The Emperor's New Markov Blankets

Manuel Baltieri

Araya Inc.

COGS seminars - 17th October 2023

ARAYA

Outline

- The free energy principle vs. active inference
- Markov blankets and conditional probabilities (Pearl blankets)
- Markov blankets define “things" (Friston Blankets)
- Possible inconsistencies and issues
- Some ways out

Behavioral and Brain Sciences

cambridge.org/bbs

Target Article

Cite this article: Bruineberg J, Dołęga K, Dewhurst J, Baltieri M. (2022) The Emperor's New Markov Blankets. *Behavioral and Brain Sciences* 1–76. doi:10.1017/S0140525X21002351

The Emperor's New Markov Blanket

Jelle Bruineberg[a], Krzysztof Dołęga[b], Joe Dewhurst[c] and Man

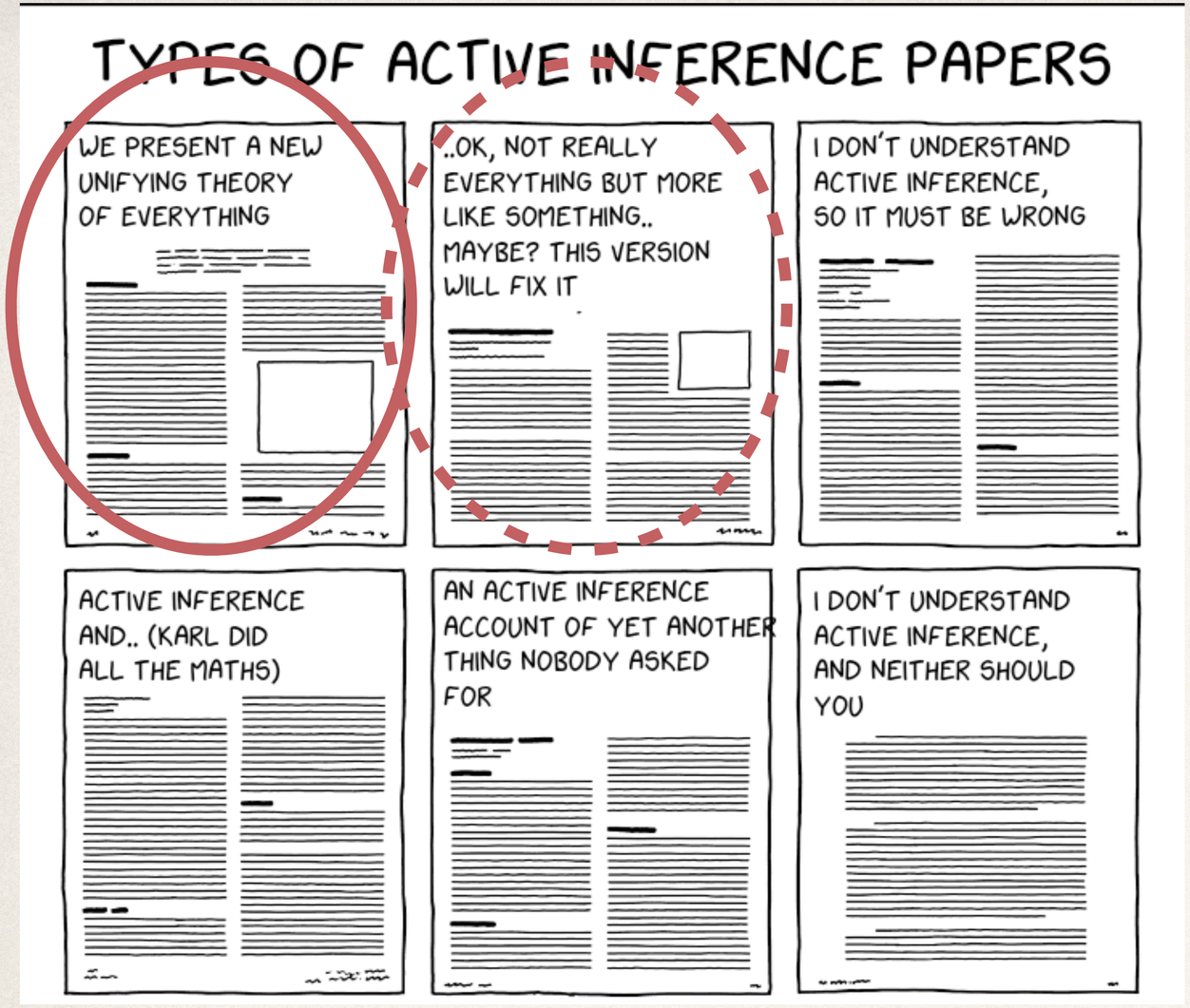
[a]Department of Philosophy, Macquarie University, Sydney, NSW 2109, Australia; [b]Insti
Fakultät für Philosophie und Erziehungswissenschaft, Ruhr-Universität Bochum, 4480
[c]Fakultät für Philosophie, Wissenschaftstheorieund Religionswissenschaft, Munich Cer
Philosophy, Ludwig-Maximilians-Universität München, 80539 Munich, Germany and [d]L
Computation and Adaptation, RIKEN Centre for Brain Science, 351-0106 Wako City, Ja
jelle.bruineberg@mq.edu.au
krzysztof.dolega@rub.de
joseph.e.dewhurst@gmail.com
manuel.baltieri@riken.jp

Authors' Response

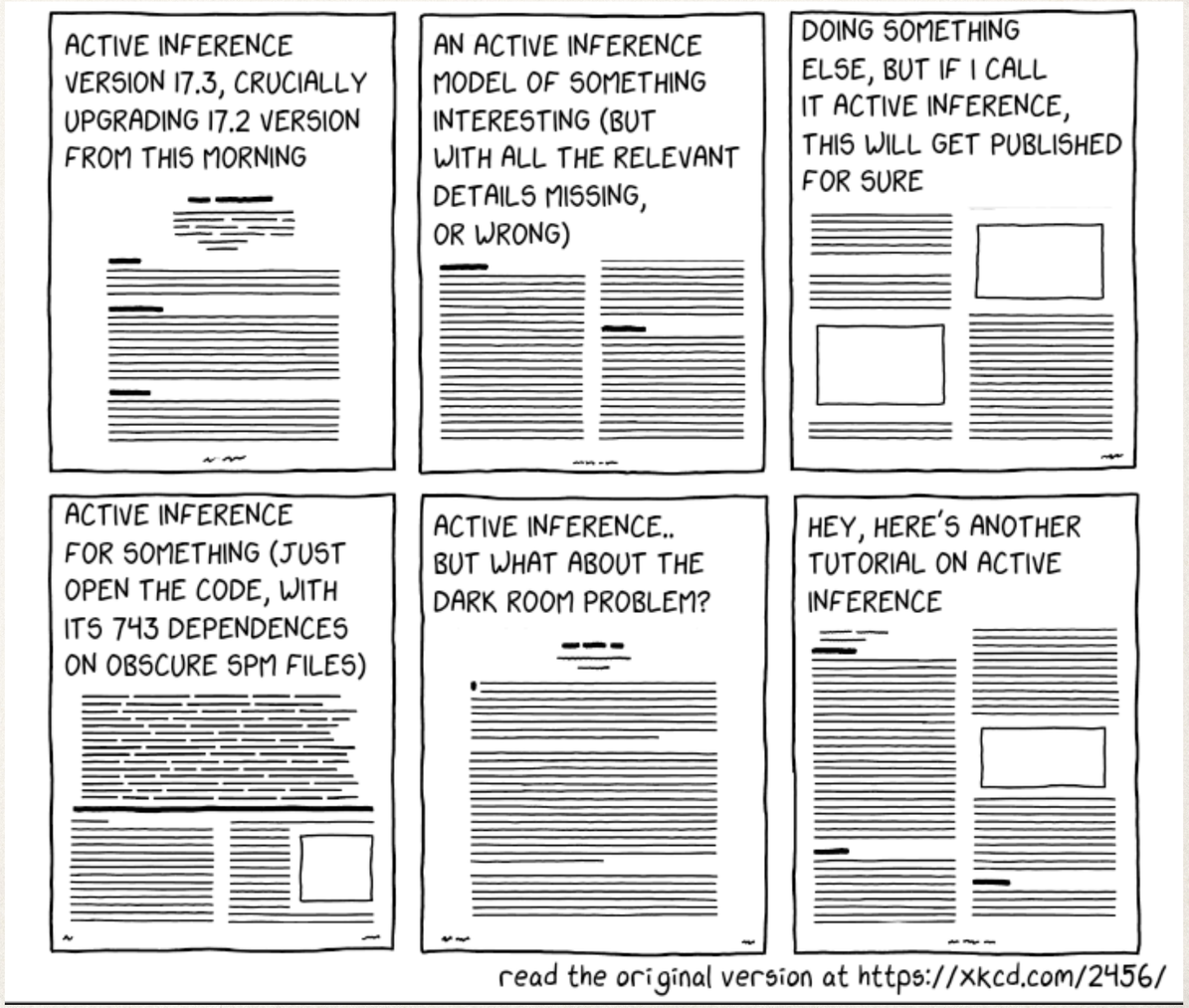
The Emperor is Naked: Replies to commentaries on the target article

Jelle Bruineberg[a], Krzysztof Dołęga[b], Joe Dewhurst[c] and Manuel Baltieri[d,e]

[a]Department of Philosophy, Macquarie University, Sydney, NSW 2109, Australia; [b]Institut für Philosophie II, Fakultät für Philosophie und Erziehungswissenschaft, Ruhr-Universität Bochum, 44801 Bochum, Germany; [c]Fakultät für Philosophie, Wissenschaftstheorieund Religionswissenschaft, Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, 80539 Munich, Germany; [d]Araya, Inc., Tokyo, Japan and [e]School of Engineering and Informatics, University of Sussex, Brighton BN1 9RH, UK
jelle.bruineberg@mq.edu.au
krzysztof.dolega@rub.de
joseph.e.dewhurst@gmail.com
manuel_baltieri@araya.org

