Simulating bounded rationality: Optimality modelling without an optimality commitment

Jason Noble
University of Southampton

Outline of the argument

- Evolved agents are boundedly rational.
- Do tools like game theory and optimization methods for agent-based models imply a rationality assumption? Can we still use them?
- Yes we can: such tools are about identifying which of a set of behavioural alternatives is best in a given context.
- They're effective ways of making contact between the two blades of the scissors: agent and environment.

Unofficial summary (why I'm really here)

- Student of fields such as agent-based modelling, behavioural ecology, artificial life, evolutionary psychology and evolutionary robotics.
- Bounded rationality sits easily: never indoctrinated into traditions like mathematical logic or economics where unbounded rationality holds sway.
- Concerned about link between bounded rationality and cognitive science as totally historical and contingent.
- This threatens to invalidate some of the tools for modelling and theorizing that I find most useful.

The myth of the rational maximizer

- The rational agent as something close to a Laplacean demon.

- Familiar story: unlimited computational resources for considering and calculating over alternatives.

- A cluster of related (and equally implausible) ideas:
  - Fully rational agents, however produced, are maximizers.
  - Evolution can produce fully rational maximizing agents.
  - Evolution produces agents that can be treated as if they were rational maximizers.

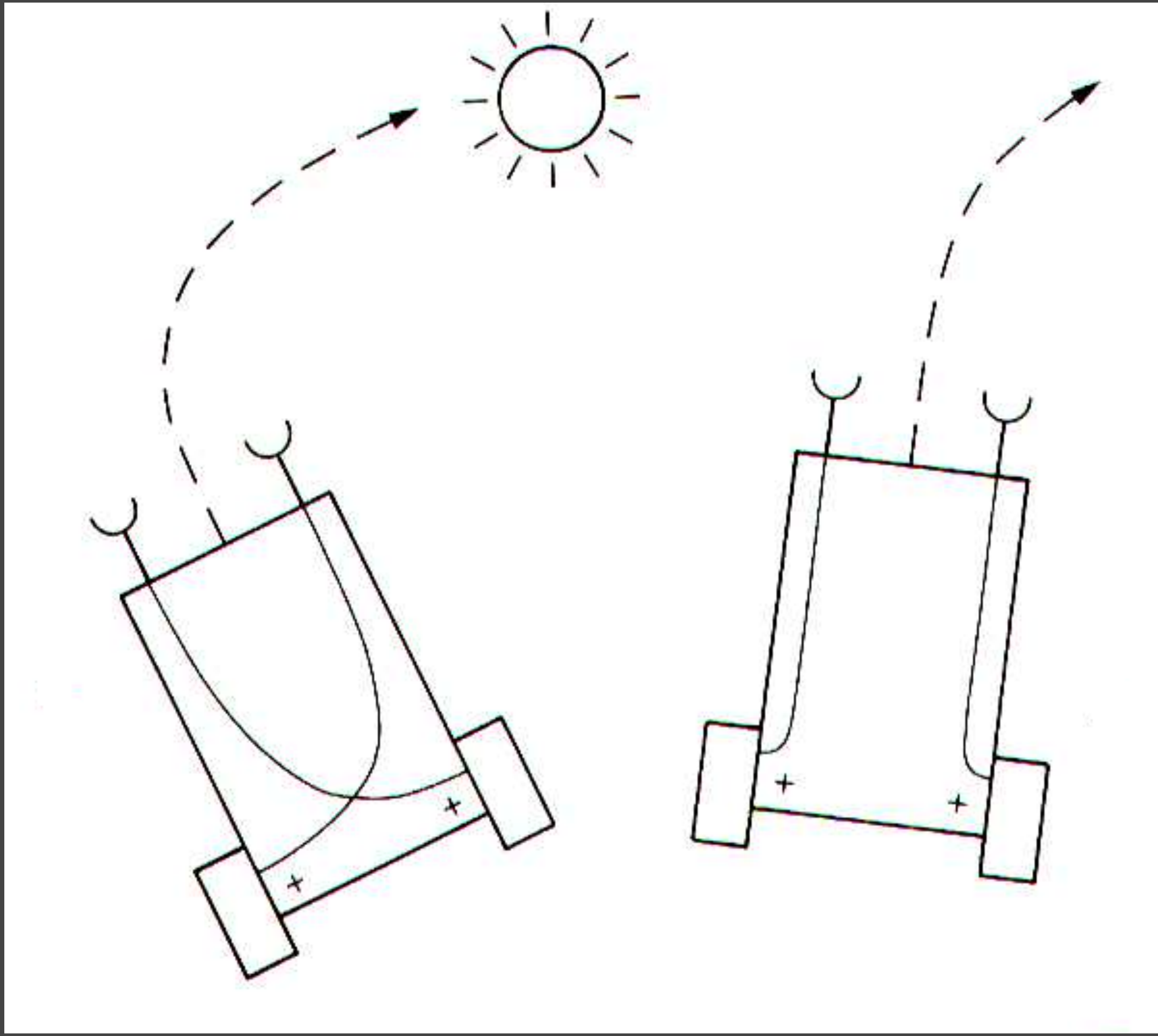
Some tools for modelling behaviour

- Rational choice theory: what is the best policy given a utility function and some alternatives?
- Game theory: what is the best policy given a utility function, some alternatives, and the assumption that one's opponent is also rationally maximizing their utility over alternatives?
- Agent-based simulation models: extensions of the above into arbitrarily complex environments and sets of strategic options. Optimization through learning, evolution, or both.
