COSC451: Artificial Intelligence

Lecture 6: The ‘who’ pathway: representations of self and other

Alistair Knott

Dept. of Computer Science, University of Otago

Outline of the lecture

1. Concepts of ‘self’ and ‘others’
2. Neural activity reflecting agency
3. The ‘match’ model of action attribution
4. The ‘mode-setting’ model of action attribution

Outline of today’s lecture

1. Concepts of ‘self’ and ‘others’
2. Neural activity reflecting agency
3. The ‘match’ model of action attribution
4. The ‘mode-setting’ model of action attribution

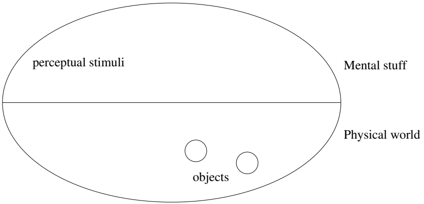
Concepts

A baby has to develop a set of concepts to represent objects and actions in the world.

- The baby receives perceptual stimuli, and must learn to interpret these as information about objects and actions.

We can distinguish between:

- Our perceptual stimuli (part of our mental life)
- The physical world of objects and actions which we construct from perceptual stimuli.

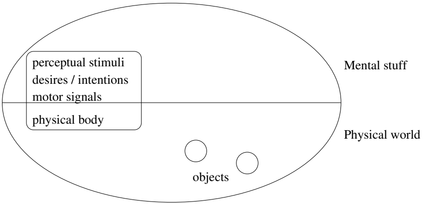


The concept of self

We also have a concept of self.

- There's a special object in the world: our own body.
- Our body is an object from which our percepts originate.
- It is also an object which we can *control*.

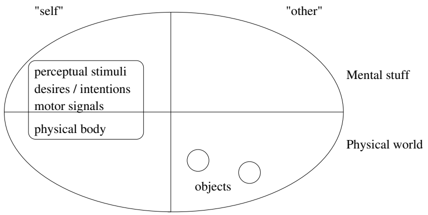
Our concept of self is partly physical and partly mental.



The concept of self

In order to develop a concept of self, we need more than just perceptual stimuli.

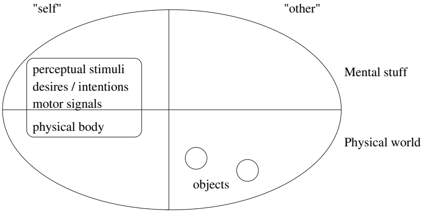
- Our mental world also contains desires, intentions, and motor signals.
  These are mental constructs which relate to our ability to *act*.
- Using these constructs, we can separate the physical world into 'self' and 'everything else'.



Concepts of other agents

We also have concepts of other agents.

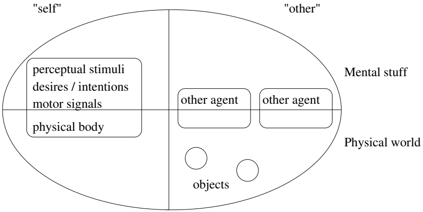
- These agents are part of the external world.
- But they are nonetheless *like us*, in that they have mental states.



Concepts of other agents

We also have concepts of other agents.

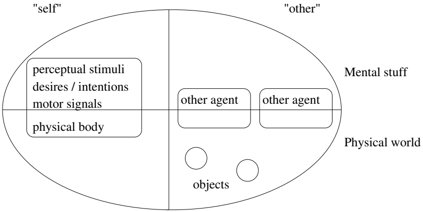
- These agents are part of the external world.
- But they are nonetheless *like us*, in that they have mental states.



Concepts of other agents

We also have concepts of other agents.

- These agents are part of the external world.
- But they are nonetheless *like us*, in that they have mental states.



We have to develop concepts that represent:

- Everything in the physical world.
- The distinction between ourselves and other objects.
- The distinction between ourselves and *other agents*.

Learning the concept of self

How do we acquire the concept of self?

There’s one particular object—our physical body—for which there are *associations* between physical events and mental phenomena.

- Events involving this body are experienced as percepts.
  (E.g. when the body is touched, we experience a sensation.)
- Motor signals which we evoke result in movements of this body.

I’ll describe some experiments which investigate these two methods of defining ‘our own body’.

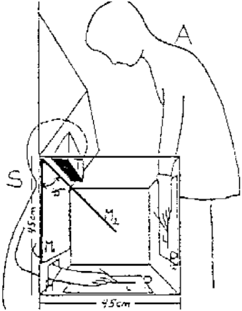
Expt 1: learning a perceptual conception of self

The rubber hand illusion (Botvinick and Cohen, 1998).

