

# Reasoning: Probability, Coherence and AI

Paul Thagard

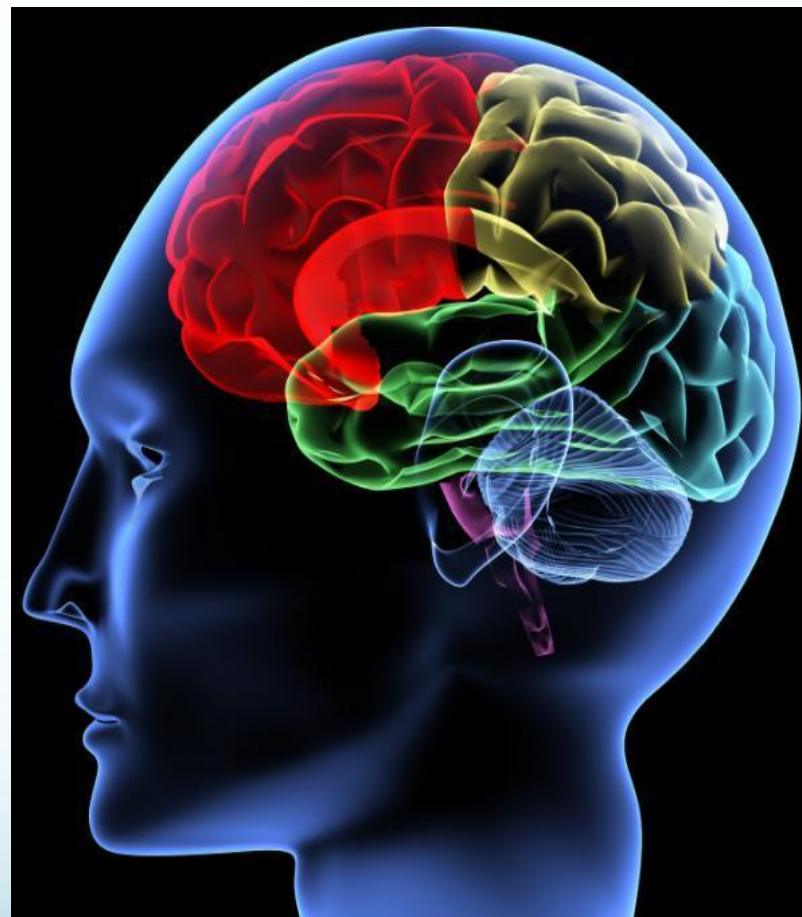
[paulthagard.com](http://paulthagard.com)

June, 2026



# Outline

1. Reason: belief and action
2. Medical examples
3. Probability
4. Coherence
5. Artificial intelligence
6. Conclusions



# *Reason: Belief and Action*

	<b>Descriptive</b>	<b>Normative</b>
What to believe?	How do people diagnose a cold?	How should people diagnose a cold?
What to do?	How do people treat a cold?	How should people treat a cold?

# GLP-1 Drugs



**1983:** Glucagon-like peptide-1 (GLP-1) identified in the gut, stimulates insulin

**2017:** *Semaglutide (Ozempic)* — once-weekly injection with high efficacy for glucose control and weight loss.

**2020s:** Expansion to obesity, cardiovascular, and neural diseases. Many similar drugs.

100 million users, 3 million in Canada, 1/8 American adults. Belief, action, descriptive, normative.

# Probability: History



1654. Pascal and Fermat invent probability theory to explain games of chance.

1738. Daniel Bernoulli explained decision making as based on expected utility.

1761. Thomas Bayes formulated a theorem that updates the probability of a hypothesis based on evidence.

1926. Frank Ramsey advocated understanding probabilities as subjective degrees of belief.

1944. John von Neumann and Oskar Morgenstern formalized decision making as maximizing expected utility.

1988. Computer scientist Judea Pearl develops AI models of probabilistic causal networks.

21<sup>st</sup> century. Psychologists such as Josh Tenenbaum, Tom Griffith, and Alison Gopnik advocate Bayesian explanations of human beliefs.

# Probability: Belief

**Bayes' Theorem:**  $P(\text{hypothesis} | \text{evidence}) =$   
 $P(\text{hypothesis}) * P(\text{evidence} | \text{hypothesis}) /$   
 $P(\text{evidence}).$

$P(\text{GLP-1s effective}/\text{studies}) =$   
 $P(\text{GLP-1s effective}) * P(\text{studies}/\text{GLP-1s eff.})$   
 $/$   
 $P(\text{evidence}) =$

5 \* 8 / 8 = 5      ??????????

