Agency and Explanation

CHAIN Winter School, Lecture 2

Dr. Nick Brancazio, Feb 7th, 2023

Outline

- The relationship between “minimal cognition” and agency
- Explanatory frameworks
- Some philosophical problems with agential explanations
  - Teleological explanations are naturalistically unfeasible
  - Fully detailing the underlying mechanisms removes the need for “agency”
  - Agency is something naturalistically inexplicable

What is Minimal Cognition?

- Methodology of minimally cognitive behaviour (Beer 1996)
- Basal cognition (Lyon et al. 2021)
- “Minimal cognition” sensorimotor coupling (van Duijn, Keijzer, & Franken 2006)
- Life-Mind continuity in autopoietic theory (Maturana and Varela 1980, Di Paolo et al. 2017)

Methodology of Minimally Cognitive Behaviour

- The term “minimal cognition” came into popular use thanks to Randall Beer’s work on modelling what he referred to as minimally cognitive agents (1996, 1997, 2003; Beer and Williams 2015). As he says of this project: “The term ‘minimally cognitive behavior’ is meant to connote the simplest behavior that raises cognitively interesting issues” (Beer 1996).

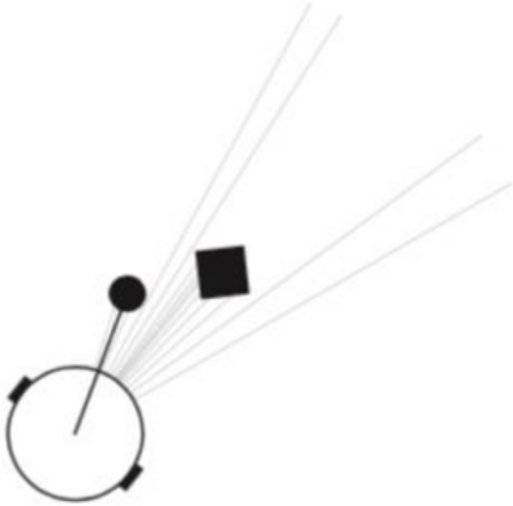


Figure 1: Basic design of a visually-guided agent. The agent (large circle) has an eye (gray lines), two motors (filled rectangles) and a transparent arm (solid line) with an opaque hand (filled circle). The arm can rotate about the center of the agent and extend or retract along its length. The hand can rotate about its point of attachment with the arm.

Basal cognition

- A **biogenic** approach to cognition
- Application of biological principles
- Urges that we treat the cognitive sciences as we do the other life sciences
- To define **cognition**, we have to look at how it **functions** to ensure survival, reproduction, adaptability - and how it has developed through natural selection
- Carve out the explanandum (capacity) in terms of its biological function for the organism

Minimal Cognition (DKF)

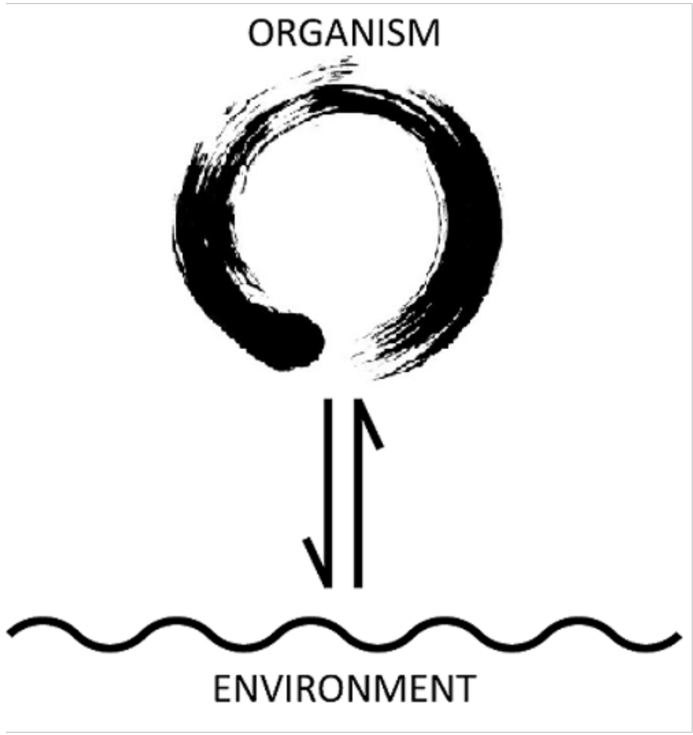
- Van Duijn, Keijzer, Franker (2006)
- Cognition as sensorimotor coordination
- “In our view, minimal cognition requires an embodiment consisting of a sensorimotor coupling mechanism that subsumes a basic metabolic/autopoietic network; this forms the basis of the increasingly popular idea that cognition revolves around sensorimotor coupling.”
- Coupling of the sensory system and motor system (ongoing feedback process)
- “lac operon” system / chemotaxis

