Computational Models Applied to Various Philosophical Topics

Nathan Gabriel

University of California Irvine

September 1, 2023

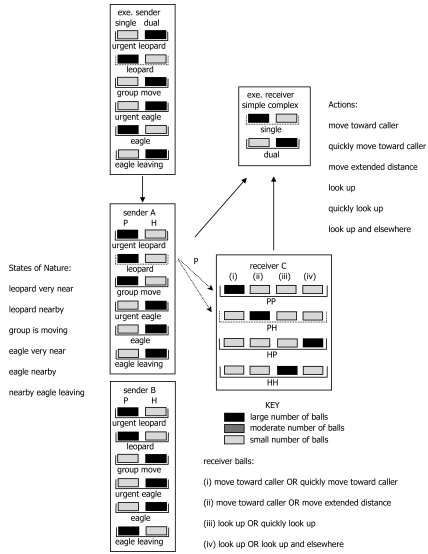
Compositionality

Putty-nosed monkeys

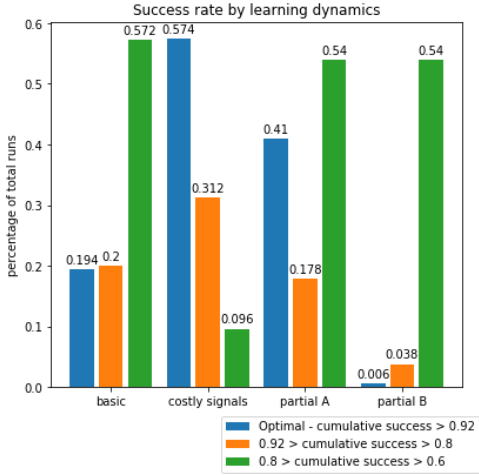
Putty-nosed monkeys have four general categories of alarm calls:

- P-sequences correspond to the presence of leopards
- H-sequences correspond to the presence of eagles
- PH-sequences correspond to group movement
- HP-sequences correspond to a formerly threatening eagle becoming more distant
- All four types of alarm call can vary based on urgency
  - P is low urgency PPPPPPPP is high urgency, PH is low urgency PPPHHH is more urgent PPPPPPPHHHHHH high urgency (or further movement)

The Pyow-hack Game



Simulation Results



Illusory Truth Effect

- Three people make a tiger:
  - Advisor to king of Wei accompanies the king's son to Handan.
  - Advisor's political rivals speak ill of him while he is away.
  - Advisor is ostracized when he returns.
- The illusory truth effect is exhibited when repeated exposure to a statement increases an individual's credence in that statement.

Simulation Results

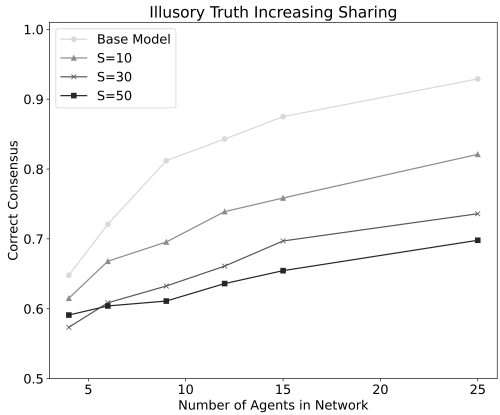


Figure: Average correct consensus rates for 2000 simulations with $M = 100$, and $D = 5$.