- Are these tools fatally wedded to the notion of the agent as rational maximizer?

Bounded rationality

- The official version via Herbert Simon: it's impossible for real agents to be rational maximizers as the required resources are not available.

- The same conclusions can be reached from quite different routes.
  - Valentino Braitenberg's *Vehicles*: a neuroscientist's fable about how simple mechanisms can produce apparently complex behaviour.
  - Evolutionary robotics: getting meaningful robot behaviour from random starting points plus a selection process.



Braitenberg vehicles: a reminder that purposeful behaviour may well be the product of an accumulated bag of tricks

Where does this leave cognitive science?

- If we throw out the myth of the rational maximizer, must the tools go too?
- Contingency and history are important, but is cognitive science reduced to ethnography?
- If the tools are viewed as doing local search in behaviour-space, and not as part of a global rationality commitment, we are OK.

- Why was the rational-maximizer view so popular anyway? Like the clockwork metaphor and later the computer metaphor for the mind, a new tool was mistakenly applied across too broad a range of phenomena.

- Externalism / internalism: the language of environmental influence pitched against the language of constraints. Agreeing that natural cognizers will only be boundedly rational doesn't entail a jump all the way over to an internalist and historical perspective.

- See also Godfrey-Smith's *Complexity and the Function of Mind in Nature*

- Paradoxically, the limitations of the methods save them from being tied to a global rationality perspective. Maynard Smith noted in *Evolution and the Theory of Games* that game theory can easily respect genetic and other constraints because the set of available strategies must be explicitly specified.

- What sort of explanation are we looking for? An account of the causal interplay between organism and environment: both blades of the scissors.

