Addressing the Subsumption Thesis

A Formal Bridge between Microeconomics and Active Inference

Noé Kuhn

10.09.2024

IWE-HSG

Subsumption Thesis

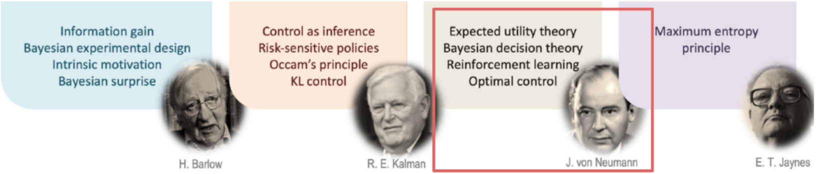


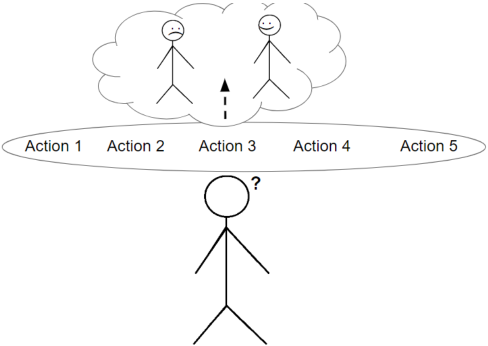
Figure 1: Figure 1: Active Inference intertwines with other accounts of behaviour. Figure from Friston et al., 2023

> "Active inference [...] englobes the principles of expected utility theory [...] it is theoretically possible to rewrite any RL algorithm [...] as an active inference algorithm." [1]

→ Expected utility theory, as seen in economics, is subsumed by active inference – it is an edge case

Motivation

- First Principles of Agency
- Exploitation vs. Exploration dilemma
- Economic welfare implications



Setting

Agent-Environment Frameworks:

Commensurable Objective Functions in a

- Markov Decision Process (MDP):
  $(\mathbb{S}, \mathbb{A}, P(s'|a, s), R(s'), \gamma = 1, \mathbb{T})$
- Partially Observable MDP (POMDP):
  $(\mathbb{S}, \mathbb{A}, P(s'|a, s), P(o|a, s'), R(s'), \gamma = 1, \mathbb{T})$

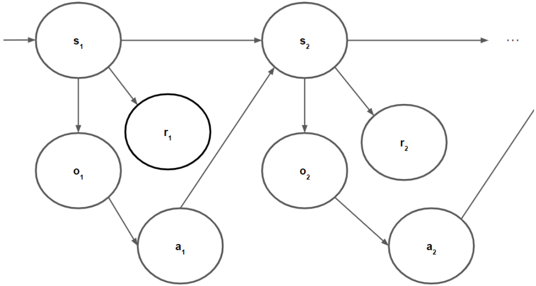


Figure 2: Graphical Model of a POMDP

Exhibit

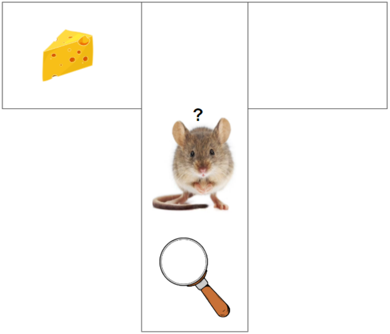


Figure 3: T-Maze POMDP [2] [3]

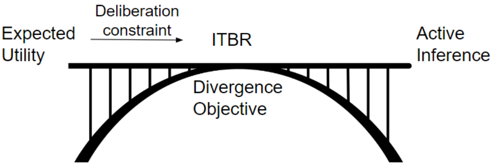
$$U(R) = R(s)^c \ : \ c \in \mathbb{R}+$$

Risky strategy: $E[U(L_1)] = 0.5 \cdot 0 + 0.5 \cdot 2^c$

Risk averse strategy: $E[U(L_2)] = 1^c$

ITBR Bridge

Connection: Divergence Objective [5]



Consider a utility maximizing agent with information-theoretic bounds on the KL-Divergence between the prior conditional distribution $P(s|a)$ and posterior distribution upon deliberation $Q(s|a)$ [6]

MDP Comparison

The agent then faces the following Lagrangian optimization problem:

$$F_{ITBR}(Q) = \sum_{s} Q(s|a) \left(U(s,a) - \frac{1}{\beta} log \frac{Q(s|a)}{P(s|a)} \right) \tag{1}$$

The solution to which is the Gibb's distribution:

$$P^*(s|a) = \frac{P(s|a) \cdot e^{\beta U(s,a)}}{Z_\beta(a)} \to P^*(s) \tag{2}$$

which is plugged into the objective function (1) to obtain the divergence objective.

