A Role for Consciousness in Action Selection

Joanna J. Bryson
Department of Computer Science
University of Bath
http://www.cs.bath.ac.uk/~jjb
A Naive Perspective

- Surely if you build action-selection for humanoid robots you are going to need consciousness.
- So if I just keep building sooner or later I’ll know what it’s for.
What is Consciousness?

Please forgive me for being brutally functionalist at least in the context of this talk, but I only have 30 minutes.

• A culturally-evolved concept of uncertain age and origin (Dennett, 2001).

• May refer to no single psychological phenomenon.

• A part of a presumably-adaptive system for action selection.
Plants can wind & unwind (reversing decisions) in pursuit of support, light, prey.
(Anthony Trewavas, Edinburgh)

Single cell organisms also pursue multiple goals & hunt prey.
Tradeoffs in Nature

- Where we see variation in expression of a trait in nature, one explanation is an evolutionary tradeoff.
- Multiple equally-adaptive solutions.
- e.g. number of offspring vs investment per offspring
Thinking is Costly on a Number of Dimensions

Big brains correlate w/ long life, slow maturation, big size & sparseness.

a, b Barrickman, Bastian, Isler & van Schaik (2008): y axes are residuals.

For this talk:

**Focus on Time**
Why is it hard to be smart?

Pretend you bought a robot, and it came with 100 things it knew how to do without being told.

For example, eat, sleep, turn right, turn left, step forward, step backward, pick things up, drop them...

Now pick a goal for your robot.

For example, flying to Tokyo.
Suppose you can’t be bothered to tell your robot exactly how to get to Tokyo, so you have it guess.

If getting to Tokyo takes one step, the robot may have to try 100 different things.

If it takes two steps, the robot may have to try each thing after each thing:

$$100^2 = 10,000$$
The hardness of smartness (3)

- If the robot doesn’t know how many steps it takes to go to Tokyo, it might get caught in an infinite loop.

- For example, it might eat, sleep, work, eat, sleep, work, eat, sleep, work... and never buy a passport.

- When computer scientists say “hard” they mean “pretty much intractable.”

Sony SDR-4Xs. Pictures from BBC
Search vs Time

• **Combinatorics** is the problem, **search** is the only solution.

• The **task of intelligence** is to focus search.
  • Called **bias** (learning) or **constraint** (planning).
  • Most behaviour has no or little **real-time search**.

• For **natural intelligence**, most **focus evolves**.
  • Physical/cognitive **constraints** limit search space.
Learning and Time

• Looking time experiments rely on reaction-time delay being indicative of surprise.

• Correlation of lost reaction-times and failure to notice shift in reward schedule, but not inhibited performance (Rapp et al 1998).
Allocating Time & Attention

I. Individuals allocate more time when less certain (Bryson 2009; 2010).

II. Species allocate in response to niche e.g. tamarins & insects (Hauser 1999).

III. Species allocate inversely with age (Rapp et al. 1998, Bryson 2009; 2010).

IV. Individuals allocate inversely with urgency (Shadlen and Newsome, 1998; Bogacz et al., 2006).
A Theory of Conscious Attention

• The basic function of conscious awareness is to update importance models (learn).

• Time is allocated in proportion to importance by inhibiting action.

• Not to choose immediate action!

• If new action is favoured due to model updates, may affect immediate behaviour.
Consciousness for AI

• Only need it if system learns and learning relies on a bottlenecked cognitive resource.

• If you need it, allocating it to tasks you are doing in proportion to how uncertain you are about them is a pretty good guess.

• Also attend to other novel / unpredicted by your internal model events.
Point of Intervention

1. Action *selection* as usual.

2. Inhibit action *expression* while selected action is in mind, *update models*.

3. If new action becomes more salient, *insight*.
   - Flush plan & start over.

4. Updated models may not have immediate impact on behaviour.
Consequences
Self Consciousness

- Consciousness of self: limited like all consciousness to likely useful search space.
- Much facilitated in humans by language & instruction ⇒ probably less in other species.
- Google Search treats its own pages like other’s: self-awareness neither necessary nor sufficient for consciousness.
Self Consciousness

• Consciousness of self: limited like all consciousness to likely useful search space.

• Much facilitated in humans by *language* & instruction $\Rightarrow$ probably less in other species.

• Google Search treats its own pages like other’s: self-awareness neither necessary nor sufficient for consciousness.
Language Helps

- Symbolic representation allows more compact and / or less emotionally-salient representations.

- Learn concepts from others; shared consciousness of events (*Dennett 2008*).

- **Not a prerequisite** for this basic functional component of action selection.
Ethics

• **Consciousness**: culturally-evolved concept of uncertain age and origin (Dennett, 2001).

• May refer to no single psychological phenomenon.

• **Ethics**: Co-evolve with social order.

• Much relies on assigning responsibility: covaries with but not determined by consciousness.
Robots are servants we own.

We build robots and determine their goals, either directly or by specifying how they acquire goals themselves (Bryson 2010b).
Because we build robots and determine their goals, our relationship to them is fundamentally different from our relationship to humans or other evolved systems.

photos: Georgio Metta (top) & Emmanuel Tanguy
Conclusions

• The basic function of awareness is not to choose actions, but to inhibit actions once selected and learn about their situation.

• A costly (in terms of time) allocation of resources for learning, varies in application by species and by individual situation.
Thanks!

Vienna Conference on Consciousness
Mark Baxter, Wayne Christianson
The Organisers
Not to Wellcome
You, For Your Attention