- Best illustrated with examples

The Norway rat

An example: social learning in rats

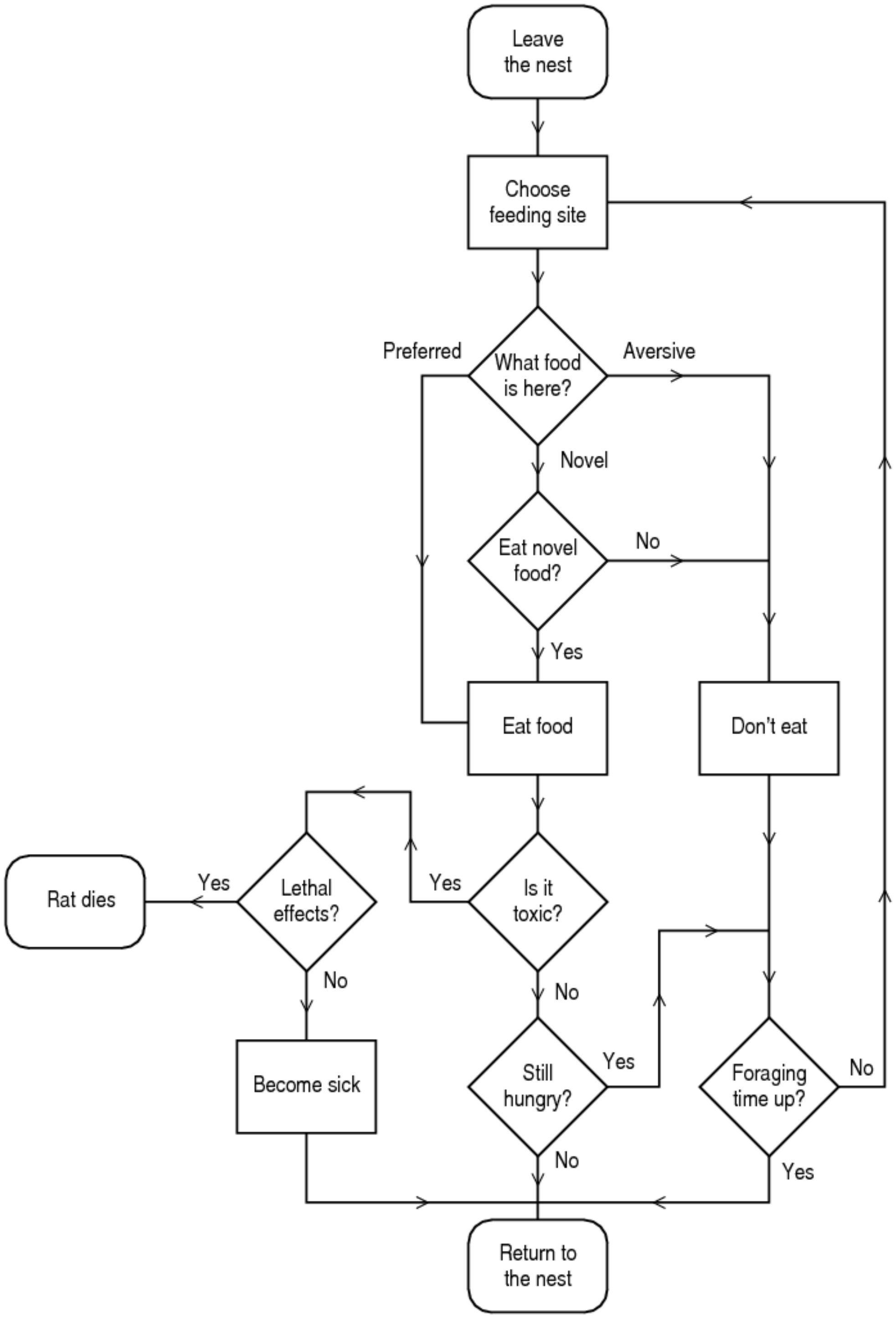
- Rats learn socially about food: neophobic but will eat novel food that others are eating.
- Mechanism discovered by Galef: social learning through smelling the breath of other rats.
- Weak point in the strategy: they still learn to like the food even if the other rat is sick. And they can tell the difference between sick and healthy rats.