Autopoiesis

- Coined by Maturana & Varela - somewhat mainstream concept in biology nowadays, way less so in cognitive science
- Living things live far from equilibrium, which requires lots of work to maintain.
- What characterises living things is that they perpetuate themselves, they are self-generating:
- A “cell produces its own components, which in turn produce it, in an ongoing circular process.” (Thompson 2007, pg. 98)
- They maintain their own boundaries, bring in resources they need, excrete things they don’t, respond to their environment, etc.

Autopoiesis

- Autopoiesis (self-creating):
- Autonomy
  - Self-organizing
  - Self-producing
- Organizational Closure
  - Self-supporting network of processes/components
  - Boundary maintenance
- Agency
  - Selective openness



Questions about Agency

- **What is agency?**
  - A capacity, a process, etc.
- **What is an agent?**
  - The demarcation criteria between agents/non-agents
- **Why does an agent do *x*?**
  - Explanation/Reason for the behaviour
- **How does an agent do *y*?**
  - Explanation of the behaviour

Different Questions, Different Answers

- Depending on the type of question and aspect of agency under investigation, we can select an appropriate explanatory framework for the phenomena
- Ex. Broad: Are plants agents?
- Ex. Specific: Why do plants change their tendril spinning behaviours at time t?
- Ex. Specific: How do plants decide when to alter their tendril spinning behaviours?

Philosophical Concerns: "Why" Questions

- Teleological explanations appeal to the purpose/function/goal of an organism
- The purpose/function/goal is treated as the cause of a behaviour
- Example: Why did Nick come to Sapporo? Answer: To lecture at the CHAIN Winter School
- Worry: The format of this answer implies that the future has caused the past. This is (supposedly) not good science - in order to empirically investigate this, we have to posit *something* previous to the behaviour that references the future.
- "Teleology is a mode of explanation in which the existence or nature of an object or event is accounted for by citing the purpose it subserves... [T]he worldview ushered in by the Scientific Revolution is inimical to teleology. That world is a realm of mechanisms and their effects: bodies in motion, the transmission of light, the flux of heat." (Walsh 2016, p. 186)

- If we accept this as a problem, then functional explanations are off the table for biology:
  - NO: any explanation of behaviour that appeals to what a behaviour is *aiming to do* for an organism

- Enter: natural selection

Etiological Theory of Function

- Griffiths (1993) and Godfrey-Smith (1994)
- This theory provides a framework for understanding claims about the function of a trait or activity in a species by specifying only adaptation in natural selection as the causally relevant mechanism underpinning its existence.
- Answers to “why” questions about a trait in a population can be answered as follows:
- “The function of x in z is to y” where
  - x is a type of process or structure,
  - z is an organism type, and
  - y is an activity