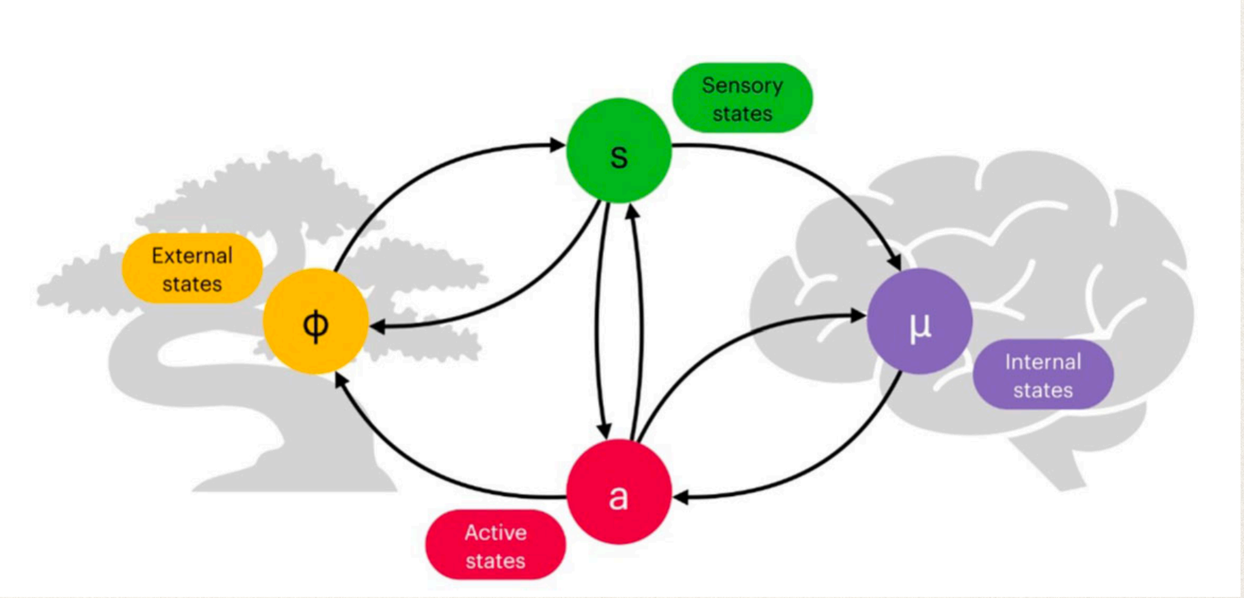


https://twitter.com/manuelbaltieri/status/1389622401915183104



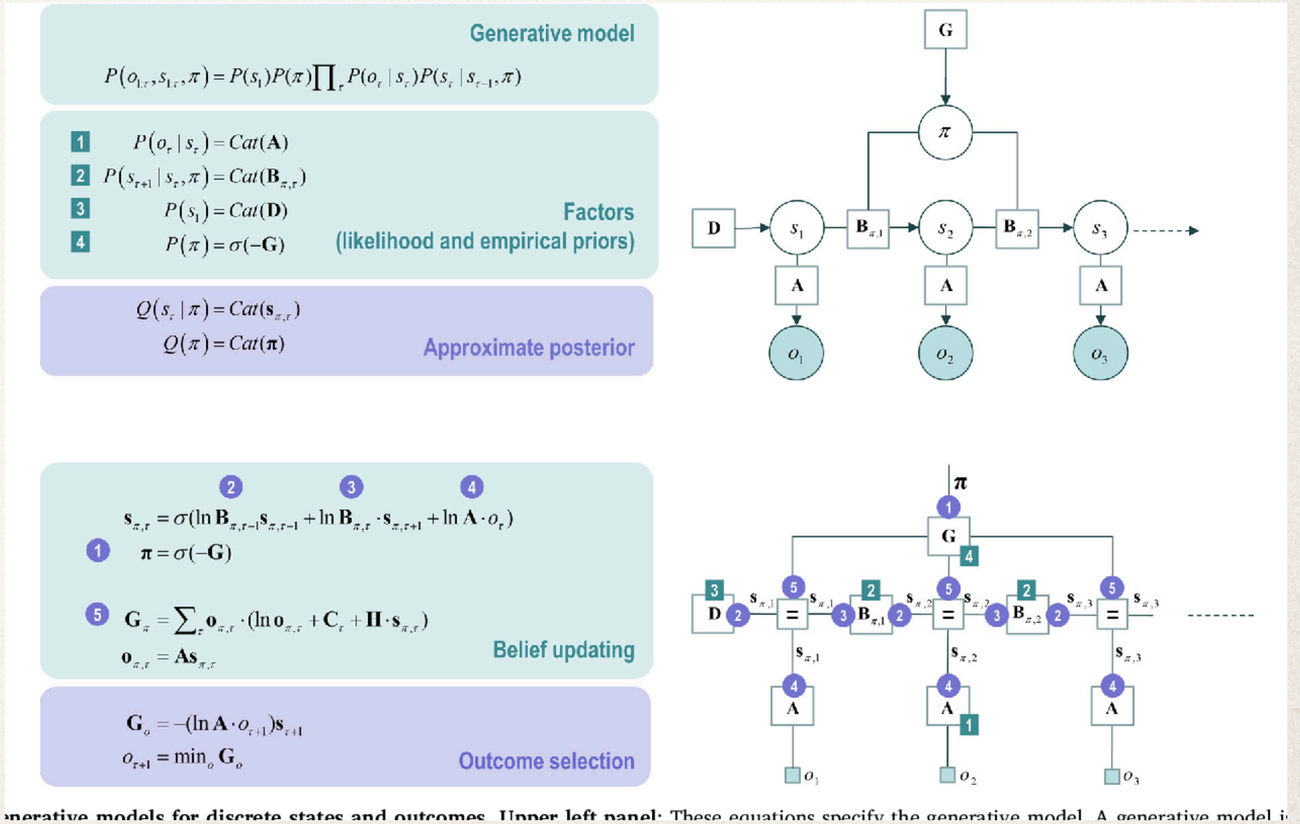
The free energy principle

- A foundational theory of agents, (living) systems, “things”
- A thing is a “thing” if and only if it (appears to) minimise(s) free energy
- Markov blankets as a “veil” that separates internal from external states



Active inference

- Assumes POMPDs/state-space models structure (~ RL setup)
- Provides an alternative cost function (expected free energy)
- …ideally one that is derived from the FEP, but <u>it can stand without it</u>



…nerative models for discrete states and outcomes. Upper left panel: These equations specify the generative model. A generative model i…

The FEP 1.01 - as of early 2021

The FEP targets:

1. systems which can be modelled as **random dynamical systems** with
2. a **unique steady-state distribution** (= weak mixing for recurrent but a-periodic Markov chains),
3. whose vector field can be **decomposed (via the Ao decomposition)**, uniquely and in a special way (= there's a number of equally valid alternatives), into orthogonal curl-free and divergence-free flows of a quasi-potential,
4. such that the set of random variables at steady-state (the stochastic process is effectively studied at steady-state) can be **partitioned into internal, external and blanket "states"** via an assumption (this is not an implication) of conditional independence between internal and external variables given the blanket (variables), based on a some **selection of** either **internal or external "states"** (the process is complementary),
5. under the additional assumption (a conjecture as seen in Friston et al. 2021, "Stochastic chaos and markov blankets") of "sparse coupling" that allows mapping of steady-state independencies to independencies on dynamical components, i.e., orthogonal curl-free and divergence-free flows,
6. and with a conditional synchronisation map assumed to connect the most likely internal and external states (see Aguilera et al. 2021 for possible issues) to try and ensure that internal variables *model* in some non-trivial sense external ones,
7. such systems can be said to contain a partition of internal states that appear to perform inference on a partition of external states via a gradient descent on variational free energy (*"Approximate Bayesian inference lemma"*).

More recent developments (~ 2022)

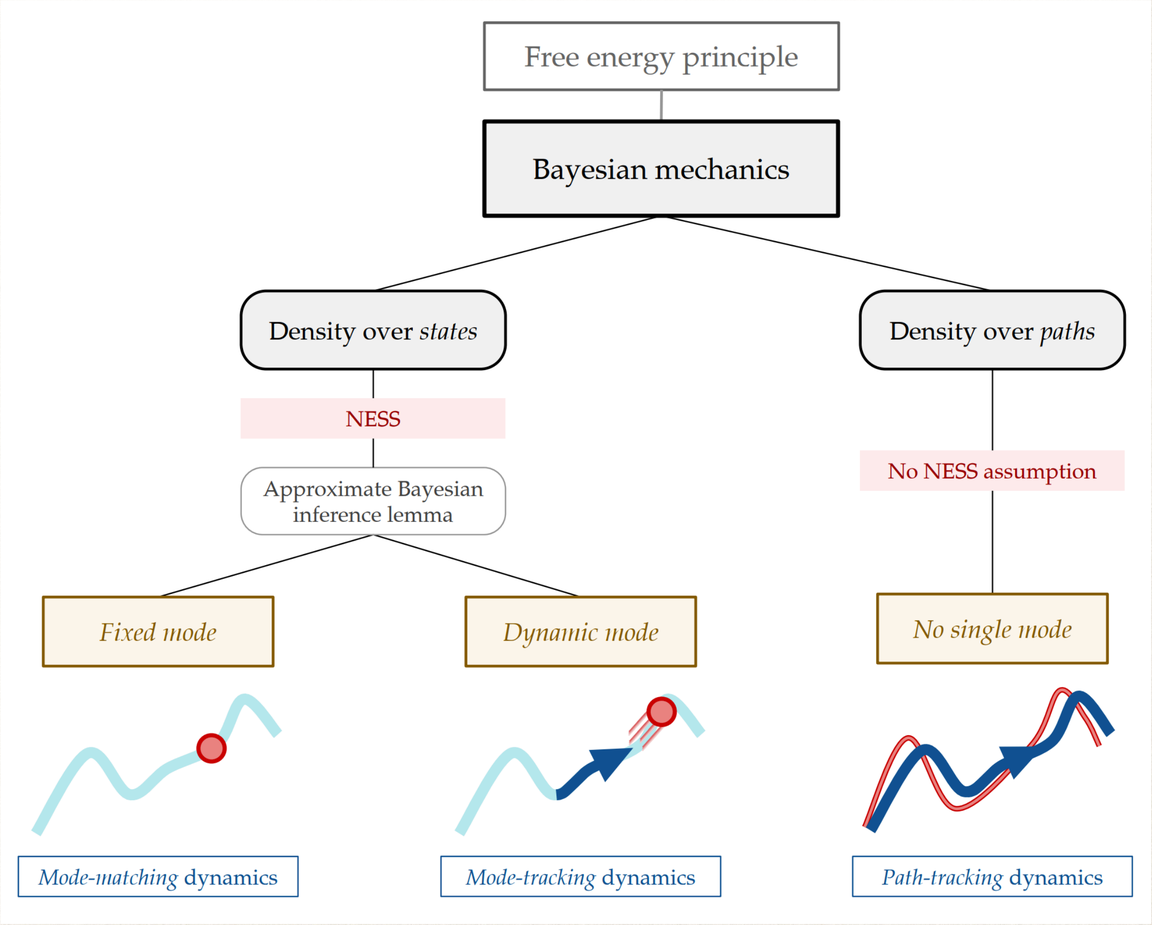
A. Da Costa et al., 2021 claims only stationarity, no uniqueness or a-periodicity required, but don't show a working example as far as I get it

