Understanding causation: the practicalities

Jackie Chappell and Aaron Sloman

Centre for Ornithology, School of Biosciences
and School of Computer Science
University of Birmingham, UK

International Workshop on Natural and Artificial Cognition, 2007

Outline

Acknowledgements

> *"Truth springs from argument amongst friends."*
>
> – David Hume

- Aaron Sloman
- Jeremy Wyatt and other members of Birmingham CoSy team
- Chris Miall
- Alex Kacelnik, Alex Weir, Ben Kenward

Introduction

Broad question:

How do animals (or humans or artificial agents) represent, use and manipulate the vast complexity of the world outside their bodies?

How do they:

- Perceive novel objects and their affordances and properties?
- Predict events?
- Plan actions?
- Represent these actions?
- What forms of representation are used?
- How can we investigate this experimentally?

Some (very broad!) definitions

Representation Encoded entity in the brain, coding for something concrete or abstract in the outside world, or relationships between representations. Could be explicit or implicit, conscious or unconscious.

Understanding/Knowledge Functional, adaptive use of representations. Again, not necessarily explicit.

Affordance All “action possibilities” latent in the environment, dependent on the agent’s capabilities (Gibson 1979)

Built-in Alternative to ‘innate’: largely independent of experience for expression

What are the possible sources of knowledge about the environment?

There are essentially three main options for evolution (or someone building an artificial agent):

1. Build almost everything in
2. Acquire from scratch from the environment (with some constraints)
3. Build in a framework for understanding structure (meta-configuration): content is acquired by learning, but framework is built in
4. Some combination of the above in varying proportions for different competences

Possible evolutionary strategies

Build almost everything in

- Advantages: Available very early or from birth/hatching, reliable
- Disadvantages: Limited flexibility, by slightly adjusting parameters. No provision for situations novel in evolutionary history, or un-anticipated by engineer
- e.g. precocial skills: flight in cliff-nesting birds, pecking, suckling

Possible evolutionary strategies

Acquire from scratch

- Advantages: Powerful and (almost) infinitely flexible
- Disadvantages: Can be slower, requires experience, less reliable and error-prone. Often puts greater parental care burden on parents.
- e.g. altricial skills: carnivores learning to hunt, orang utans learning distribution of fruiting trees in canopy.

Possible evolutionary strategies

Framework for structured learning: meta-configuration

- Allows rapid acquisition of an appropriate response in a novel situation (because the agent can probably predict what will happen without trying it)
- Works in situations never encountered in evolutionary history
- If ‘chunks’ of new knowledge/representations can be re-combined, can build up powerful new competences very quickly

Evolutionary strategy depends on type of competence required

- Strong selection pressure on some competences to be performed correctly first time, and early in the animal's life (e.g. suckling), though competence may be calibrated by experience (e.g. pecking in domestic chicks)
- Other competences aren't subject to this selection pressure
- Or, environment is so variable that built-in competences do not remain adaptive within generations

Precocial species tend to have many built-in competences, and altricial species many learned competences, but the real distinction is between *competences* and not *species*:

Meta-configuration: what is involved?

To recap part of Aaron's talk:

- Detailed knowledge of kinds of objects, properties, affordances etc. are probably learned rather than built-in
- But the following might be built-in:
  - Types of representation
  - Basic framework for classifying 'stuff'
  - Actions to attempt, kinds of things to explore
  - Ways in which knowledge can be combined

What is special about Kantian causation?

- Structure of objects play an important part in determining an animal's action
- Prediction is usually possible without trying an action, by understanding the role of structure
- Interventions can be made to test hypotheses - unlike in Humean causation, these can specifically target *functional* aspects of the situation (c.f. Hauser et al. 1999)
- Ability to monitor multi-strand relationships and processes: dynamic relationship between objects or parts of objects important (e.g. shapes of tool and aperture)

What might we predict?

If we suspect that an animal has meta-configured, Kantian competences, what kinds of behaviour might we expect?

- Exploratory behaviour specifically directed towards novel objects or novel parts of objects → strategies for forming hypotheses
- If an object or material looks like one previously experienced but has unexpected properties or effects, would expect another bout of exploratory behaviour → hypothesis testing, debugging.
- Animals which learn about a new property, material or affordance (type of structure) should be able to re-use that knowledge in perceptually very different situations → ontology formation and extension

The value of sensible 'defaults'

Why does knowing something about structure help?

We know that evolution does not usually provide a 'blank slate': many examples even in associative learning where some associations are made more readily than others, if at all (e.g. taste aversion conditioning in rats, Domjan and Wilson 1972)

Excluding putative causes because they cannot possibly be (or are extremely unlikely to be) the cause of an action reduces possible causes and helps to focus attention on the most likely candidates. See also work by John McCarthy, Thomas Kuhn and Alison Gopnik on human causal reasoning.

What kinds of defaults?

- Probably very rich
- Varieties of spatial concepts (e.g. near, far, on top of, underneath, next to etc.)
- Temporal sequencing of events (e.g. before, after, concurrent etc.)
- That actions *have* a cause?
- Solidity, contact, collision etc.
- And many more...

Similar ideas

Gopnik on the 'Theory Theory'

> "Moreover, there may still be some overall constraints on the kinds of representations that are generated, not every logically possible theory will be formulated or tested by human beings. These constraints reflect the basic presuppositions of scientific inquiry, for example, that the world has a causal structure that can be discovered." – Gopnik, In *Chomsky and His Critics* 2003

She also argues (as we do) that new competences (theories) are built on previous ones, and continually refined with experience and testing of hypotheses. But the mechanism she proposes for inferring new causal facts involves Bayesian networks.

See also similar, older ideas by Thomas Kuhn and John McCarthy.

