われているのかに留意することも重要であると思われる。

1.2. 本研究で対象とする意識の機能

本研究では、意識が持つと考えられる機能の一部を汎用人工知能の性能向上に利用することを目的とし、意識の情報処理の方法をヒントとしてアルゴリズムを考える。

本研究で対象とする意識の機能は、主にドゥアンヌの[16]で詳しく説明されているコンシャスアクセスと自由意思の一部に関係する機能である。

2. 一貫した情報の生成と選択機能

前回の発表では、意識の持つ機能の一部として一貫した情報の生成と選択の機能について、グレゴリーのダルメシアン¹の知覚、両眼視野闘争、カプグラ症候群を例として説明した。また、このような機能を実現するためのアルゴリズムとして、深層生成モデルを用いて候補を生成する機能と、生成した候補を基準値で検証する機能と、基準値を満たした候補が得られるまで一定時間候補生成と検証を反復する機能と、基準値を満たした候補だけを意識に上らせる機能を持つアルゴリズムについても説明した[30]。

前回との重複を減らすため、今回はグレゴリーのダルメシアン知覚についてだけ記載しておく。

2.1. グレゴリーのダルメシアン知覚

グレゴリーのダルメシアン知覚は、白黒のまだら模様があるだけに見える絵で、犬がいて見るとダルメシアンが姿が見えてくる知覚現象である[17]。そして、一度ダルメシアンが意識に上ると、その姿は見え続ける。

これは、意識の持つ機能の一部として、ノイズの多い輪郭線の中から犬のイメージと一貫性がある輪郭線を選択する機能があるためと考えられる。



図 1 Gregory R (1970) "The intelligent eye" McGraw-Hill, New York (Photographer: RC James)

3. 一貫した情報の学習機能

前回の発表では、リベットの自由意思の実験[9]を例にして、前章のアルゴリズムによって基準値を満たした候補だけを意識に上らせ、それを用いて自由意思モデルと整合するように無意識処理を学習するアルゴリズムを提案した[30]。このアルゴリズムにより、時間遅れが許容できる次回から自由意思モデルによる結果と無意識処理による結果が一致するように学習した処理が反映されるようになり、一貫した情報を用いたノイズの少ない高精度の学習との両立が可能となる。

前回との重複を減らすため、今回はこの部分の説明は省略する。

4. 深層生成モデルを用いた意識の持つ機能の一部のアルゴリズム

アルゴリズムについては前回の発表でも説明したが[30]、今回も再度記載しておく。

なお、前回の記載では、深層生成モデルの生成フェーズの処理で z の事前分布からサンプリングするという説明になっていたが、今回は入力データにノイズを含む画像を与えて z を算出する形に変更した。

4.1. 一貫した情報の候補の生成

一貫した情報の候補の生成には、深層生成モデルを用いた。深層生成モデルは、深層ニューラルネットワークを用いて確率モデルを学習するもので、高い表現能力に特徴がある。例えば、VAE[5][6]、GAN[4]、GQN[3]などが有名である。本研究では、条件付き生成を行うため、Conditional Variational Auto Encoder (CVAE) [6]を用いた。

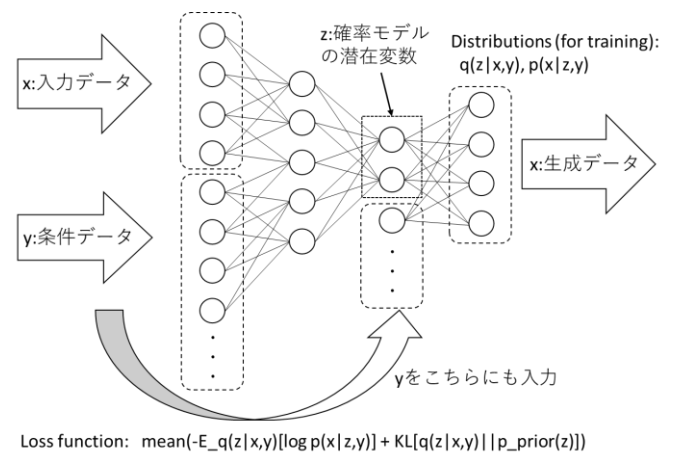


図 2 深層生成モデルの処理

図 2 に深層生成モデルの処理の概念図を示しておく。

CVAE を用いて候補を生成することにより、例えば、グレゴリーのダルメシアン知覚では、犬という条件データを指定して候補を生成できるようになる。

4.2. 一貫した情報の選択と学習

深層生成モデルを用いて候補を生成した後は、生成した候補の一貫性を基準値で検証し、基準値を満たした候補が得られるまで一定時間候補生成と検証を反復し、基準値を満たした候補だけを意識に上らせ、それを用いて自由意思モデルと整合するように