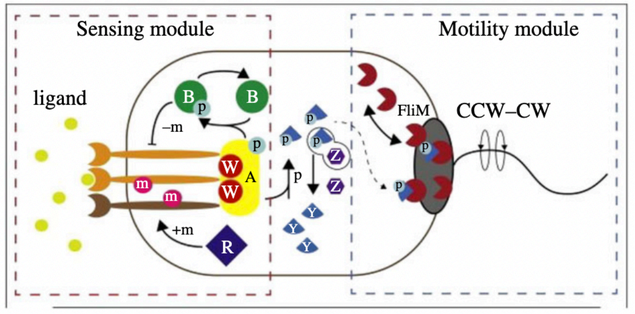
Pushback

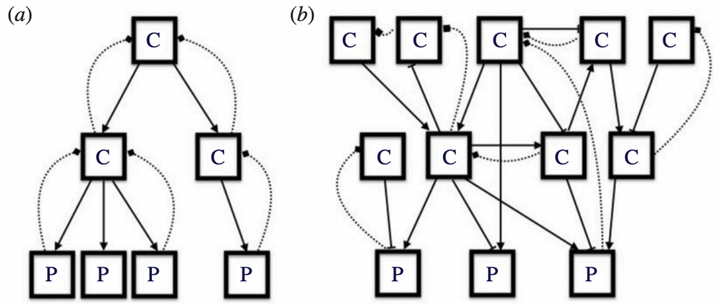
- Walsh (2014) points out that functional explanations that apply to the activities or part-whole relations of individuals cannot take the ETF form, and thus, answering "why" questions about individual behaviors necessitates "genuinely teleological explanations in biology"
- "Natural selection tells us about the process that gives rise to teleology. But it does not explain how it works. It does not point to any particular mechanism behind the mysterious moment-to-moment behavior, the seeking behavior, of seemingly purposeful entities. Selection also cannot explain the apparent goal-directedness of the Darwinian process itself, of selection itself." (McShea 2016, p. 65)

Answering “How” Questions

- The ETF does not answer our “How” questions (at population or individual level)
- In other words, we may still need to posit goals or purpose at times

- How does a slime mould navigate a maze to find food?
- Explanatory frameworks (examples):
  - Organisational frameworks (e.g. autopoiesis, interactivism)
  - Formalizations / Modeling (regularities and/or probabilities)
  - Information-processing
  - Mechanism





Toolkit of Basal Cognitive Capacities

Lyon et al. 2021

capacity	function
orienting response	Ability to selectively attend to a state of affairs to the exclusion of others. Biological basis of attention.
sensing/perception	The capacity to sense and recognize (re-cognize) existentially salient features of the external and internal milieux.
discrimination	The ability to determine that a state of affairs affords an existential opportunity or presents a challenge requiring a change in internal state or relation to the external environment.
memory	The capacity to retain information about the immediate (and possibly distant) past, and to calibrate the sensorium to take account of this information, at a minimum via habituation or sensitization.
valence	An organism's capacity to assign a value (advantage/good, harm/bad) to a particular stimulus or the summary of information about its surroundings relative to its own current state. The 'building block' of affect [47].
decision making	The capacity to combine information from multiple sources and act, typically in furtherance of an implicit or explicit goal.
behaviour	An organism's capacity to adapt to changes in its internal or external milieux by changing its own structure and function (internal) and/or its spatial and interactive relations (external).
problem solving	Behaviour selection in circumstances with multiple (possibly conflicting) parameters and high degrees of uncertainty.
error detection	The capacity to determine whether a behaviour has succeeded or failed. Together with homeostasis/allostasis, this is the basis of biological normativity.
motivation	Implicit goals arising from existential conditions. Teleonomic striving. The inclination of a thing to continue to exist and enhance itself.
learning	The capacity to adapt behaviour to salient stimuli according to past experience by altering the threshold for or the nature of a response.
anticipation	The capacity to predict what is likely to happen next based on an early stimulus.
communication	The capacity to interact productively with other organisms via forms of signalling, notably (but not limited to) conspecifics, including initiating collective action, which may or may not include an explicit means of differentiating 'like me' (us) from 'not-like me' (them).