# Probability: Action

Expected utility (action) =  $p_1u_1 + p_2u_2 + \dots + p_iu_i$  , where  $p_i$  is the probability of outcome  $i$  and  $u_i$  is the utility of outcome  $i$ .

Do the action if it has higher expected utility than the alternatives.

E.g. adopt GLP-1s if they have high expected utility compared to others.

Descriptive and normative?

# Problems for Probability

1. Practical: we rarely know the probabilities and utilities.
2. Psychological: people reason badly with probabilities and utilities (Kahneman & Tversky).
3. Computational tractability and availability.
4. Philosophical: interpretation of probability.
5. Philosophical: objectivity of probability.
6. Philosophical: ethics and normativity, e.g. justice.

# Coherence



Coherence is **making sense** of a complex situation, e.g. the sentence “The pen is in the bank.” Old/young woman.

Coherence is a computational process of **constraint satisfaction**. Positive constraints hold between elements that fit together, while negative constraints hold between incompatible elements.

Coherence is naturally **computed** by neural networks via excitatory and inhibitory links.

# Coherence: History



1807. Georg Hegel introduces coherentist epistemology with the doctrine that “the truth is the whole”.

Early 1900s. British idealists such as F. H. Bradley and Bernard Bosanquet advocate coherence as the basis of knowledge.

Mid-1900s. Gestalt psychologists such as Wolfgang Köhler and Leon Festinger emphasize coherence ideas such as grasping wholes and cognitive dissonance.

Mid-1900s. Philosophers such as Wilfred Sellars, W. V. O., Gilbert Harman, and Keith Lehrer emphasize coherence of beliefs.

1989. Thagard formalizes coherence as constraint satisfaction and uses computational models to show how to evaluate it.

2000s. Psychologists including Keith Holyoak, Stephen Read, Dan Simon, and Matthew Lieberman use coherence as constraint satisfaction to explain psychological phenomena.

# Coherence: Belief

Hypotheses are evaluated by explanatory coherence.

Positive constraints: how much explained.

Negative constraints: competing hypotheses.

We should believe that GLP-1s are effective because that is the best explanation of many rigorous studies.

Many applications in science, medicine, law, and ordinary life.

# Coherence: Action

Hypotheses are evaluated by deliberative coherence.

Positive constraints: how much actions accomplish goals.

Negative constraints: competing actions.

We should adopt GLP-1s as treatments because studies show that they are more effective than alternatives.

# Coherence: Problems

1. Is coherence really better descriptively and normatively for belief and action than probability?
2. Coherence is too vague: no – specify constraints and algorithms.
3. Coherence does not guarantee truth and optimality.
4. Is coherence computationally tractable?



# Generative AI

2022, ChatGPT 3.5. 2024, ChatGPT 4o.