Optimal agency is then equivalent for ITBR and active inference (Divergence objective):

$$a^* = \underset{a \in \mathbb{A}}{\arg\min} D_{KL}[P(s_\tau|a_t)||P^*(s)] \tag{3}$$

POMDP Comparison

The Lagrangian and preference distribution can analog be applied to a POMDP:

$$a^* = \underset{a \in \mathbb{A}}{\arg\min} D_{KL}[P(o_\tau, s_\tau | a_t) || P^*(o, s)] \tag{4}$$

Whereby the relationship between the expected free energy G and $-F_{ITBR}$ is as follows:

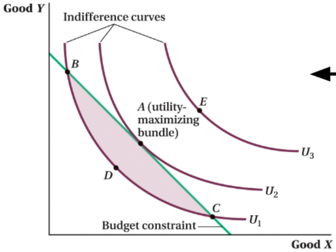
$$G - E_{Q(o|a)} \mathfrak{H}[Q(o|s)] = -F_{ITBR} \tag{5}$$

And the extrinsic value differs as follows [5]:

$$P^*(s) = \frac{e^{U(R(s))}}{\sum_s e^{U(R(s))}} \quad \text{and} \quad P(s|C) = \frac{e^{R(s)}}{\sum_s e^{R(s)}} \tag{6}$$

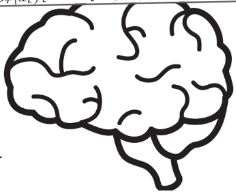
Conclusion

- Bridge: Reward maximization $\rightarrow$ Expected utility $\rightarrow$ ITBR $\rightarrow$ Active inference
- Key differences for POMDPs (5) - Which free energy?
- Potential for active inference economies [7]



$$\boxed{-E_{P(o_\tau|a_t)}[D_{KL}[P(s_\tau|o_\tau)||P(s_\tau|a_t)]] -}$$

$$\boxed{E_{P(o_\tau|a_t)}[D_{KL}[P(s_\tau|o_\tau)||P(s_\tau|a_t)]]}$$



END

Addressing the Subsumption Thesis: A Formal Bridge between Microeconomics and Active Inference

Slides available on *https://github.com/noeromeo*

QUESTIONS?

Citations:

1: Da Costa & Tenka & Zhao & Sajid, 2024, pg. 2

2: Da Costa, Friston, Parr, Sajid, 2022, pg. 10

3: Da Costa & Tenka & Zhao & Sajid, 2024, pg. 9

4: Da Costa & Tenka & Zhao & Sajid, 2024, pg. 10

5: Millidge & Buckley, 2021

6: Ortega, Braun, Dyer, Kim, Tishby, 2015

7: Hyland, Gavenciak, Da Costa, et al., in preparation

References:

Barto & Sutton (2018): An Introduction to Reinforcement Learning

References ii

https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

Da Costa (2022): Reward Maximisation through Discrete Active Inference
https://arxiv.org/abs/2009.08111

Da Costa, Parr, Sajid et al (2020): Active inference on discrete state-spaces: A synthesis
https://doi.org/10.1016/j.jmp.2020.102447

Friston (2006): A Free Energy Principle for the Brain
https://doi.org/10.1016/j.jphysparis.2006.10.001

Friston (2015): Active inference and epistemic value
10.1080/17588928.2015.1020053

Ortega & Braun (2011): Information, Utility & Bounded Rationality https://arxiv.org/pdf/1107.5766

Ortega et al. (2015): Information-Theoretic Bounded Rationality https://arxiv.org/pdf/1512.06789

Sajid, N., Da Costa, L., Parr, T., Friston, K.: Active inference, bayesian optimal design, and expected utility. In: Cogliati Dezza, I., Schulz, E., Wu, C.M. (eds.) The Drive for Knowledge: The Science of Human Information Seeking, pp. 124– 146. Cambridge University Press, Cambridge (2022)

Barto & Sutton (2018): An Introduction to Reinforcement Learning https://www.andrew.cmu.edu/course/10-703/textbook/BartoSutton.pdf

Friston (2006): A Free Energy Principle for the Brain https://doi.org/10.1016/j.jphysparis.2006.10.001

Ortega & Braun (2011): Information, Utility & Bounded Rationality

https://arxiv.org/pdf/1107.5766
Ortega et al. (2015): Information-Theoretic Bounded Rationality
https://arxiv.org/pdf/1512.06789
Mann et al. (2022): Free Energy: A user's guide
https://link.springer.com/article/10.1007/s10539-022-09864-z
Da Costa, Parr, Sajid et al (2020): Active inference on discrete state-spaces: A synthesis https://doi.org/10.1016/j.jmp.2020.102447
Da Costa (2022): Reward Maximisation through Discrete Active Inference https://arxiv.org/abs/2009.08111
Friston (2015): Active inference and epistemic value 10.1080/17588928.2015.1020053
Friston, K., Da Costa, L., Sajid, N., Heins, C., Ueltzhöffer, K., Pavliotis, G.A., Parr, T.: The free energy principle made simpler but not too simple. Physics Reports 1024, 1–29 (June 2023). https://doi.org/10.1016/j.physrep.2023.07.001