Visible and invisible structure/affordances

Visible and invisible structure/affordances

If structure/affordances are not directly perceivable, can non-human animals use them?

- Geometry or structure of an object may give some information (circular objects roll) – this can be directly perceived, but may originally have been learned
- Some structure tends to co-vary with appearance (e.g. diameter of tree branch tends to indicate its rigidity) – this can be learned
- Some structure can be discovered during exploration or explicit testing, including exceptions to rules
- Some affordances might be one step away (doing x makes y possible)

Learning about structure and affordances is probably a very dynamic process, because some depend on action for their discovery

Role of exploration and play

- Allows testing of hypotheses
- Allows discovery of new properties and processes
- Some structure or affordances important for causal understanding are dynamic (or hidden – e.g. partially-occluded objects) must be tested by manipulation

Image source: http://icanhascheezburger.com/page/2/

Testing causal reasoning

Problems

1. Experimentally distinguishing causal reasoning from other mechanisms (e.g. associative learning etc.) – overt behaviour may be very similar
2. Separating rapid within-trial or within-session learning from causal reasoning
3. Getting animals to mentally simulate ‘what would happen if...’ would be ideal, but obviously we need to get them to act to get an output

Testing causal reasoning

Possible solutions

Very difficult to design good experiments, but kinds of things we need to think about are:

- If something 'violates expectations', expect burst of exploratory behaviour and new actions to test hypotheses
  - How do animals try to correct errors and problems?
  - But allowing error correction generates opportunity for associative learning
- 'Prediction' of events – do animals look in the correct place, act on the right part of the apparatus?
- Re-use of competences in perceptually very different configurations of apparatus, and re-combination of two or more competences to form a new one (e.g. secondary tool use, etc.)

Kantian causation

The role of structure

What is different about Kantian causation is that potential actions are constrained by structure of:

- Objects and processes in the world
- Internal representations, either built-in, or generated through experience

Exploration also has a key role, and qualitatively, we expect exploration to be directed towards specific objects or parts of objects.

Examples

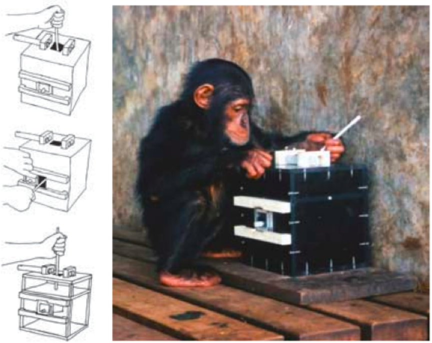
Which domains are promising for finding evidence of causal reasoning?

- Physical cognition: tool use, but also understanding of physical structure and affordances in non-tool using species, object manipulation, nest-building etc.
- Foresight and planning
- Social cognition

Perhaps all involve the same underlying mechanisms of causal understanding, making arguments about whether technical or social pressures are responsible for the evolution of intelligence redundant?

Examples

Physical cognition

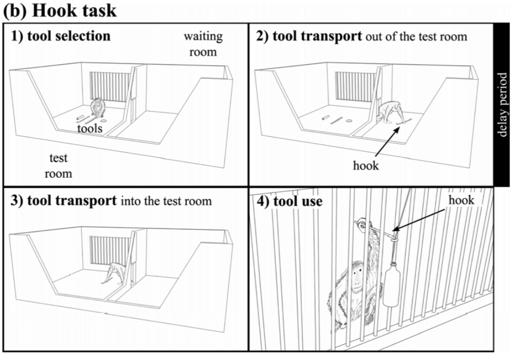


Horner and Whiten 2005

- Tested with opaque and transparent versions of the box
- Chimps (not children) stopped performing functionally useless action when they could see the structure of the box

Examples

Planning/Physical cognition - apes

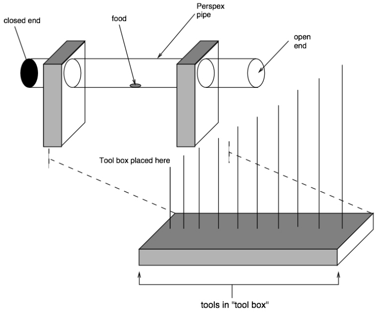


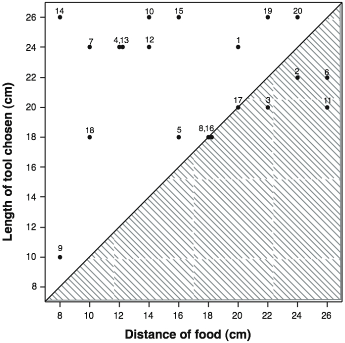
Mulcahy and Call 2006

- Subjects significantly targeted suitable tools to transport from the room
- 2 subjects succeeded even with a 14h wait

Examples

Planning/Physical cognition - New Caledonian crows





Chappell and Kacelnik 2002

Summary

- Good, adaptive reasons why some species might have developed meta-configured competences and Kantian-type causal understanding
- Experiments to test this are very hard to design (but possible!)

Many experiments already going in the right direction, but:

- We need to look in more depth at what happens when animals encounter problems or make errors
- More qualitative work on exploration needed, with qualitative detail on actions in experiments
- Problem: how to analyse this kind of data statisically?

References I

Aaron Sloman and Jackie Chappell
Altricial self-organising information-processing systems
Proceedings IJCAI'05, pp. 1187-1192
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0501

Jackie Chappell and Aaron Sloman
Natural and artificial meta-configured altricial information-processing systems
To appear in *Int. J. Unconventional Computing*, 3(3), 2007
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609

References II

Aaron Sloman and Jackie Chappell
Computational Cognitive Epigenetics
To appear in *Behavioral and Brain Sciences*
(Commentary on Jablonka and Lamb: Evolution in four dimension)
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0703