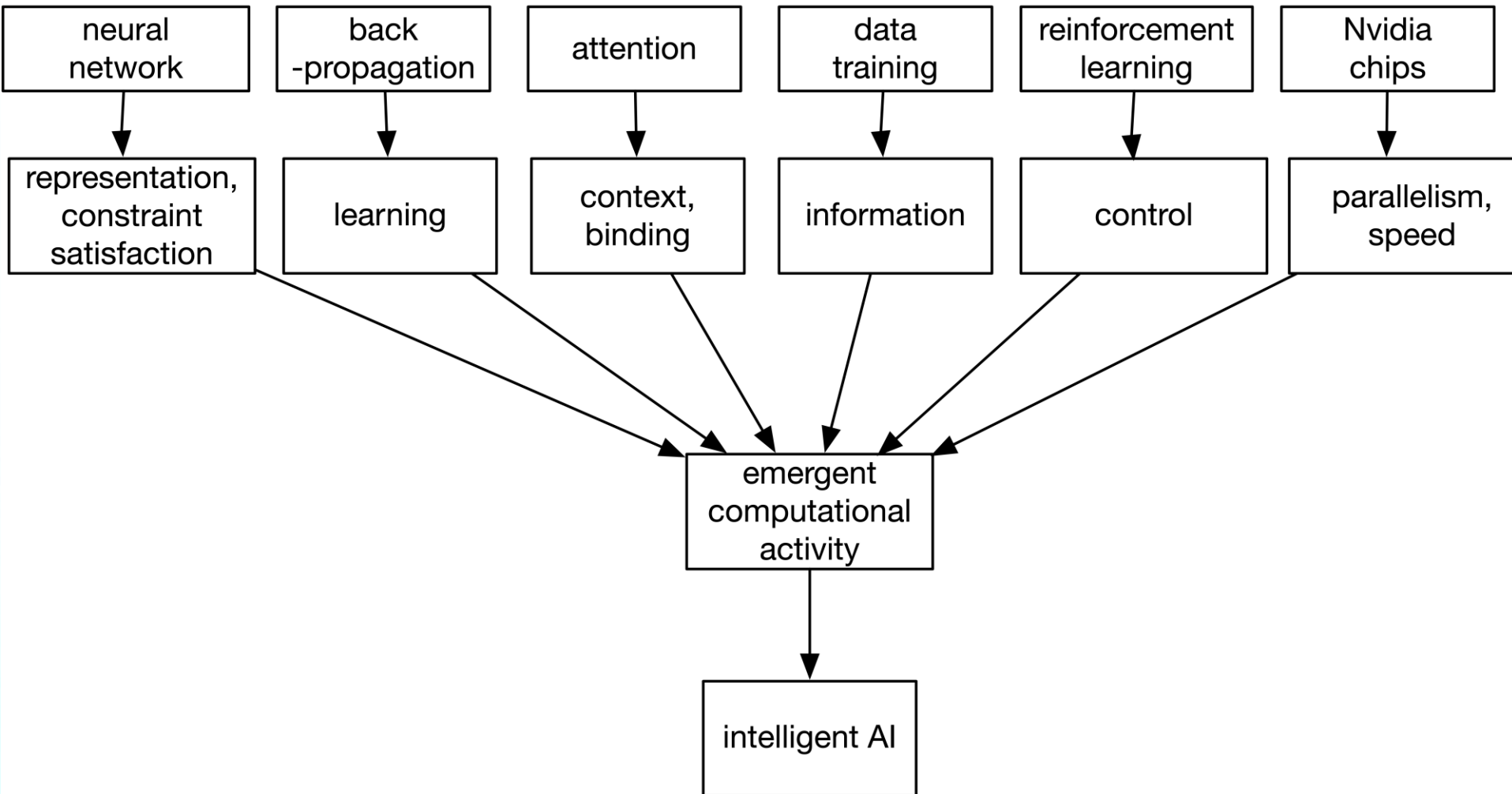
2026, ChatGPT 5.5, Codex. 900 million users.

**Similar:** Google Gemini, Anthropic Claude, Meta Llama, xAI Grok, Mistral le Chat, DeepSeek.

Do the benefits of these models outweigh the risks? Thagard, *Boom or Doom?* MIT Press,

Why are these models so powerful?

# Power = mechanisms + causal networks + emergence



# AI and GLP-1 Drugs (1)

ChatGPT, Claude, Gemini, and Grok are all **excellent** at reasoning about GLP-1s, for belief and action, and for obesity and depression. How do they do this?

ChatGPT says that GLP-1 drugs are effective treatments based on a combination of (1) empirical induction from clinical trial results about treatments and (2) an abductive inference that the best explanation of these results is that GLP-1 drugs are effective.

# AI and GLP-1 Drugs (2)

ChatGPT, Claude, Gemini, and Grok can also reconstruct their reasoning in terms of probability and coherence.

Grok says that it reaches its conclusions using both probability and coherence.

ChatGPT even writes LISP code for modeling GLP-1 drug effectiveness in ECHO (explanatory coherence).