B. New species of blankets keep on appearing (later, "**the zoo**")

C. FEP for non-stationary processes, but

   I. Friston (+ Pearl) Blankets are not meaningfully defined

   II. No "Approximate Bayesian inference lemma"? (What's the FEP without this again?)

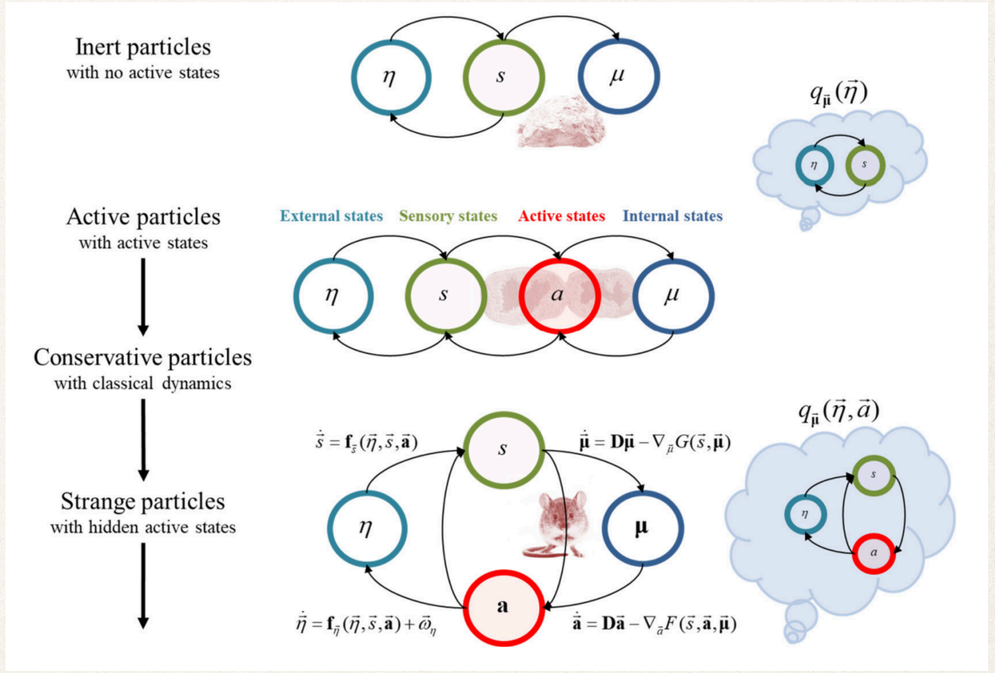


More (more) recent developments (~ 2023)

A. Path integrals solve all the issues? Not sure

B. New species of particles (inert, active, conservative, strange) and associated blankets? (Not in "**the zoo**")

C. Reviews à la "The free energy principle made simpler but not too simple", with the usual questions



Some basics

- A (joint) probability $p(x, y)$

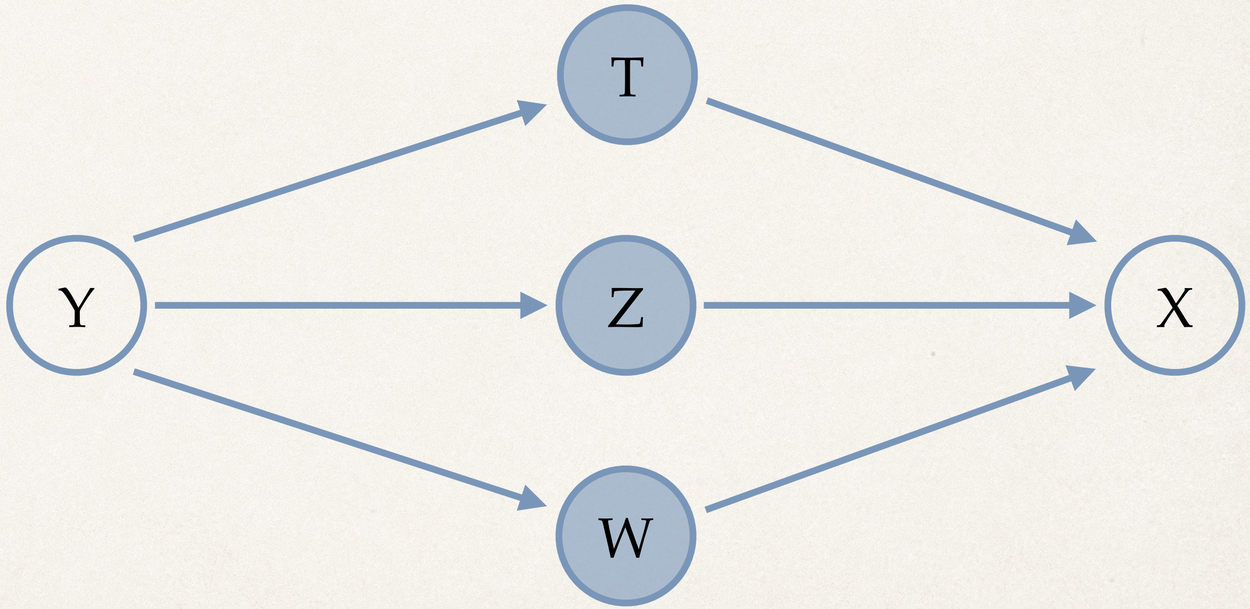
- A conditional probability $p(x \mid y) = \frac{p(x, y)}{p(y)}$

- Marginal independence $p(x \mid y) = p(x) \Rightarrow p(x, y) = p(x)p(y)$

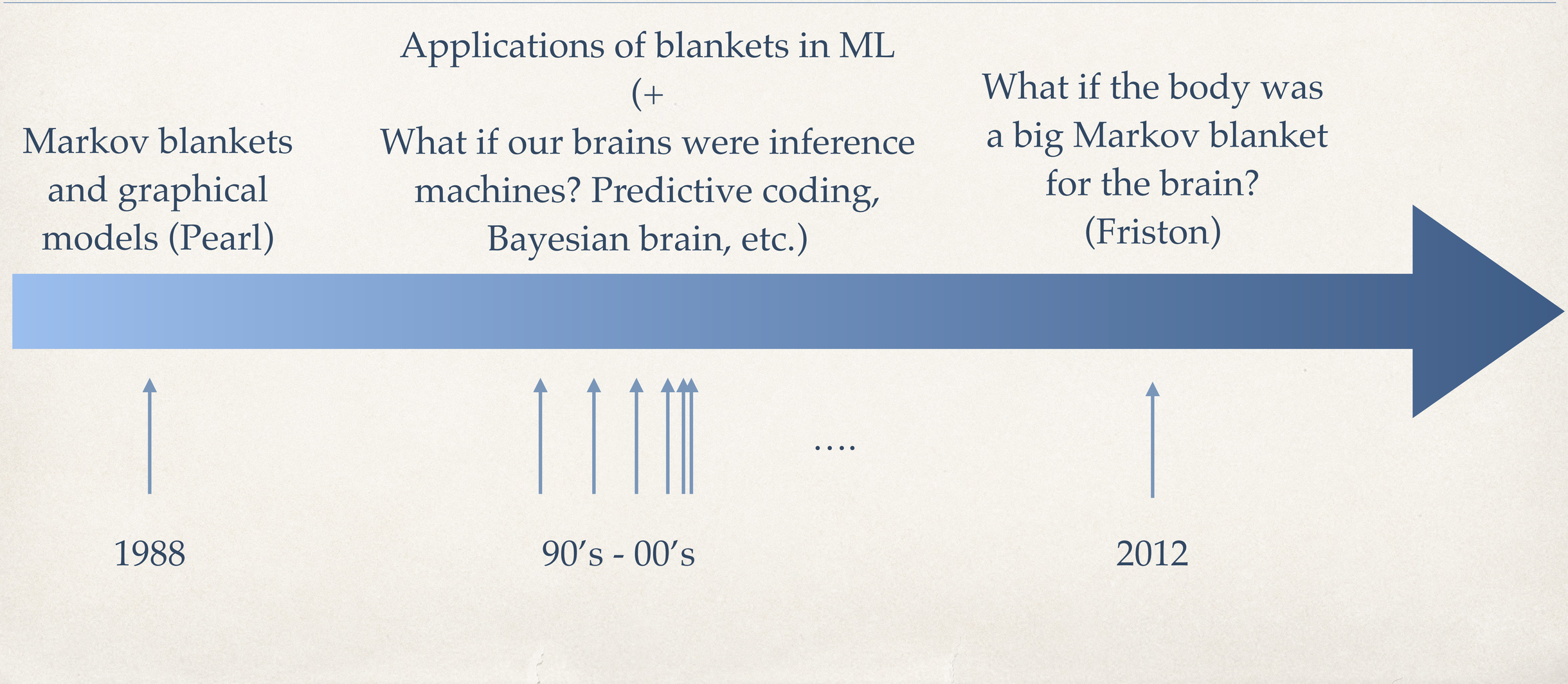
- Conditional independence (example) $p(x \mid y, z) = p(x \mid z) \Rightarrow p(x \mid z)p(y \mid z)$

What is a Markov blanket?