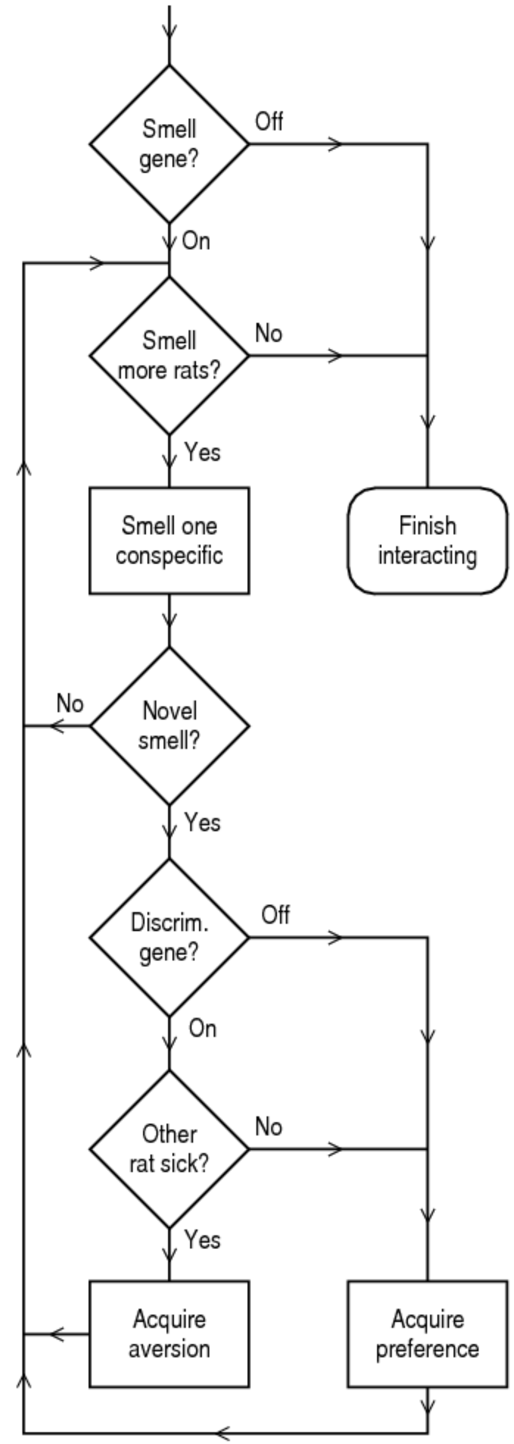
Learning about food by smelling each other's breath

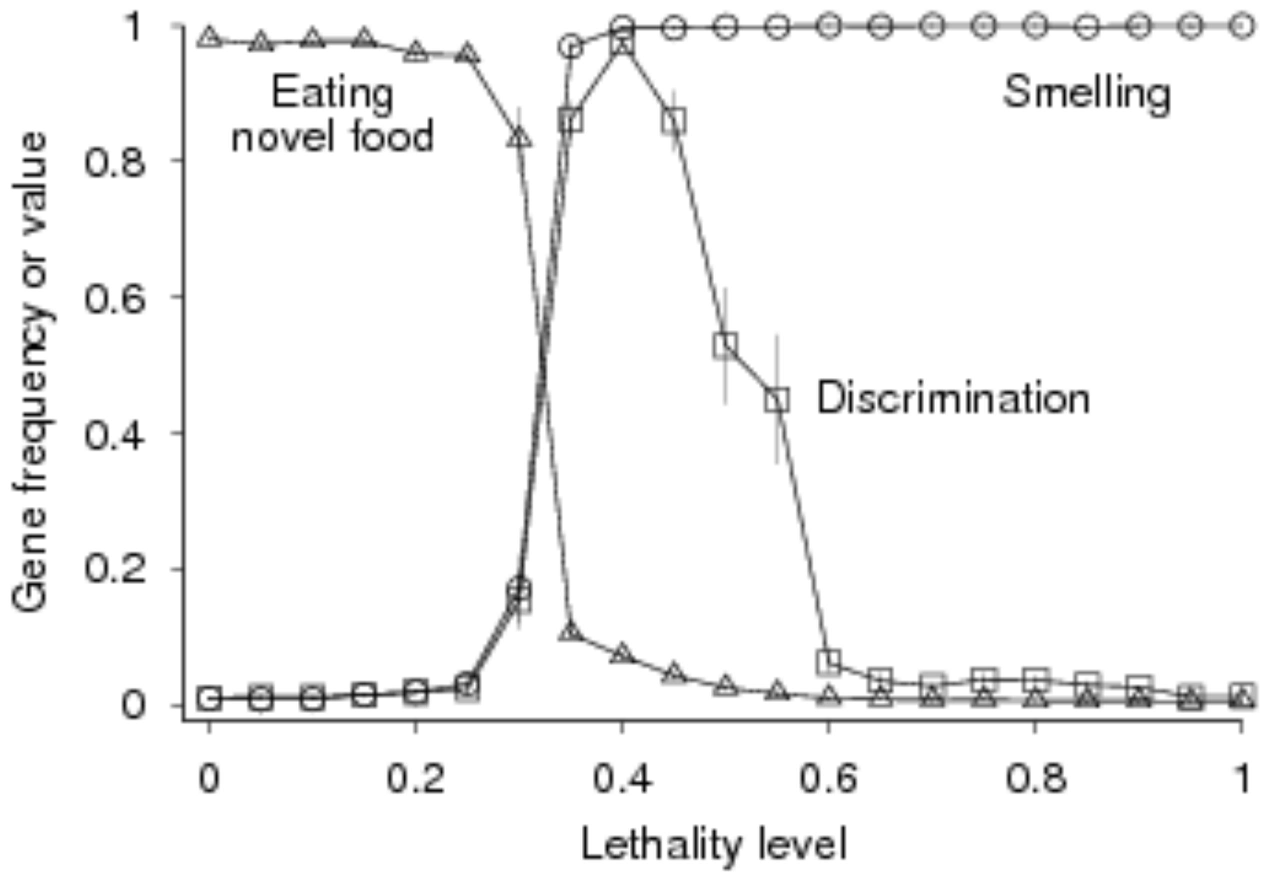
- In parallel with the Kahneman and Tversky approach in humans, we could say: "Look at the silly animals, see the dumb mistake they're making!"