- The subject can't see their real hand: they watch a rubber hand instead.
- The experimenter gives the same tactile stimuli on the real and rubber hands.
- After a while, many subjects begin to perceive the rubber hand as their own.

Expt 2: learning a motor conception of self

Nielsen (1963).

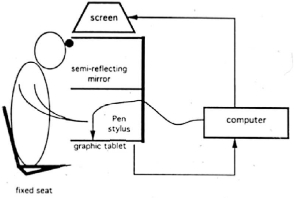


- Subjects draw a line on a piece of paper.
- Instead of seeing their own hand, they see the experimenter's hand (at roughly the same position).
- If the experimenter performs the same action as the subject, subjects often mistakes his hand for their own.

Aside: The primacy of visual information

In the rubber hand illusion, subjects trust their eyes more than their proprioception.

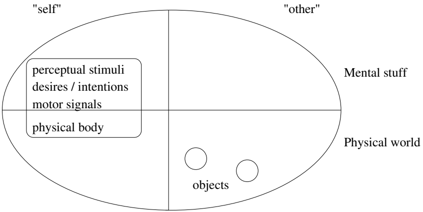
Here’s another experiment by Fourneret and Jeannerod (1998).



- Subjects draw on a graphics tablet, and see a line on a computer screen, at roughly the position where they expect it.
- The line can be *perturbed* away from the position it should appear in. Subjects correct for small perturbations unconsciously.

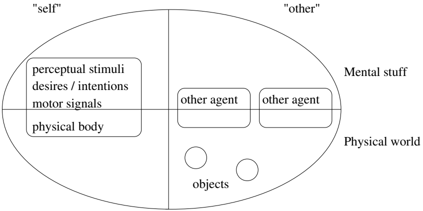
Learning the ‘self/other-agent’ distinction

We must learn to distinguish between our body and ‘external’ objects.



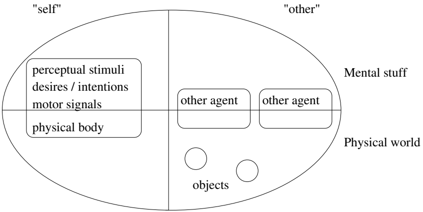
Learning the ‘self/other-agent’ distinction

We must learn to distinguish between our body and ‘external’ objects.
We must also learn that some external objects are *other agents.*



Learning the 'self/other-agent' distinction

We must learn to distinguish between our body and 'external' objects.
We must also learn that some external objects are *other agents.*



When should we attribute agency to an external object?

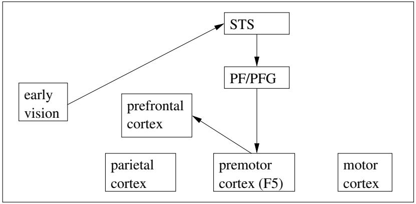
- When it generates its own movements. . .
- When its movements can be well explained by *positing mental states*.

‘Explaining a movement by positing a mental state’

Some observed events can be well explained by simple physical laws.

- E.g. a stone rolling down a hill.

Other events are better explained by *assuming* that they are caused by an agent like ourselves.

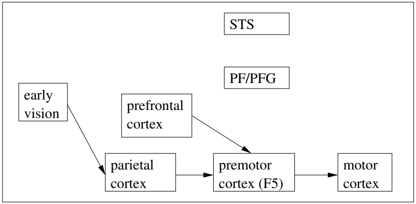


We can use our own motor/intentional representations to help us *explain* certain observed events, even if we didn’t *cause* them.
This amounts to hypothesising the existence of other agents.

Attributing actions to onesself or to another agent

We are now using our motor/intentional representations for two purposes:

- To generate our own actions.
- To explain the actions of other people.

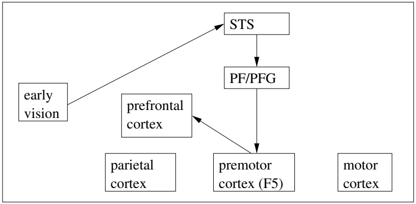


Question: how do we distinguish between these two situations?

Attributing actions to onesself or to another agent

We are now using our motor/intentional representations for two purposes:

- To generate our own actions.
- To explain the actions of other people.



Question: how do we distinguish between these two situations?

Outline of today's lecture

Brain areas active during action execution / observation

Are there any brain areas which are more active during action execution than action observation, or vice versa?

- An obvious idea: Action execution involves the motor cortex; action observation involves STS.