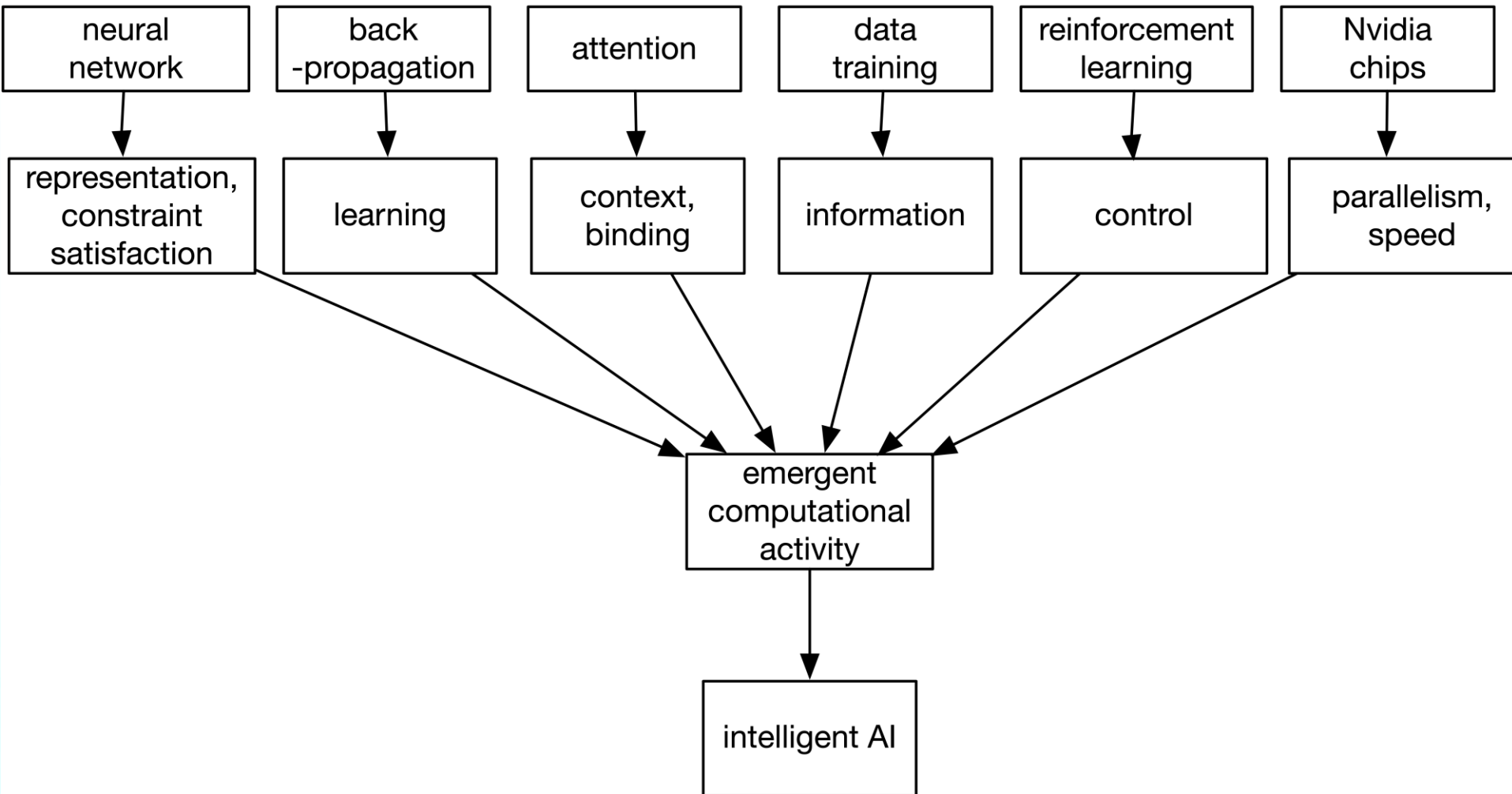
# AI and GLP-1 Drugs (3)

But how are AI models actually reasoning? Their introspection is unreliable. ChatGPT, Claude, and Grok recommend combining probability and coherence, but how?

Conjecture: AI models have emergent intelligence that we do not yet understand.

Ethical conclusion: Do not allow these models to make important decisions, e.g. bombing targets.