無意識処理を学習する。

図 3 に、以上の処理をまとめたフローチャートを示す。このアルゴリズムにより、時間遅れが許容できる次回から自由意思モデルによる結果と無意識処理による結果が一致するように学習した処理が反映されるようになり、一貫した情報を用いたノイズの少ない高精度の学習との両立が可能となる。

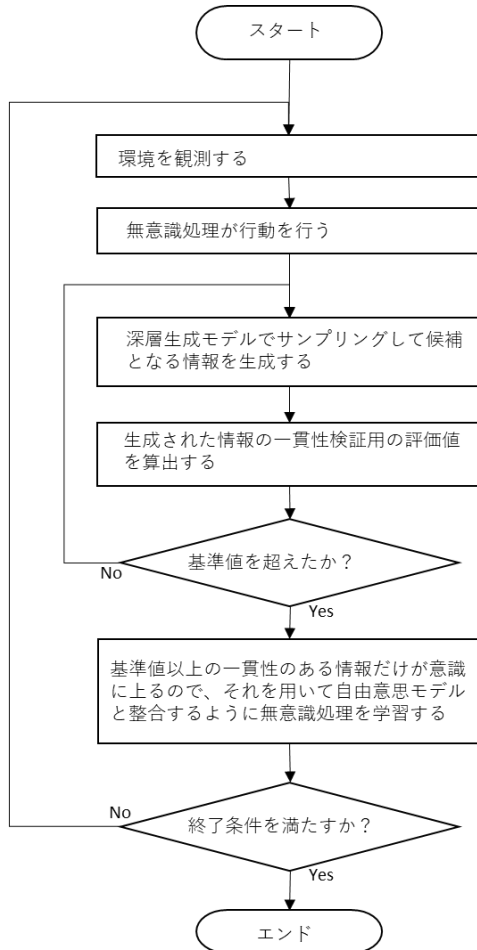


図 3 深層生成モデルを用いた意識の持つ機能の一部のアルゴリズムのフローチャート

5. 実験方法と結果

意識の持つ機能の一部と考えられる一貫した情報の生成と選択機能を実現するアルゴリズムの実験として、ダルメシアン知覚を単純化し、ランダムな線の中から深層生成モデルで候補を生成し、輪郭線の一貫性が偶然の確率と比較して基準値以上の場合に選択を行う実験を行った。図 4 に実験の概念図を示す。

5.1. 学習フェーズ

入力データにノイズを含まない画像とノイズを含む画像を、画像の種類を条件データとして与え、正解データにノイズを含まない画像を与えて深層生成モデル (CVAE) を学習した。

5.2. 認識フェーズ

入力データにノイズを含む画像を与えて画像を生成し、生成した候補の一貫性を基準値で検証し、基準値を満たした候補が得られるまで一定時間候補生成と検証を反復し、基準値を満たした候補だけを選択した。

今回の場合は、生成画像の輪郭線と入力データの輪郭線が一致する確率が偶然よりも十分に高い場合に基準値を満たすように設定した。

5.3. 実験ツール

深層生成モデルのツールには Pixyz[32]を用いた。Pixyz を用いることにより、深層生成モデルが簡単に記述できるようになる。

5.4. 実験結果

なお、実験結果については時間的な問題で研究会の発表時に示す予定である。

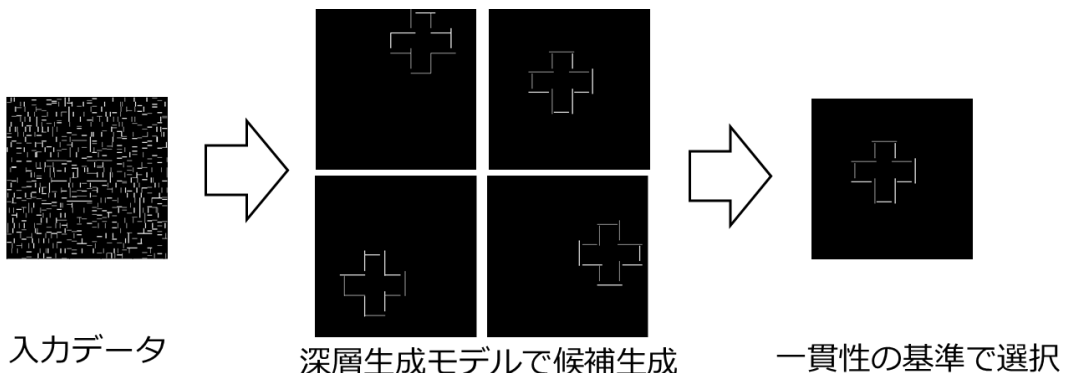


図 4 深層生成モデルを用いた一貫した情報の生成と選択の実験の概念図

6. まとめ

前回の発表では深層生成モデルを使用して、意識の機能の一部を用いて AI の性能を向上させる方法について報告したが、前回は時間の関係上実験が間に合わなかったため、今回はその実験に関して示した。