Philosophical Problem with "How" Questions

- If we can answer the "how" question, what work is the "why" question doing?
- Worry: Agential explanations are those that require explaining the *reason* for action ("how" questions do not actually explain agency)
- Worry: We are merely using an *intentional stance*
- Worry: If we can give a full explanation of the mechanisms behind agency, then perhaps there is no agency

Goals = Representations

- Fodor (1987): *intentional behavior* supported by *representational mental states* is the hallmark of agency
- Representations enable agents to behave in ways that are not merely governed by "lawful covariation of a property of the stimulus with a property of the response" (1986, p. 9).
- it just would not be possible for paramecia to be agential, because purposeful action would require the possibility of *representing an aim*
- That "goal" must be capable of guiding an organism's activity. Otherwise, the system is merely responding to environmental stimuli.

Response (1)

- **Fred Keijzer** (2001) argues that a major flaw in what he calls Agent Theory (AT) involves using our *folk psychological understanding* of *reasons for action* (beliefs, knowing, planning, goals) to guide our understanding of the processes underpinning behavior.
- Keijzer claims that "[w]hat AT does is to change the normal, personal-level context of representation and use it at a sub-personal level" (2001, p. 40-41). That is, it assumes without further evidence that our social conventions of propositional goal-representation are supported by an empirically concrete representational mental state with roughly the same content.

- Fodor's cognitivist approach gives us a naturalistically plausible way to answer both "why" and "how" questions
- But Keijzer cautions that the folk psychological terms that we use to describe agential behaviors are assumed to directly correspond to sub-personal mechanisms that support the behavior
- And framework-specific assumptions about representations in cognition are used to define what belongs in the category of agents (Lecture 1)

The Intentional Stance

- “The intentional stance is the strategy of interpreting the behavior of an entity (person, animal, artifact, whatever) by treating it as if it were a rational agent who governed its ‘choice’ of ‘action’ by a ‘consideration’ of its ‘beliefs’ and ‘desires.’” (Dennett 2009)
- “The intentional stance is thus a theory-neutral way of capturing the cognitive competences of different organisms (or other agents) without committing the investigator to overspecific hypotheses about the internal structures that underlie the competences.” (Dennett 2009)

Response (2)

- Helen Steward (2009) advocates for moving away from this kind of connection between agency and a highly intellectualized conception of representational intentions or beliefs as the motivators of action.
- Her view is that when we attribute agency to non-human animals, it is because we see them as a “centre of some form of subjectivity” (p. 226)
- We are justified in attributing some rudimentary intentional states to them, such as wanting, perceiving, or trying, in a rudimentary (non-propositional) sense

- Steward argues that **our intuition to attribute** non-propositional reasons for action “deserves not to be lost underneath the mountain of epistemological scrupulousness, mech-anomorphism, anti-dualist fervour, and behaviouristic scientific methodology” (ibid., p. 229).
- As an example, she contrasts this intuition about our attributions of agency for some non-human animals with our intuitions about the paramecium:
- “[O]nce we find out that a few simple equations govern the movement of the paramecium through the water, there is no reason to suppose that there is any role left for a **paramecium-self** to play in the control of the paramecium body. … None of its interactions with the environment need be mediated by anything desire-like, and none need involve anything like **an act** on the part of the paramecium.” (Steward 2009, p. 227)

- In other words, if the bodily activity of the paramecium requires no further explanation than pointing to the biochemical processes between the system and environment, then there is simply no further explanatory work for agential language to do. **There is no role for subjectivity, thus no agency.**

- Ex: “Even when they are doing something highly complex, the behaviour of ants and spiders and bees does not seem to be under the individual’s control. Their evolved biology is in control.” (Tomasello 2022, p. 1)

- The idea that all life is agentive does not seem to sit well with many philosophers and scientists.
- However, there is no clear consensus on what the “magic ingredient” is that humans have when they undertake agential behaviours that non-human organisms do not possess.

Discussion Questions

- “[S]ometimes the reduction eliminates the phenomena of interest. Thus models of cognitive science built on the computer metaphor or neuroscientific explanations often do not have room for a decision-making agent. If a model has a decision-maker that is not somehow mechanical, the charge of ‘homonculus’ is levied, and there is an attempt to eliminate it. But I believe that for the most important phenomena of psychology, the decision-making agent cannot be eliminated without losing those phenomena.” (Tomasello 2022, p. 132)
- What is being said in this quote? Do you agree or disagree, and why?