- Instead our approach was to ask how would the environment have to be structured to make sense of this apparently crazy behaviour.

- The alternative strategies we considered were not the whole universe of possible strategies (indeed, how could we specify this?) but general foraging and learning strategies that we knew other mammals exhibited.


Leave the nest
Choose feeding site
What food is here?
Preferred
Aversive
Novel
Eat novel food?
No
Yes
Eat food
Don't eat
Is it toxic?
Yes
No
Lethal effects?
Yes
No
Rat dies
Become sick
Still hungry?
Yes
No
Foraging time up?
No
Yes
Return to the nest


Smell gene?
Off
On
Smell more rats?
No
Yes
Smell one conspecific
Finish interacting
Novel smell?
No
Yes
Discrim. gene?
Off
On
Other rat sick?
No
Yes
Acquire aversion
Acquire preference



The results of a simulated evolutionary process

What did we learn about rats?

- Rat behaviour makes sense given an assumption of reasonably high sensitivity to toxic food.

- Not established as true through the simulation, but a productive and reasonable way of generating hypotheses.

- Both blades of the scissors are in play: Galef's work on the mechanisms behind social learning in rats fixes one blade, and we asked what environmental factors would make sense of the other.

The putty-nosed monkey

Another example: intentional communication in monkeys

- Work in progress: a model of alarm calling behaviour in putty-nosed monkeys.
- Their near relatives, vervet monkeys, are famous for their apparently referential signalling.
- Putty-nosed monkeys have recently been observed to have something like a primitive syntax in their calls.
- Interesting animals for anyone concerned about continuity between animal communication and human language.

These guys have much to be worried about...

Monkey alarm calls

- Canopy dwelling monkeys that are preyed upon by leopards and eagles.
- Use a system of alarm calls to warn each other of danger.
- A range of systems can underlie the behaviour of signallers and receivers:
  - From hard-wired responses…
  - To complex cognition such as concepts, beliefs, and perspective-taking

- Currently involved in building a model to look at what level of cognitive sophistication we should expect from the monkeys given the environmental challenge they face.

- Levels of intentionality (from Dennett)
  - Zero-order system: signaller can't be said to believe anything; he is hard-wired to signal as soon as he sees a leopard
  - First-order system: signaller believes that there is a leopard nearby and decides to signal (difficult to distinguish from the zero-order case)
  - Second-order system: signaller believes the receiver is not aware of the leopard
  - Third-order system: signaller believes that the receiver will believe that the signaller has seen a leopard

Conclusions

- Optimality modelling can be safely employed as long as it doesn't come with a commitment to global rationality.
- These types of models are a useful way of getting both blades of the scissors in contact when thinking about bounded rationality.

Acknowledgements

- Peter Todd and Elio Tuci (social learning in rats)
- Kate Arnold and J.P. de Ruiter (communication in monkeys)