- If this were my entire model, z would be Markov blanket of y (or x): the set of random variables “shielding” y from x
- More in general however, we can have complicated models, and in that case z is only a part of the Markov blanket
- So, Markov blanket ~ the set of random variables (e.g., t, w, z) that render a (set of) random variable(s) (e.g., y) conditionally independent of a (set of) random variable(s), (e.g., x)

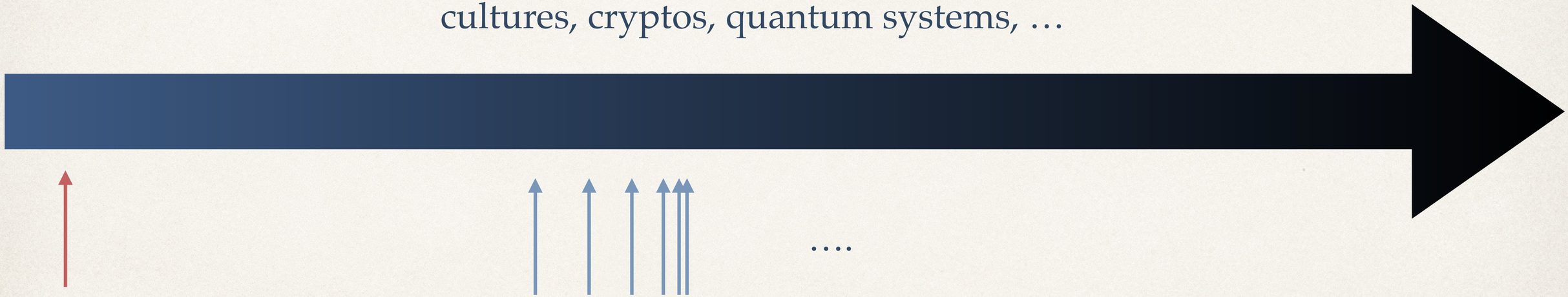


Forensics of blankets



Markov blankets of… [*your favourite system*]

Markov blankets of life, mind,
self, sex and gender, pain experience, religious practices
climate and ecosystems, social systems,
cultures, cryptos, quantum systems, …



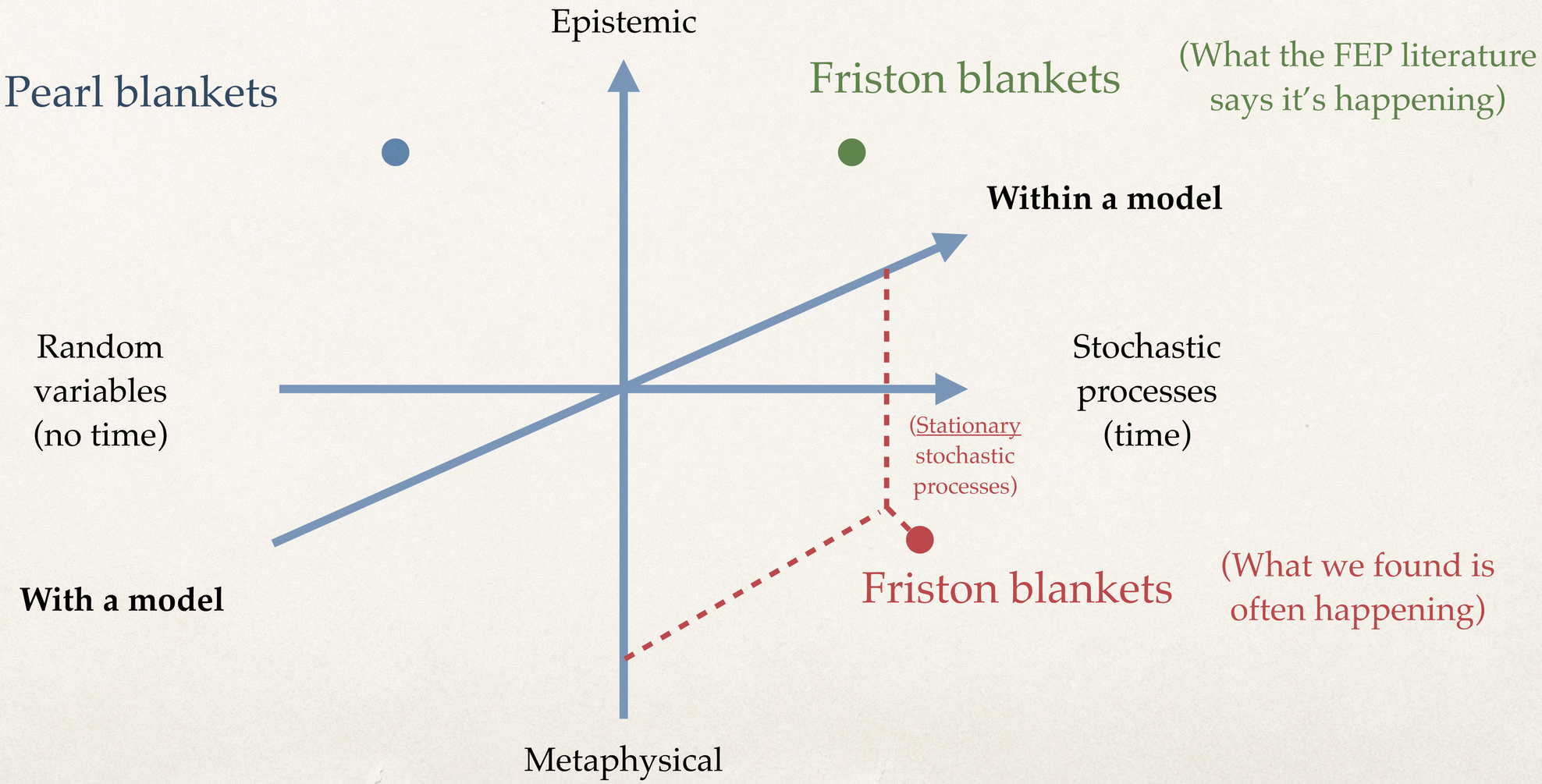
….

2013

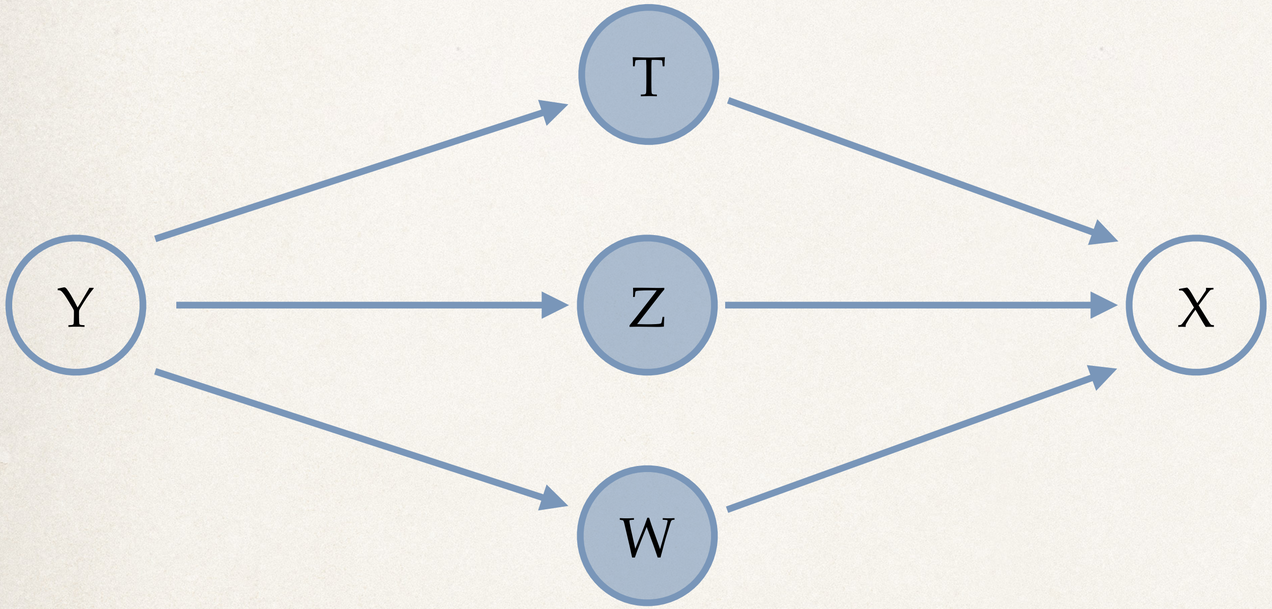
“Life as we know it”

2013 - now

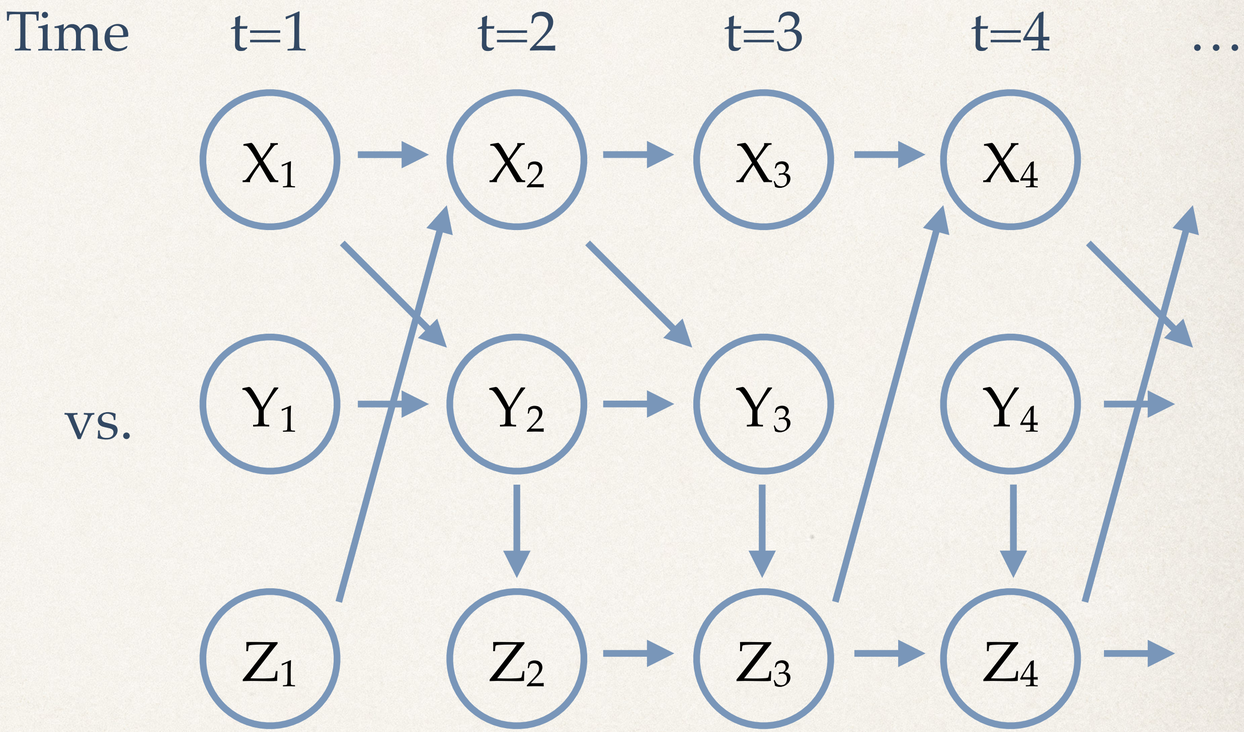
From Pearl to Friston blankets: just maths?