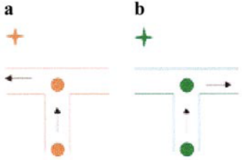
This turns out to be true (see e.g. Rizzolatti *et al.*, 1996).
BUT:

- We often watch our own actions. (So STS is often also active during action execution.)
- Our body is often moving while we watch someone else's action. (We don't attribute the observed action to ourselves in this situation!)

Brain areas associated with the ‘experience of agency’

Farrer and Frith (2002) attempted to manipulate ‘sense of agency’ while controlling for all other sensory/motor factors.

- Subjects used a joystick to drive a circle along a T-shaped path.
- The visual stimulus they saw was either generated by their own movement, or a (different) movement made by the experimenter.

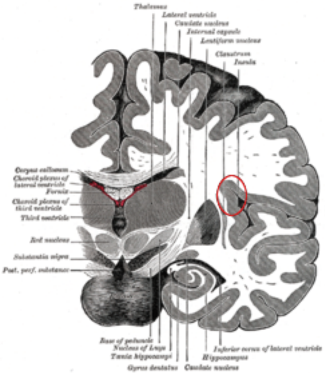


- They saw a movement, and made a movement, in both cases.
- But they only experienced a ‘sense of agency’ in the former case.

Brain areas associated with the ‘experience of agency’

Results:

- The experience of agency was associated with activity in the anterior insula.

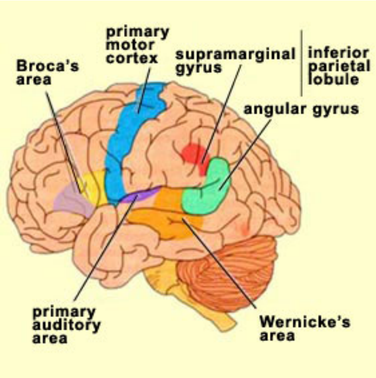


(The insula is a region of cortex that lies inside the lateral sulcus.)

Brain areas associated with the ‘experience of agency’

Results:

- The feeling of ‘watching someone else’s action’ was associated with activity in an area of the inferior parietal cortex called the angular gyrus.



Some other findings about the anterior insula

Blakemore *et al.* (2005) investigated a patient who had vision-touch synesthesia: she experienced tactile sensations when she *saw* people being touched.

- This patient showed abnormally high activity in the anterior insula. (As well as in somatosensory cortex and premotor cortex.)

If anterior insula activation is associated with self-attribution:

- Abnormally high insula activity might cause an observed touch action to be registered as an 'experienced' action.
- Activation in premotor/somatosensory cortex might be the result of this.

Some other findings about inferior parietal cortex

Schizophrenia is a condition where attributions of agency to self and other are often confused.

- Some schizophrenics show ‘passivity’—a condition where they think their own actions are under control of an external agent.
- Such patients have increased activation of inferior parietal cortex.

If IP cortex activation is associated with attribution of an action to others, perhaps abnormally high IP activation is the cause of passivity phenomena.

Outline of today’s lecture

The 'match' model of action attribution

How do we attribute agency to ourselves or to others?

One idea: the sense of agency results from a *match* between signals generated in the motor system and signals arriving from perceptual (visual/somatic) sources.

- Recall Nielsen (1963): subjects identified the observed hand as their own if its actions were correlated with their motor commands.
- Recall the lecture on motor control: an efferent copy of motor commands is given to a forward model, which calculates its expected sensory consequences.

Maybe we attribute an evoked premotor signal to ourselves if the sensory stimuli we experience are those which are predicted from this motor signal.

The 'match' model of action attribution

One problem: according to our model of the mirror system, we expect correlations between visual stimuli and motor signals whether we perform an action ourselves, or watch another agent performing it.

One way to save the match model of agency: the perceptual stimuli have to be *somatosensory*, not visual.

- Normal agents only experience somatosensory stimuli when they perform an action.
- So maybe the anterior insula is involved in detecting a match between motor signals in premotor cortex and reafferent *somatosensory* signals.
- Note: the anterior insula does receive a lot of somatosensory information.

The ‘match’ model of action attribution

So why is the inferior parietal cortex active when we attribute an action to someone else?

Recall:

- Inferior parietal cortex is on the ‘action recognition’ pathway (STS→PF/PFG→F5)
- Posterior parietal cortex is on the ‘action execution’ pathway
- Maybe these pathways compete with one another, with high IP activation favouring action recognition.

Problems with the 'match' theory of attribution

1. The issue of whether somatosensory signals match motor signals feels more like a *consequence* of a decision about who the agent is, not its cause.