# Power = mechanisms + causal networks + emergence



# Micro/Macro Emergence

Why does change occur? MME

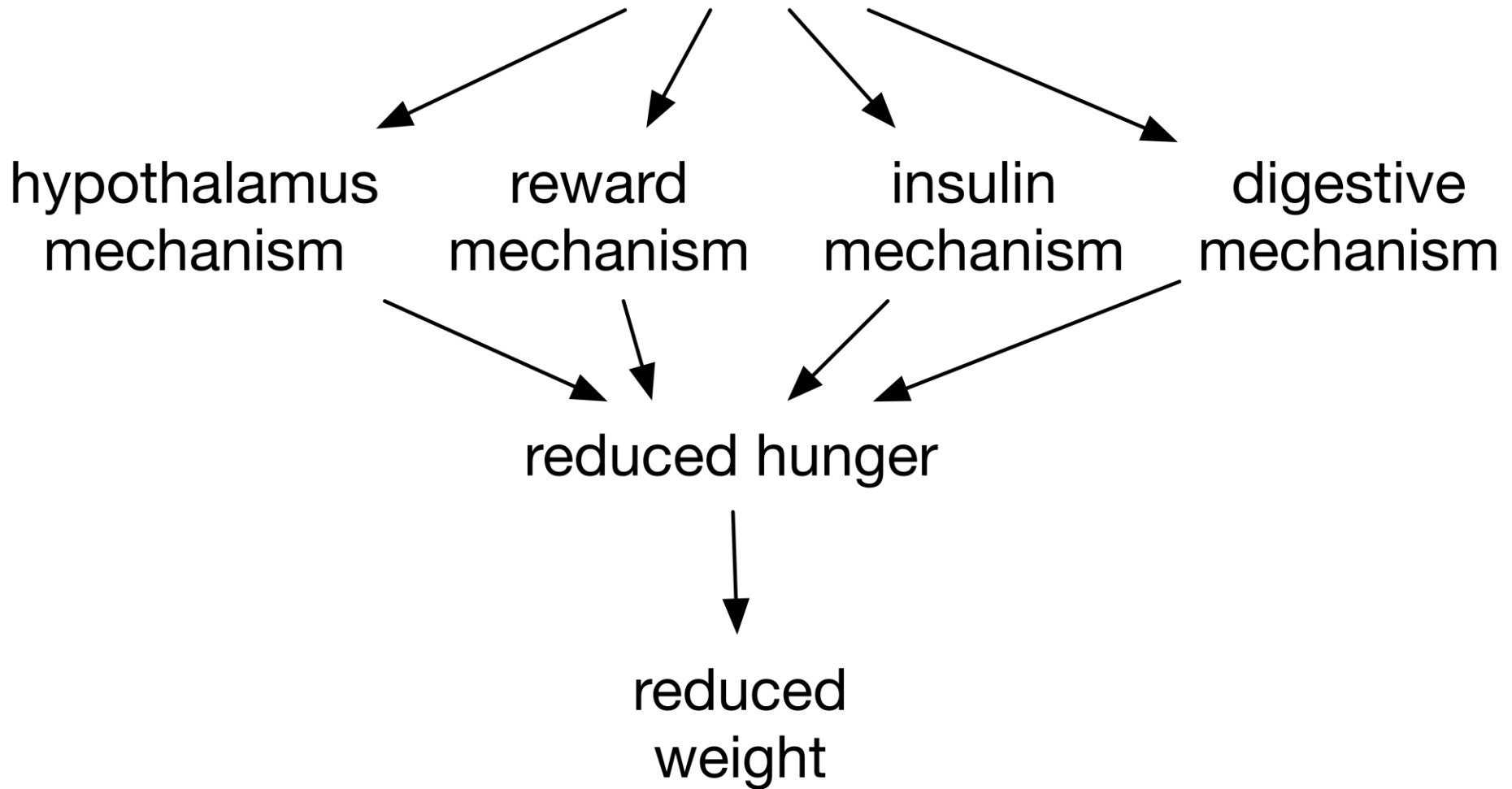
**Micro-mechanisms** describe interactions of parts to produce regular changes.

**Macro-mechanisms** describe causal networks of interacting mechanisms.

Networks result in feedback, resilience, tipping points, and **emergence**.

Applies to consciousness, AI (deep learning + attention mechanism), desire, habits, disease.

# GLP-1 drugs



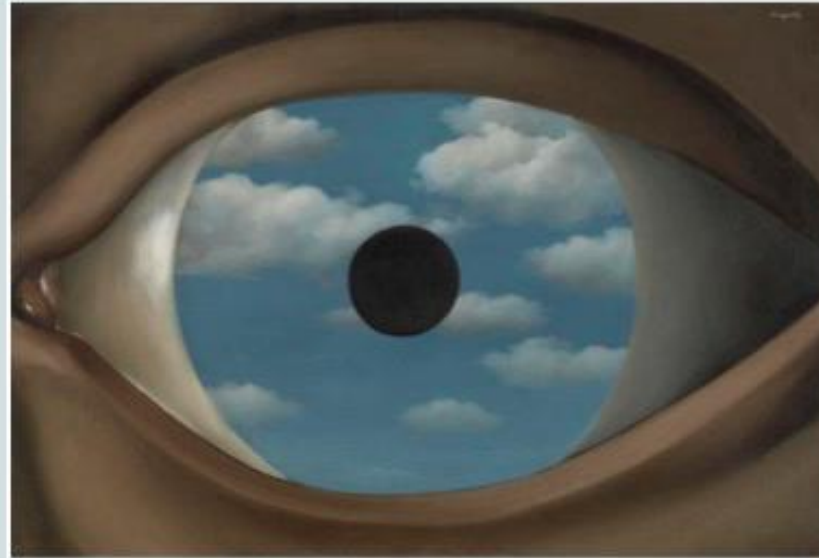
# Conclusions

**Coherence** is a strong descriptive and normative alternative to probabilistic accounts of human reasoning about belief and action.

**AI** reasoning can present itself as probabilistic or coherentist, but its actual reasoning is not well understood, even by itself.

**Micro/macro emergence** explains AI reasoning and other powerful effects, e.g. consciousness, disease, addiction.

OXFORD



# DREAMS, JOKES, AND SONGS

HOW BRAINS BUILD CONSCIOUSNESS

---

PAUL THAGARD