1. From random variables to stochastic processes



vs.

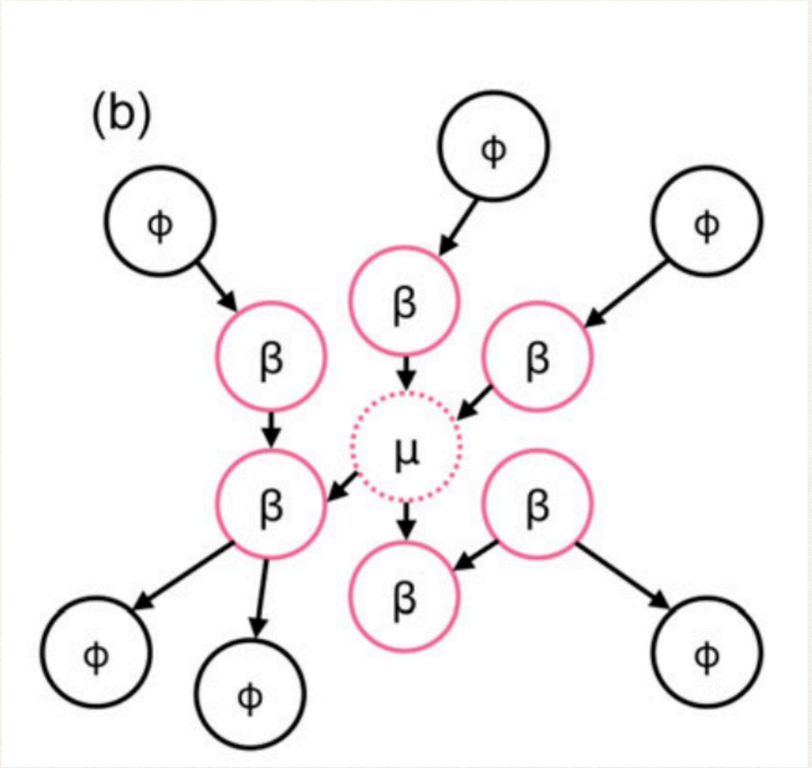


What should be conditionally independent of what given what?

See Biehl et al., 2021; Aguilera et al., 2021; Virgo et al. 2022 (or ask me at the end)

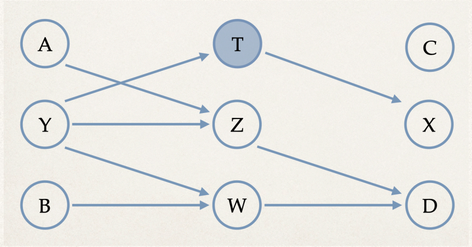
A zoo of blankets

- Unclear relation between Pearl and Friston blankets
- Inconsistencies among different definitions of Friston Blankets
- General concerns about the application of most definitions of Friston Blankets (e.g., steady-state assumption)



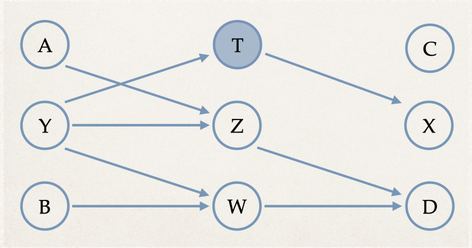
	Pearl blanket	Friston blanket	New blanket
Markov blankets as conditional independent for random variables (no time involved)	O		
Markov blankets within a Markov chain (the present shields future from past, see Pearl et al., 1989)	O	X (after Biehl et al., 2021)	
Markov blankets within a steady-state distribution (Friston, 2013, "Life as we know it")	O	O?	
Markov blankets within a stochastic process with off-block-diagonal solenoidal couplings and extra constraints (Biehl et al., 2021)	required on steady-state distribution	X (after Biehl et al., 2021)	
Markov blankets within a stochastic process from conjectured lack of off-block-diagonal solenoidal couplings (Friston et al., after 2021)	required on steady-state distribution	O?	
Asymptotic approximation to a weak-coupling equilibrium (Friston et al., 2021, "Parcels and particles: Markov blankets in the brain)	required on steady-state distribution	O?	
Causal blanket (Rosas et al., 2020)			O
History-dependent blanket (Virgo et al., 2022)			O
Standard definitions of conditional independence for stochastic processes (see our reply for a few references)			O?

2. Inference with or within a model?



Pearl blankets are used by a modeller to do inference on a system of interest with a model (~ active inference)

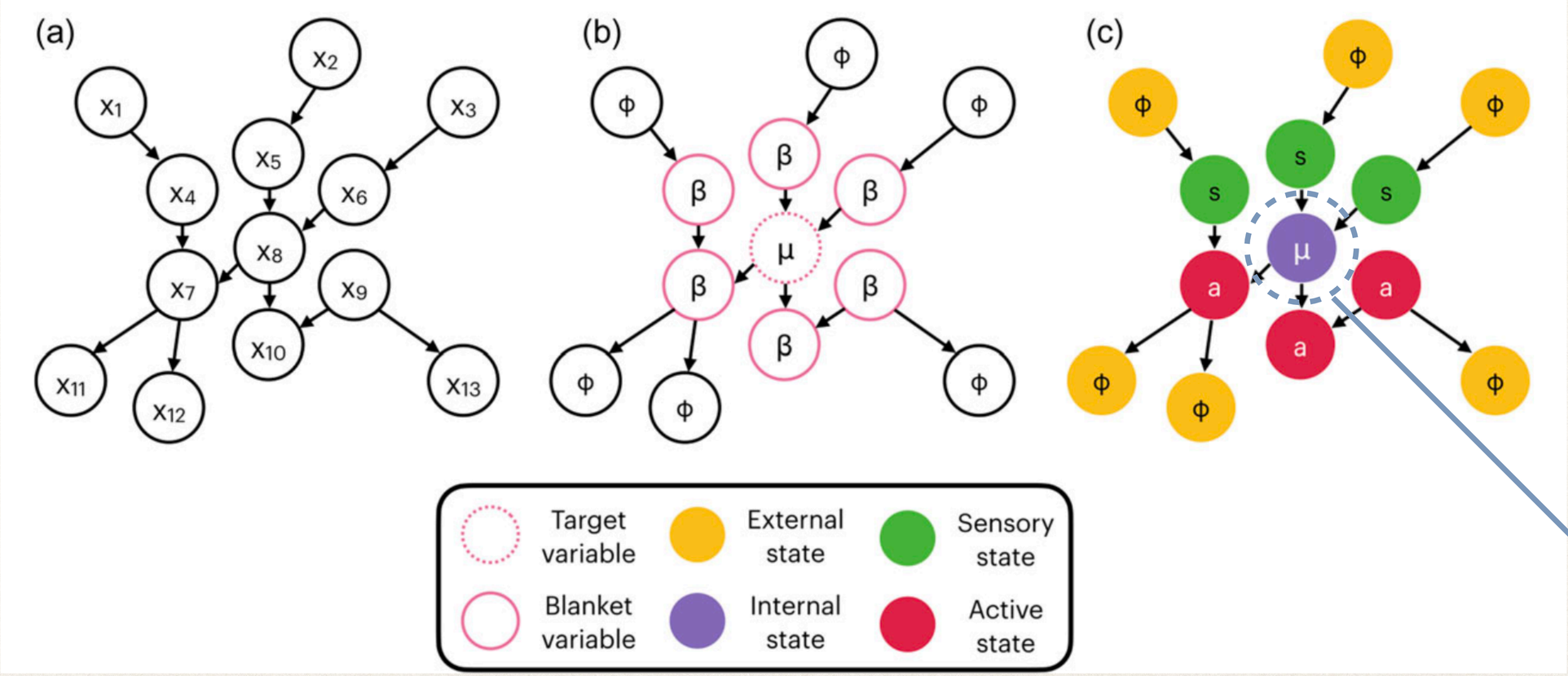
2. Inference with or within a model?



??