2. Models of the mirror system all assume that there are two separate modes: recognition mode and execution mode.

- The circuitry of the system is completely different in the two modes.
- There should be explicit mechanisms for *selecting* one mode or another, so that (e.g.) the agent can learn when to act and when to observe.

Problems with the 'match' theory of attribution

3. The match theory provides a means for distinguishing between onesself and external agents, but it doesn't specify how we distinguish *between different external agents*.

- Presumably we have to identify the agent using the object classification pathway in IT. But how does this process relate to the process of action monitoring?

4. We can *imagine* actions, or *remember* actions, without there being any somatosensory stimuli at all.

- We have no problem at all distinguishing between our own actions and those of others in such situations.

Outline of today’s lecture

1. Concepts of ‘self’ and ‘others’
2. Neural activity reflecting agency
3. The ‘match’ model of action attribution
4. The ‘mode-setting’ model of action attribution

The ‘mode-setting’ model of attribution

I suggest that there are special operations which put our mirror system into ‘recognition mode’ or ‘execution mode’.

The operation which puts us into recognition mode is an action of attention to an external object.

- The object captures our attention, because it is salient. This puts us into action recognition mode.
- Thereafter, this object will be understood as the agent of any action evoked in our mirror system.
- Attending to the object also provides us with an opportunity to categorise it.

The operation which puts us into execution mode is a decision to act.

The mode-setting model of attribution

This mode-setting model of attribution solves the problems raised by the match model.

- The agent can now learn when to act and when to observe.
- The agent can distinguish between different external agents.
- In order to distinguish between self and other in imagined or remembered actions, the agent can *simulate* the operations of moving into action execution or action recognition mode.

What's a 'decision to act'?

Is it really plausible that agents 'decide to act', prior to deciding what action to do? Here's some evidence.

- If you record EEG signals from a subject and ask them to make a spontaneous movement, you see a wave of activity in precentral and parietal areas which precedes the action, called the cortical readiness potential (Shibasaki 1992). It's the same for all actions.
- There is a second EEG signal called the lateralised readiness potential which occurs later in action preparation, and which reflects which hand will be used.
- Incredibly, the general readiness potential occurs *before the agent is aware of their decision to act* (Libet *et al.*, 1983).
- The lateralised readiness potential occurs *after* the agent becomes aware of their decision (Trevena and Miller, 2002).

EEG signals in observed and imagined actions

The EEG signals which precede the observation of an action are very different from those which precede the execution of the same action (Babiloni *et al.*, 2003).

However, the EEG signals which precede the *mental simulation* of an action are quite similar to those which precede actual execution of the action (Jankelowitz and Colebatch, 2002).

- In fact, there are surprisingly few differences between neural activity for actual and simulated actions.

The crucial question: are there differences between *imagining executing* an action and *imagining observing* the same action?

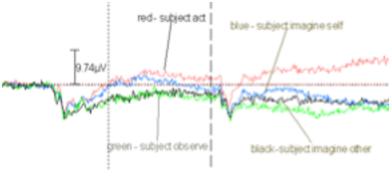
EEG signals in observed and imagined actions

We did a preliminary EEG study to look at this question.

Subject and the experimenter sat on opposite sides of a table, with a cup in the middle. A preliminary signal indicated one of four conditions:

Subject grabs cup	Experimenter grabs cup
Subject imagines grabbing cup	Subject imagines exp'ter grabbing cup

The action (real or simulated) was triggered by a second signal.



Between the two signals, there's a main effect of agent (self vs other), but not of modality (actual vs imagined).

Mode-setting and match models combined

Problem for the mode-setting model:

- It doesn’t seem to explain the difference between the two conditions in Farrer and Frith’s experiment.
  (In each condition, the subject presumably ‘decides to act’, but there’s only a ‘sense of agency’ in one condition.)
- The ‘sense of agency’ is much better explained in terms of the match model.

One way of reconciling the two models: distinguish between the *experience* of agency and the *representation* of agency.

- We only get the *experience* of agency when we actually act.
  (When we simulate an action we don’t get it.)
- To *represent* agency, we need to make reference to a mode-setting operation.

Summary

- Babies have to *learn* a concept of ‘self’.
- They also have to develop concepts of other agents.
- These concepts seem to be *acquired* using correlations between perceptual and motor signals.
- During actual experience, correlations between motor signals and reafferent *somatosensory* signals appear to be associated with a sense of agency.
- But I suggest that the main mechanism for distinguishing between one’s own actions and those of others is a ‘mode-setting’ operation, which configures the circuitry in the mirror system for action execution or action observation.