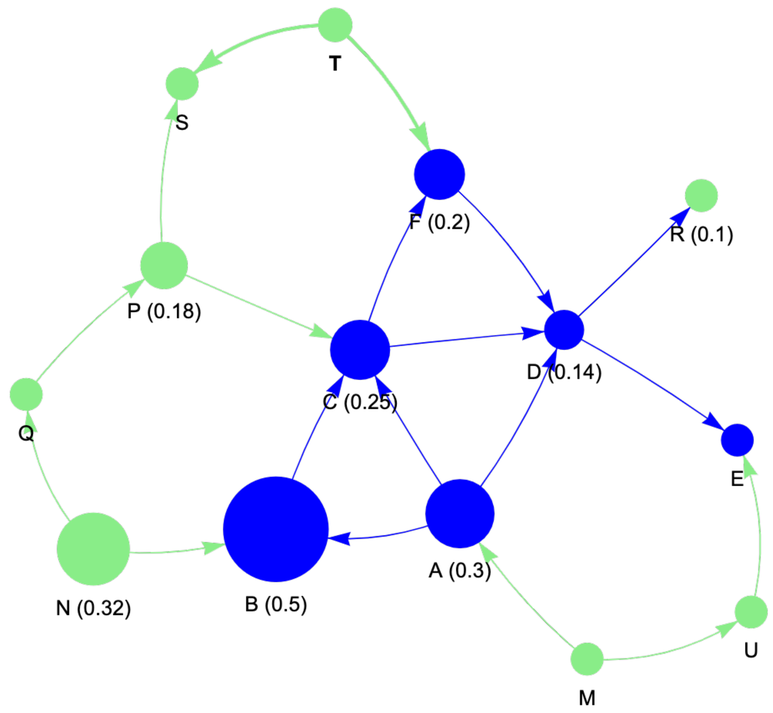
System 3: a mental simulation module for estimating the causal structure of the world

The literature

- Schelling, *The Mind as a Consuming Organ*: brains generate their own reward
- Ainslie, Loewenstein, Kimball: much of consumption is mental
- Sloman & Lagnado: people represent the world as a causal network
- Schultz, Dayan, Montague: dopamine activation ‘migrates’ back along a causal chain
- Gopnik: children easily learn causal relations
- Sutton & Barto, Rescorla-Wagner: RL predicts then updates reward
- Daw: model-based more tractable vs model-free
- K&T: mental simulation heuristic
- Pham: HDIF heuristic
- Hamrick: mental simulation in AI solves problems well
- Dayan, Stachenfeld: the successor representation



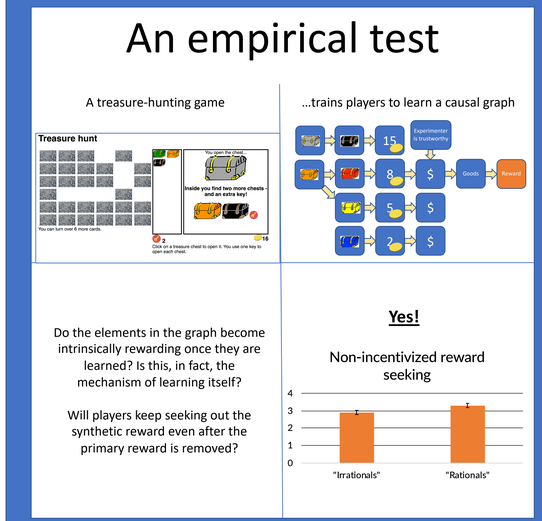
The causal subgraph (blue) is embedded in a larger implication graph (green).

Learners cannot easily distinguish cause from implication; mental simulation helps them do so.

Key claims of this theory:

1. Not just causal, also implication relations
2. Synthetic reward generated at every node
3. Reinforcement learning updates the graph weights during simulation
4. Truncation prevents perfect updating: synthetic reward gets ‘stuck’ e.g. liking money for its own sake
5. Motivates consumption of fiction, mind-wandering, prospection, replay, empathy
6. Decisions over the graph are neither a Type 1 nor 2 process

An empirical test



Think of Pac-man running through the maze picking up power pills. In a similar way, your brain traverses the implication graph gathering reward.

Each time a node is active, it predicts how much reward will be found at the next node. The graph weights update according to how good the guess was. Eventually, the predicted rewards will become accurate.

This learning process has the delightful side-effect of generating lots of mental reward. This can be called ‘cognitive goods production’.