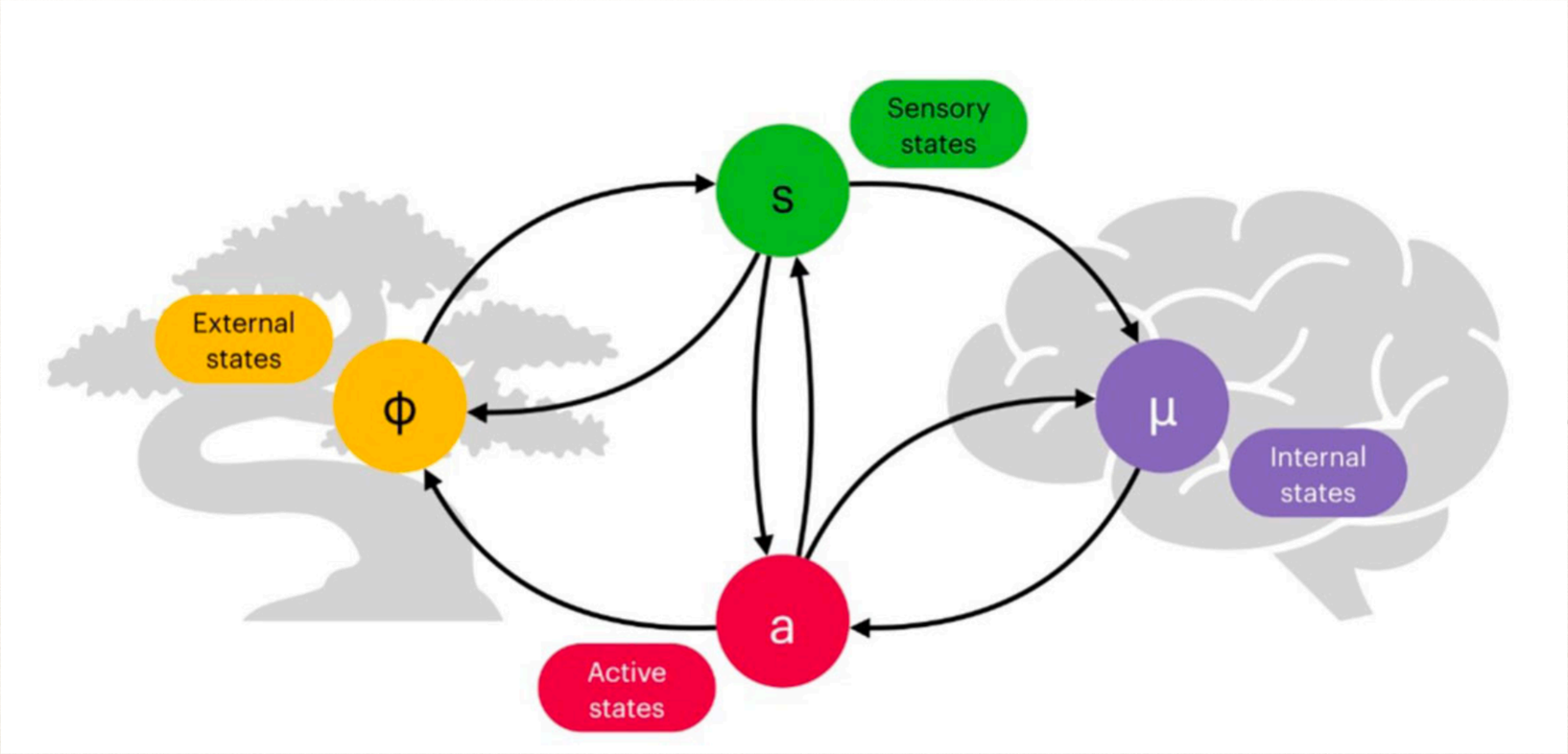
Friston blankets are claimed to define a “thing” (an agent, a mind) doing inference on everything else within a model of a system (~ free energy principle)

“Things” and particles



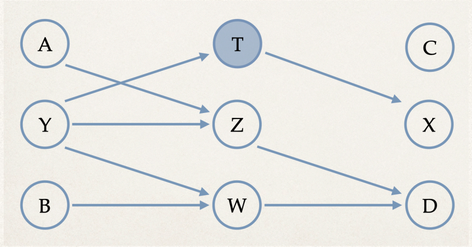
A “thing”

The usual FEP story ("things" in time)



3. Epistemic or metaphysical?

What is the relation between a model and the system of interest?



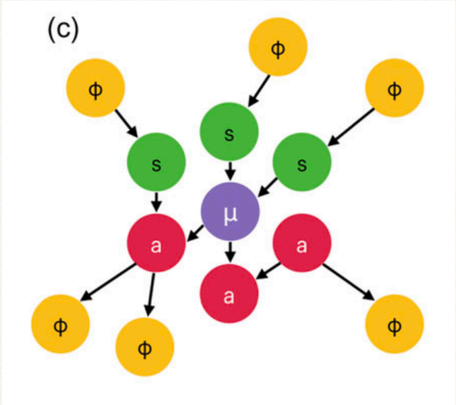
??

Do blankets exist in a model or “out there”?

3. A blanket-oriented ontology (BOO)?

"A Markov blanket **defines the boundaries of a system** (e.g., a cell or a multi-cellular organism) in a statistical sense."

"In short, the **very existence of a system depends upon** conserving its boundary, known technically as **a Markov blanket**, so that it remains distinguishable from its environment—into which it would otherwise dissipate."

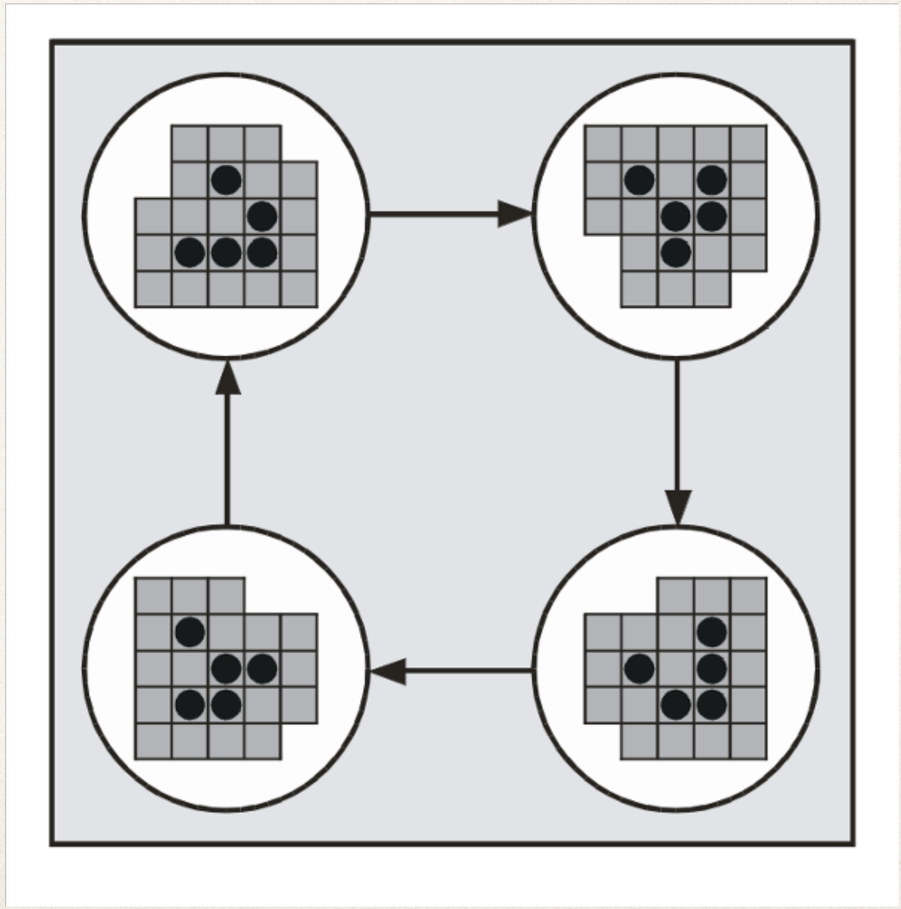


"The claims we are making about **the boundaries of cognitive systems are ontological**. We are using a mathematical formalism to answer questions that are traditionally those of the discipline of ontology, but crucially, we are not deciding any of the ontological questions in an a priori manner. The Markov blankets are a result of the system's dynamics. In a sense, **we are letting the biological systems carve out their own boundaries** [= Friston blankets] in applying this formalism. Hence, we are endorsing a dynamic and self-organising ontology of systemic boundaries."

A little excursus: a literalist fallacy?

Are we saying that Friston blankets appear to only be applied if the universe is a big Bayesian network?

> No, just noting that the mapping between (properties of) a model and (properties of) the universe is not trivial and certainly doesn't come for free when "doing the maths".



Pearl vs Friston blankets - claims

Pearl blankets

- Random variables (no time)
- (Usually) Epistemic
- Systems of interest are assumed
- Inference algorithms applied by a scientist after selecting a blanket for a modelled "thing"
- (Roughly, not exhaustively) active inference

Friston blankets

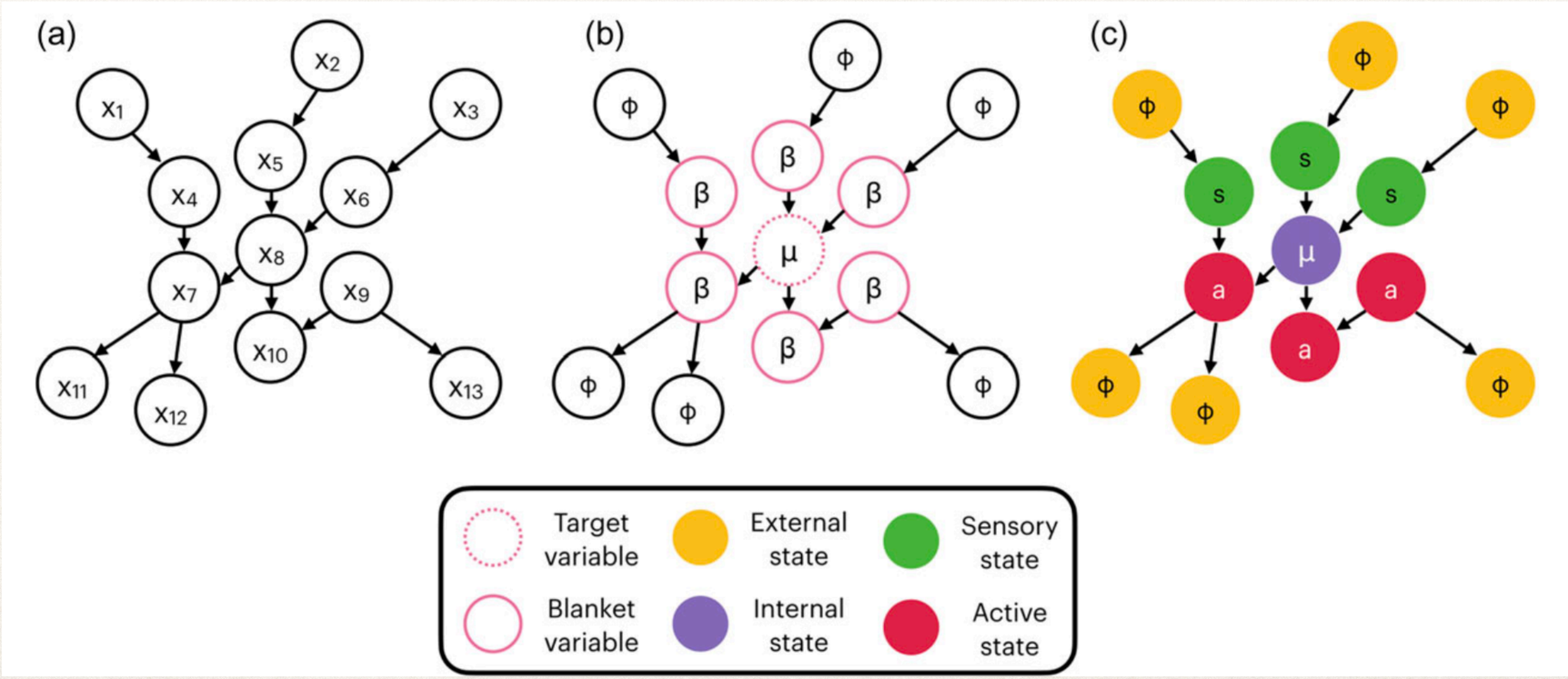
- (Stationary) Stochastic processes (time)
- (Usually) Metaphysical
- A foundational theory of "things"
- Inference emerging as the interaction between things/agents and their environments (no scientist)
- (Roughly, not exhaustively) free energy principle