参考文献

- [1] Bengio, Y., The consciousness prior, *arXiv preprint arXiv:1709.08568*, (2017)
- [2] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding, *arXiv preprint arXiv:1810.04805*, (2018)
- [3] Eslami, S. M. A., Rezende, D. J., Besse, F., ... & Viola, F.: Neural scene representation and rendering, *Science*, Vol. 360, Issue 6394, pp. 1204-1210, (2018)
- [4] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y.: Generative adversarial nets, In *Advances in neural information processing systems* (pp. 2672-2680), (2014)
- [5] Kingma, D. P., & Welling, M.: Auto-encoding variational bayes, *arXiv preprint arXiv:1312.6114*, (2013)
- [6] Kingma, D. P., Mohamed, S., Rezende, D. J., & Welling, M.: Semi-supervised learning with deep generative models, In *Advances in neural information processing systems* (pp. 3581-3589), (2014)
- [7] Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., ... & Dieleman, S.: Mastering the game of Go with deep neural networks and tree search, *nature*, 529(7587), 484-489, (2016)
- [8] Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., ... & Lillicrap, T.: Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm, *arXiv preprint arXiv:1712.01815*, (2017)
- [9] B.リベット, (訳)下條信輔, *マインド・タイム : 脳と意識の時間*, 岩波書店, (2005)
- [10] C.コッホ, (訳)土谷尚嗣, 金井良太, *意識の探求 : 神経科学からのアプローチ*, 岩波書店, (2006)
- [11] C.コッホ, (訳)土谷尚嗣, 小畑史哉, *意識をめぐる冒険*, 岩波書店, (2014)
- [12] D.C.デネット, (訳)山口泰司, *解明される意識*, 青土社, (1998)
- [13] D.カーネマン, (訳)村井章子, *ファスト&スロー : あなたの意思はどのように決まるか?*, 早川書房, (2012)
- [14] M. ガザニガ, (訳) 藤井留美, 「わたし」はどこにあるのか : ガザニガ脳科学講義, 紀伊國屋書店, (2014)
- [15] M.マッスイミーニ, G.トノーニ, (訳)花本知子, *意識はいつ生まれるのか : 脳の謎に挑む統合情報理論*, 亜紀書房, (2015)
- [16] S. ドゥアンヌ, (訳)高橋洋, *意識と脳—思考はいかにコード化されるか*, 紀伊國屋書店, (2015)
- [17] V.S.ラマチャンドラン, (訳)山下篤子. *脳のなかの天使*. 角川書店, (2013)
- [18] 磯村拓哉, 自由エネルギー原理の解説 : 知覚・行動・他者の思考の推論, *日本神経回路学会誌* 25(3) 71-85, (2018)
- [19] 乾敏郎, *感情とはそもそも何なのか : 現代科学で読み解く感情のしくみと障害*, ミネルヴァ書房, (2018)
- [20] 乾敏郎, 知覚・認知・運動・感情・意思決定をつなぐ自由エネルギー原理, *日本神経回路学会誌* 25(3) 123-134, (2018)
- [21] 吉田正俊, 田口茂, 自由エネルギー原理と視覚的意識, *日本神経回路学会誌* 25(3) 53-70, (2018)
- [22] 浅田稔, 痛みを感じるロボットの意識・倫理と法制度, *人工知能学会誌* 33(4), 450-459, (2018)
- [23] 前野隆司, *脳はなぜ「心」を作ったのか : 「私」の謎を解く受動意識仮説*, 筑摩書房, (2004)
- [24] 大羽成征, 自由エネルギー原理は脳型人工知能の基盤となるか〜アクティブインファレンスの意義について〜, *日本神経回路学会誌* 25(3) 113-122, (2018)
- [25] 大泉匡史, 統合情報理論から考える人工知能の意識, *人工知能学会誌* 33(4) 460-467, (2018)
- [26] 大槻知史, 三宅陽一郎 監修, *最強囲碁 AI アルファ碁 解体新書 増補改訂版*, 翔泳社, (2018)
- [27] 渡辺慧, (訳)村上陽一郎, 丹治信春, *知識と推測 : 科学的認識論*, 東京図書, (1987)
- [28] 渡辺正峰, *脳の意識機械の意識 : 脳神経科学の挑戦*, 中央公論新社, (2017)
- [29] 疋田聡, 深層学習とモンテカルロ木探索を用いた強化学習の組合せ最適化問題での実験とフレーム問題に関する 1 考察, *人工知能学会研究会資料 SIG-AGI-009-07*, (2018)
- [30] 疋田聡, 深層生成モデルを用いた意識の持つ機能の一部の実験, *人工知能学会研究会資料 SIG-AGI-013-05*, (2019)
- [31] 疋田聡, 深層生成モデルを用いた人型ロボットの動作生成実験, *人工知能学会研究会資料 SIG-AGI-012-03*, (2019)
- [32] <https://github.com/masa-su/pixyz> (2020/2/13 現在)