References v

Hyland, D., Gavenciak, T., Da Costa, L., Heins, C., Kovarik, V., Gutierrez, J., Wooldridge, M., Kulveit, J.: Multi-agent active inference. Forthcoming Manuscript in preparation

Appendix

Tangent: Preferences in Active Inference

Consider the following example:

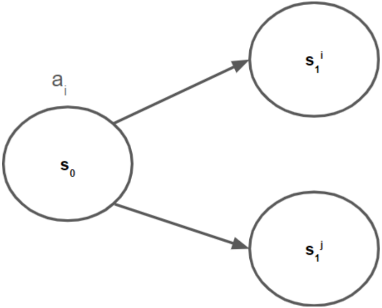


Figure 4: Transition between states given some action

$$P(s_1^i|a_1), P(s_1^j|a_1) = 0.6, 0.4 \quad P(s_1^i|a_2), P(s_1^j|a_2) = 0.1, 0.9$$
$$P(s_1^i) = 0.6 \cdot P(a_1) + 0.1 \cdot P(a_2) = 0.1$$
$$\rightarrow P(s_1^i) \leq 0.6$$

2nd Exhibit

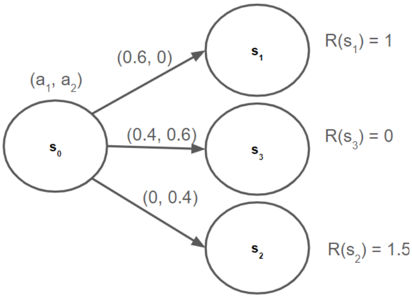


Figure 5: Paraglider POMDP

Desiderata[4]: Risk Aversion

Paraglider calc

Expected Utility and Active Inference for the 'Paraglider' MDP
The single-step MDP is specified as follows. Therefore note that the subscript does not pertain to the period:

$$\mathbb{S} = s_1, s_2, s_3$$
$$\mathbb{A} = a_1, a_2$$
$$\{P(s_1|a_1), P(s_2|a_1), P(s_3|a_1)\} = \{0.6, 0, 0.4\}$$
$$\{P(s_1|a_2), P(s_2|a_2), P(s_3|a_2)\} = \{0, 0.4, 0.6\}$$
$$\{R(s_1), R(s_2), R(s_3)\} = \{1, 1.5, 0\}$$

Consider an expected utility agent with utility function $U(R(s)) = R(s)^c$ where $c \in \mathbb{R}^+$. As such,

$$E[U(a_1)] = 0.6 \cdot 1^c$$
$$E[U(a_2)] = 0.4 \cdot 1.5^c$$
$$\text{For } c < 1 \rightarrow \underset{a \in \mathbb{A}}{\arg\max}\ E[U(a)] = a_1$$
$$\text{For } c > 1 \rightarrow \underset{a \in \mathbb{A}}{\arg\max}\ E[U(a)] = a_2$$

So a risk-averse expected utility agent will scale the smaller but safer mountain.

The active inference agent however is indifferent between the two actions. If we assume the preference distribution to be a softmax on the rewards, then we can ignore the normalizing denominator as it is constant w.r.t to action. Therefore we can write the relevant objective function as:

$$G(a_t) = -\sum_s P(s_\tau|a_t) \cdot R(s_\tau) - \sum_s P(s_\tau|a_t) log \frac{1}{P(s_\tau|a_t)}$$
$$G(a_1) = -0.6 - 0.3065 - 0.366 = G(a_2)$$
$$\rightarrow \underset{a \in \mathbb{A}}{\arg\min} G(a) = \{a_1, a_2\}$$

Therefore the optimal action of the risk-averse expected utility agent is a subset of the optimal active inference agency.

MDP math

(9)

$$P^*(s|a) = \frac{P(s|a)e^{\beta U(s)}}{Z_\beta}$$

$$\frac{1}{\beta} ln(P^*(s|a) \cdot Z_\beta) = U(s)$$

plug in for $U(s)$ (7)

$$argmax_a \frac{1}{\beta} E_{Q(s|a)}[lnP^*(s|a) + ln(Z_\beta)] - \frac{1}{\beta} ln \frac{Q(s|a)}{P(s|a)}$$

$$= argmax_a E_{Q(s|a)}[lnP^*(s|a) + ln(Z_\beta) - lnQ(s|a) + lnP(s|a)]$$

$$= argmin_a E_{Q(s|a)}[-lnP^*(s|a) - ln(Z_\beta) + lnQ(s|a) - lnP(s|a)]$$

$$argmin_a D_{KL}[Q(s|a)||P^*(s|a)]$$

$\rightarrow$ (10)

POMDP math

(13.5)

$$P^*(o,s|a) = \frac{P(o,s|a)e^{\beta U(o,s,a)}}{Z\beta}$$

plug in 13

end

$$= argmin_a D_{KL}[Q(o,s|a)||P^*(o,s|a)]$$