The elephant in the room

Draw a Bayesian network (if it helps)

Assume a-priori a set of variables of interest (target variables)

Apply a sensorimotor interpretation



“Who tailors the blanket?” (Suzuki et al., 2022)

Pearl blankets

- Modeller chooses variables of interest
- They find its blanket
- They do inference
- …

Friston blankets

- Modeller **still chooses** variables of interest
- They find its blanket
- They claim that the chosen variables are doing inference instead of them
- …

Other possible issues

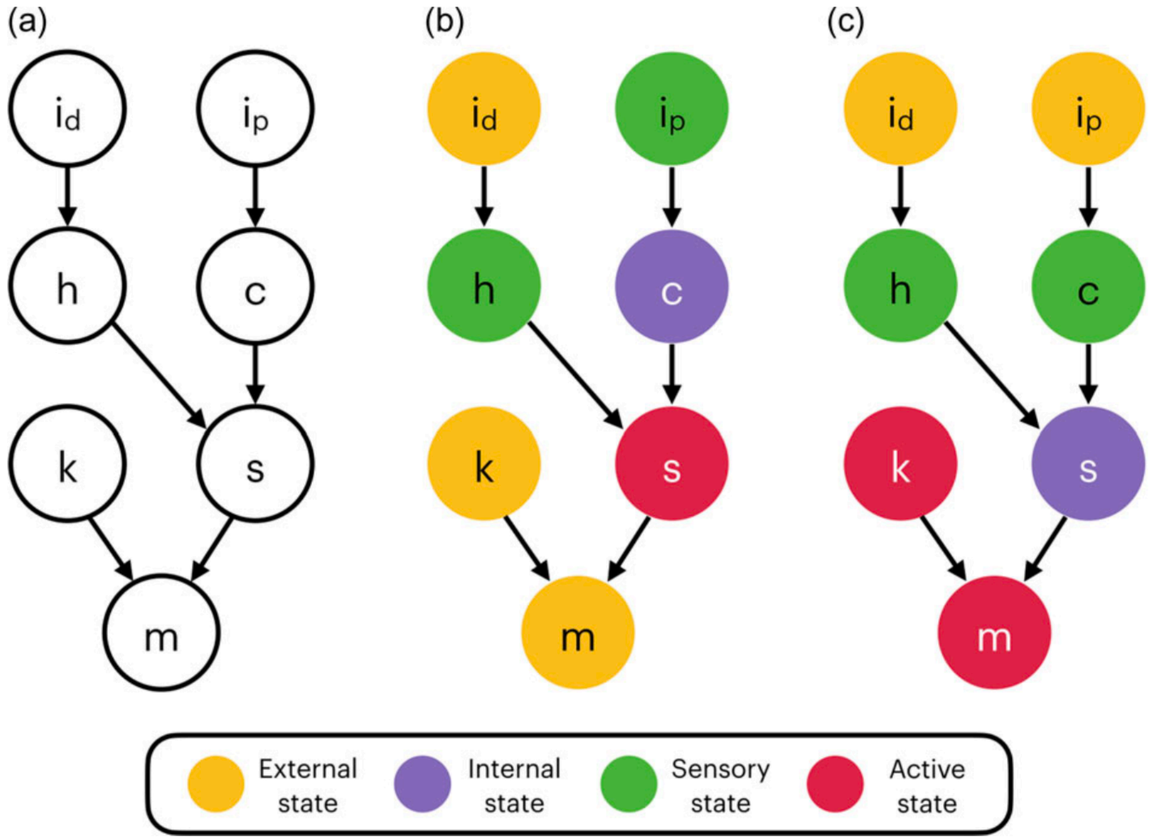


Figure 7. Conditions leading up to the knee-jerk reflex. On the left, a Bayesian network where i_d and i_p denote the motor intentions of the doctor and the patient respectively. Node *s* denotes the spinal neurons that are directly responsible for causing the kicking movement *m*. Node *h* indicates a medical intervention with a hammer, while *c* stands for a motor command sent to *s* from the central nervous system. Finally, node *k* stands for a third way of moving the patient's leg, for example, by someone else kicking it to move it mechanically. The middle (b) and the right figures (c) with the coloured-in nodes show two different ways of partitioning the same network using a "naive" Friston blanket with different choices of internal states, *c* and *s* respectively.

What's next? Or "*not everything needs to be a blanket*"

1. More clarity
2. FEP without Friston blankets
3. Active inference without the FEP
4. Beyond the FEP

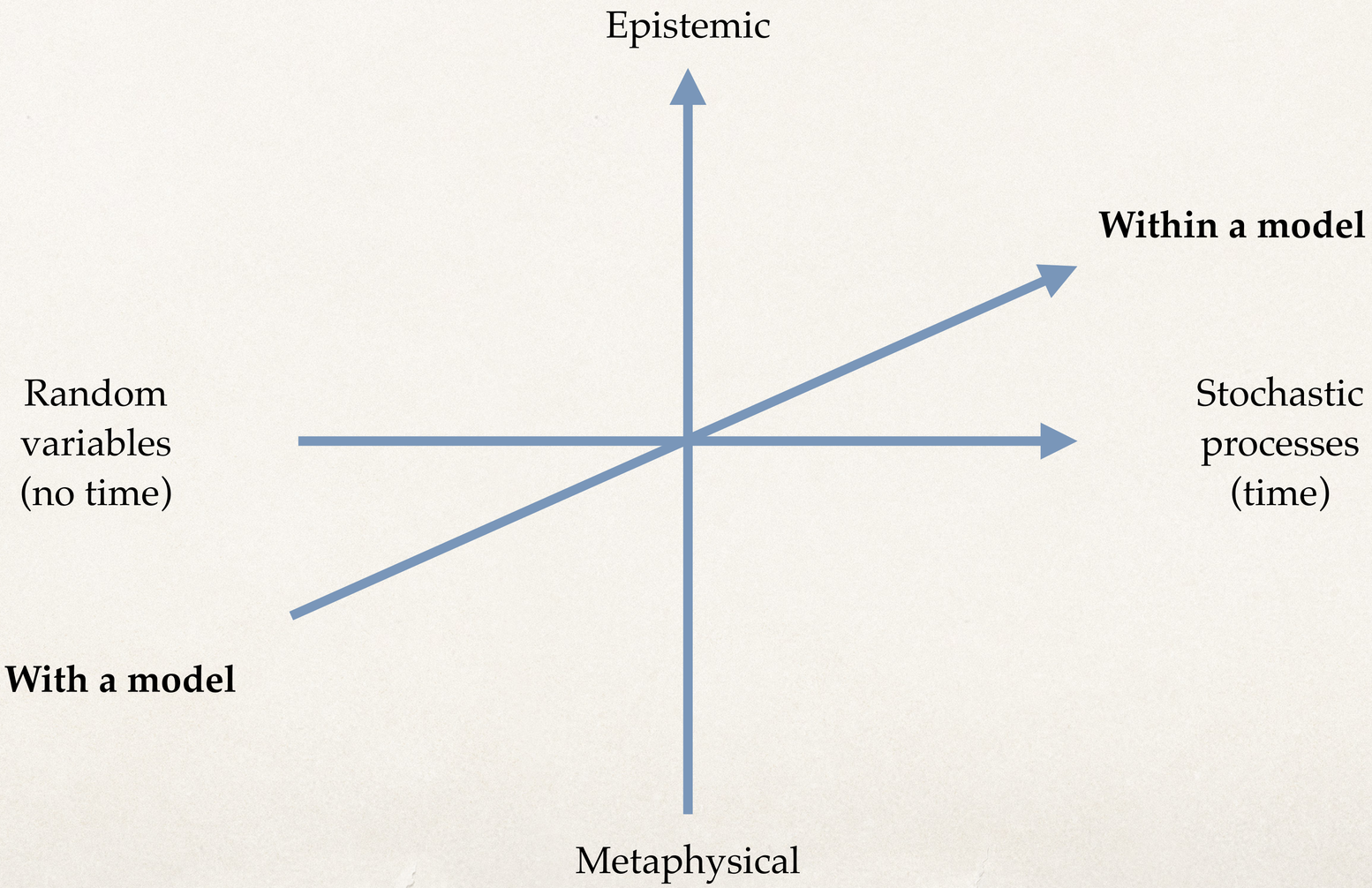
Authors' Response

The Emperor is Naked: Replies to commentaries on the target article

Jelle Bruineberg[a], Krzysztof Dołęga[b], Joe Dewhurst[c] and Manuel Baltieri[d,e]

[a]Department of Philosophy, Macquarie University, Sydney, NSW 2109, Australia; [b]Institut für Philosophie II, Fakultät für Philosophie und Erziehungswissenschaft, Ruhr-Universität Bochum, 44801 Bochum, Germany; [c]Fakultät für Philosophie, Wissenschaftstheorieund Religionswissenschaft, Munich Center for Mathematical Philosophy, Ludwig-Maximilians-Universität München, 80539 Munich, Germany; [d]Araya, Inc., Tokyo, Japan and [e]School of Engineering and Informatics, University of Sussex, Brighton BN1 9RH, UK
jelle.bruineberg@mq.edu.au
krzysztof.dolega@rub.de
joseph.e.dewhurst@gmail.com
manuel_baltieri@araya.org