Robustness

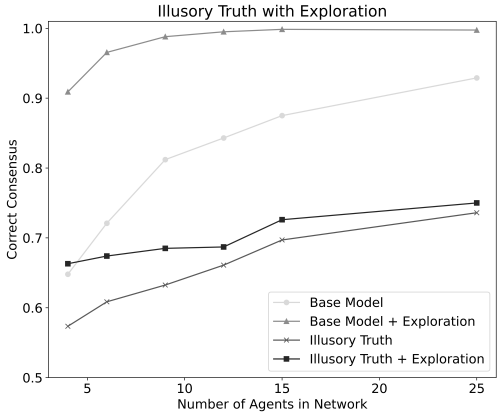
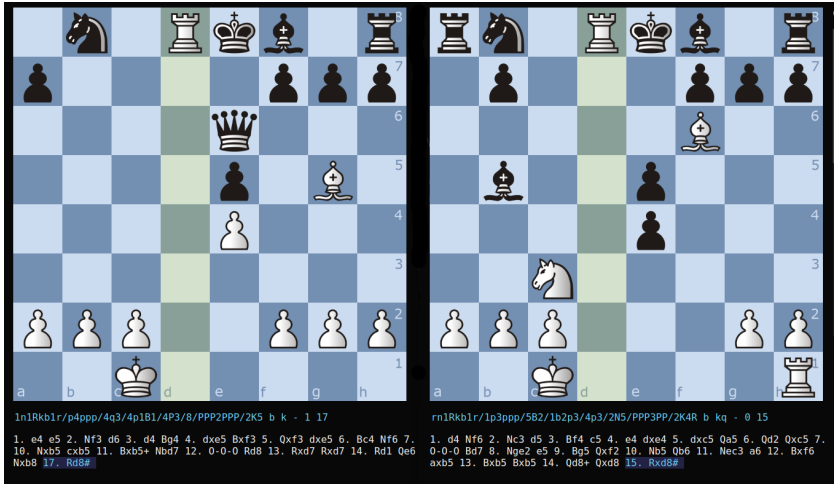


Figure: Average correct consensus rates for 2000 simulations of ϵ-greedy agents with $\epsilon = 0.001$, $M = 100$, $S = 30$, and $D = 5$.

Abstraction



1n1Rkb1r/p4ppp/4q3/4p1B1/4P3/8/PPP2PPP/2K5 b k - 1 17

1. e4 e5 2. Nf3 d6 3. d4 Bg4 4. dxe5 Bxf3 5. Qxf3 dxe5 6. Bc4 Nf6 7. 10. Nxb5 cxb5 11. Bxb5+ Nbd7 12. O-O-O Rd8 13. Rxd7 Rxd7 14. Rd1 Qe6 Nxb8 17. Rd8#

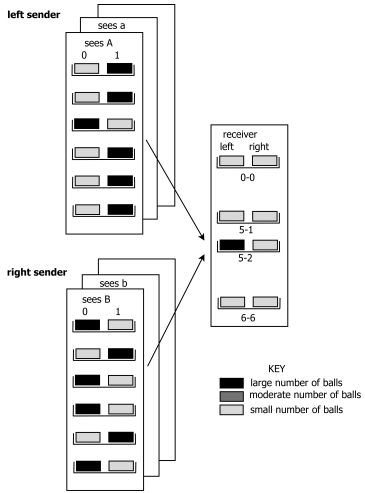
rn1Rkb1r/1p3ppp/5B2/1b2p3/4p3/2N5/PPP3PP/2K4R b kq - 0 15

1. d4 Nf6 2. Nc3 d5 3. Bf4 c5 4. e4 dxe4 5. dxc5 Qa5 6. Qd2 Qxc5 7. O-O-O Bd7 8. Nge2 e5 9. Bg5 Qxf2 10. Nb5 Qb6 11. Nec3 a6 12. Bxf6 axb5 13. Bxb5 Bxb5 14. Qd8+ Qxd8 15. Rxd8#

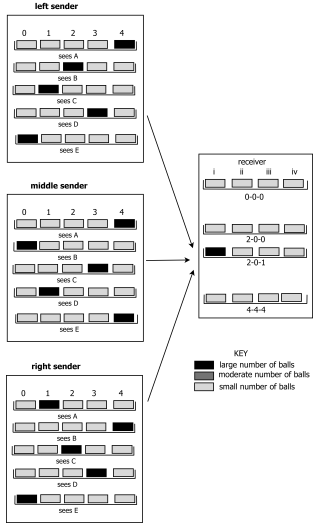
Transitive Inference

- serial ordering: A, B, C, D, E
- a second Serial ordering: a, b, c, d, e
- Experimental paradigm:
  - Five objects/stimuli are arranged in a serial order: *A*, *B*, *C*, *D*, *E*.
  - The subject to be tested is conditioned on adjacent pairs in the serial ordering through rewards for selecting the first object in a presented pair. (E.g. If presented with *A*-*B* the agent is rewarded for choosing the item on the left. If presented with *D*-*C* the agent is rewarded for choosing the item on the right.)
  - After the agent has learned to respond correctly to pairs of objects that are adjacent in the serial order, she is then tested on the non adjacent pairs *B*-*D* and *D*-*B*.

Transitive Inference Model



Nonsense Grammar Model



Questions?

Thank you