1. More clarity



2. FEP without Friston blankets

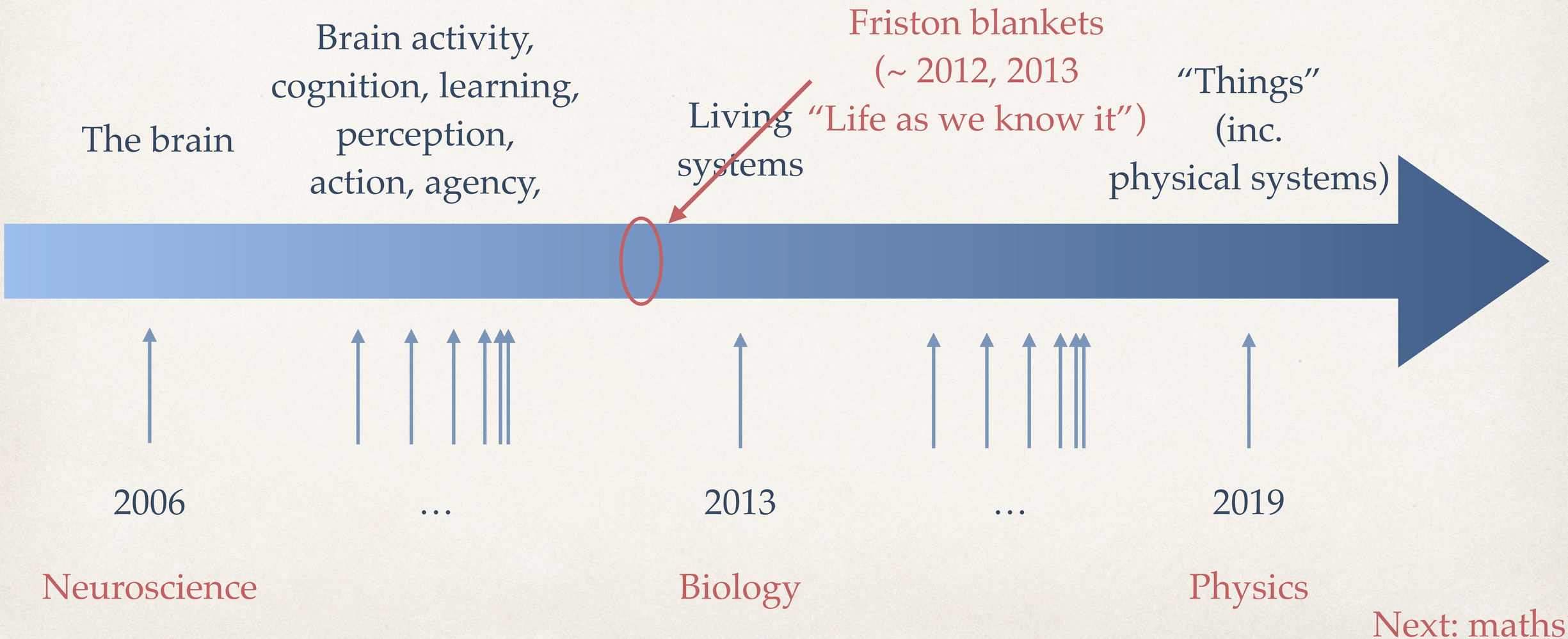
With them:

- Stationarity is required
- Inference/learning must happen away from stationary state

Without them:

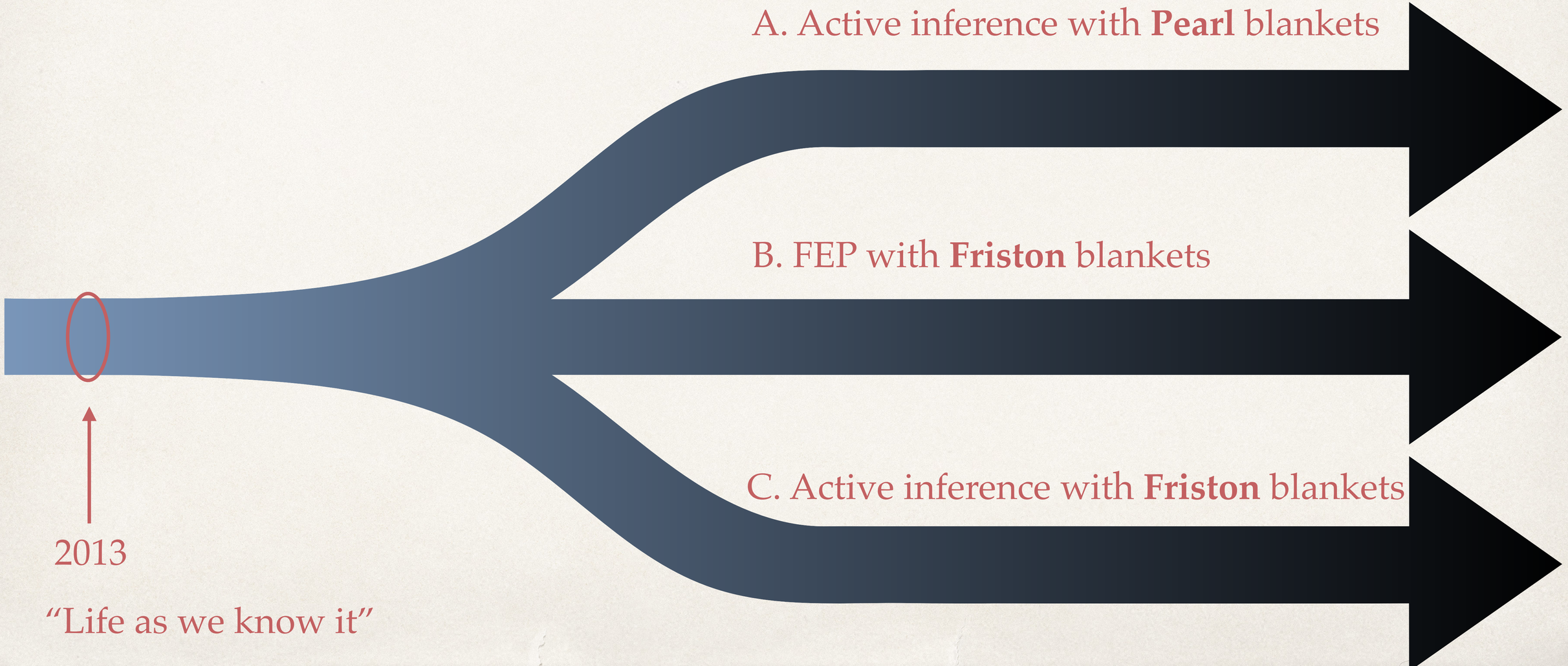
- No obvious partition between internal/external?
- No “Approximate Bayesian inference lemma”?
- No FEP?

3. Active inference without the FEP?

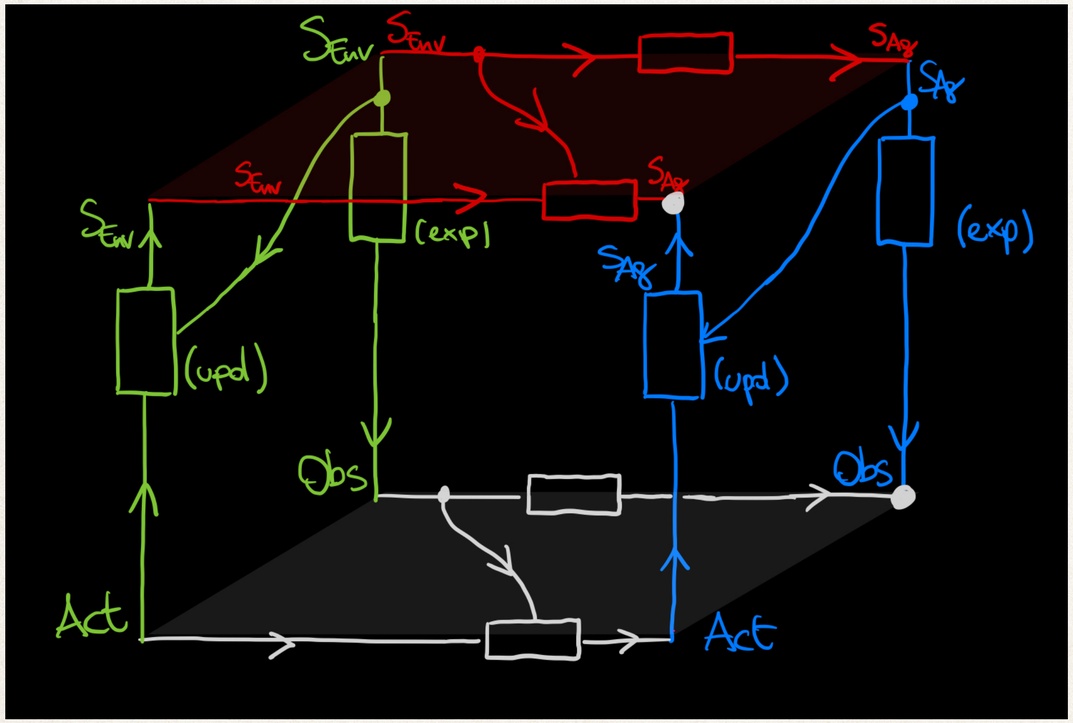


Next: maths?

3. Active inference without the FEP



4. Beyond the FEP



Interpreting Dynamical Systems as Bayesian Reasoners

Nathaniel Virgo[1][0000−0001−8598−590X], Martin Biehl[2][0000−0002−1670−6855], and Simon McGregor[3]

[1]Earth-Life Science Institute, Tokyo Institute of Technology, Tokyo 152-8550, Japan
[2]Araya Inc., Tokyo 107-6024, Japan
[3]University of Sussex, Falmer, UK

Abstract. A central concept in active inference is that the internal states of a physical system parametrise probability measures over states of the external world. These can be seen as an agent's beliefs, expressed as a Bayesian prior or posterior. Here we begin the development of a general theory that would tell us when it is appropriate to interpret states as representing beliefs in this way. We focus on the case in which a system can be interpreted as performing either Bayesian filtering or Bayesian inference. We provide formal definitions of what it means for such an interpretation to exist, using techniques from category theory.

Keywords: Bayesian filtering · Bayesian Inference · Category Theory.