Functionalism, Intentionality and artificial intelligence[1]

This handout follows the one on 'Functionalism'. You should read that handout first.

INTENTIONALITY

Many mental states are 'about' something, objects or events in the world. For example, I might have a belief about Paris, a desire for chocolate, be angry at the government, or intend to go to the pub. In all these cases, my state of mind is 'directed' towards an 'object', the thing I'm thinking about (Paris, chocolate, the government, going to the pub). This idea of 'directedness' is known as 'Intentionality' (from the Latin *intendere*, meaning 'to aim at'). Intentionality is not about intentions (to mark the difference, I shall always use a capital 'I' for 'Intentionality'). If I have an intention, I am 'aiming at' doing something. With Intentionality, it is the thought or mental state which 'aims at' its object, what it is about, and no 'doing' needs to be involved. Beliefs, desires, emotions all have Intentionality; they are all 'Intentional mental states'.

An Intentional mental state is a mental state with Intentional content. So what is this? Whenever we think of, have a belief about, or desire something, we always conceive of it in a certain way, under a particular description. For example, in Sophocles' famous play *Oedipus Rex*, Oedipus kills his father and marries his mother. He doesn't want to do this. But it turns out that he doesn't know who his parents are. He doesn't know that the people he thinks are his parents aren't. On his journey, he meets an old man in the road who gets in his way. Oedipus becomes very angry, and kills the old man. In fact, the man was his father, Laius. Oedipus was angry at the old man. Was he angry at his father? From his point of view, he wasn't – he didn't think of the old man as his father.

So Intentional states represent the world in particular and partial ways. It's like seeing something from a particular aspect; you can see it, but not all of it. What Intentional states represent – Paris, the government, Laius, snow – is called the 'Intentional object'. The way they represent that object we can call the 'aspectual shape' of the object. The Intentional object + the aspectual shape comprise the Intentional content.

We can have different mental states with the same Intentional content if we take different 'attitudes' to that content. For example, I can believe I'm arriving late; I can want to be arriving late; I can fear I'm arriving late; I can be pleased I'm arriving late. An Intentional state, then, comprises a particular 'attitude' or 'mode' towards a particular Intentional content.

---

[1] This handout is based on material from Lacewing, M. (2010) *Philosophy for A2: Unit 3* (London: Routledge), Ch. 1, pp. 30-4

PROBLEMS FOR REDUCTION

Intentionality poses this challenge: how is it possible for anything physical to have the property of Intentionality? Physical things are never 'about' anything. To say what it is for a physical thing or state to be the thing or state that it is doesn't require reference to something else. A particular molecular structure or physical process, described in these terms, is not about anything. But the states and processes of your brain are just chemical states and processes. So how could they ever be about anything? So how could Intentional mental states be states of your brain?

Functionalism claims that Intentionality is a functional property. A physical thing can have the property of Intentionality because of its role in a network of causes and effects. But how do we get from the idea that a belief about dogs, say, is caused by certain experiences of dogs and causes certain behaviour towards dogs to the idea that beliefs are about dogs? Compare: clouds are caused by water evaporating from the sea and they cause rain; but clouds aren't about the sea, and they don't represent rain.

Functionalists maintain that if the causal network is complicated enough, then states within the system are not just causal, but Intentional. The last 35 years have seen large research programmes attempting to develop an account of the necessary causal network.

Intentional states 'represent' the world. We can understand this in terms of information. So philosophers have started from thinking how natural things can carry information, e.g. smoke carries the 'information' that there is a fire; fingerprints carry 'information' related to the identity of the person, and so on. However, notice that it makes no sense to say that smoke is 'mistaken'. By contrast, Intentional states can be mistaken. For example, someone might believe that the capital of Germany is Frankfurt, not Berlin. They would be wrong. Getting this idea of a mistake out of a model of natural 'information' has proven very difficult.

Some philosophers argue the function of beliefs is to represent the world truly. We can get the idea of 'mistaken' (false) beliefs from this; such beliefs are the result of our belief-forming process 'malfunctioning'. We can compare this to biological organs. The function of the eye is to see, yet eyes can malfunction in many ways. We identify the function of something biological in terms of evolution – what was the organ 'selected for'? This gives us a standard by which we can talk about something being (or going) 'wrong'. We can then use this idea of function as a starting point for Intentionality.

However, this is still not enough. For example, the stomach has the function of digestion, and it can fail to perform this function. But states of the stomach are not about digestion (or food). They are not Intentional states at all.

WHETHER ARTIFICIAL INTELLIGENCE IS INTELLIGENT

Artificial intelligence is a research project in computer science that aims to create computers that display behaviour that is 'intelligent'. Some philosophers and scientists argue that the test for whether a computer is intelligent is the 'Turing Test'. A person, a computer and another person (the interrogator) are each in a different room. The interrogator puts the same questions, in turn, to the person and the computer, not knowing which is which. If, after five minutes, the interrogator cannot tell from his conversations with the person and the machine which is which, then the machine has passed the Turing Test.

Whether the Turing Test is a good test for intelligence is very controversial. But we shall look just at an objection that claims that, unless we develop computers with consciousness, they cannot be genuinely intelligent.

John Searle argues that Intentionality is not reducible to functions. (*Minds, Brains and Programs*) To illustrate his argument, he describes a room with two holes in the wall; through one, pieces of paper are passed in (inputs), through the other, pieces of paper are passed out (outputs). There is someone in the room, who has to respond to the inputs by sending the outputs. The inputs are question in Chinese; the person doesn't understand Chinese, but has a huge book which correlates every question with an answer. He finds the output that is that answer, and sends that piece of paper out. The room as a whole – the system – 'behaves' as if it understands Chinese! But it doesn't – the person doesn't, the rulebook doesn't, the room doesn't. Even if the person memorized the rulebook, he wouldn't understand Chinese; he wouldn't know what the questions mean. This is what real Intentionality requires. Yet the room performs the same functions as someone who does understand Chinese, answering questions. So performing functions isn't enough for understanding meaning, for real Intentionality.

Some functionalists have rejected Searle's conclusion: the room (operating as it does, with the person inside) does understand Chinese. Consider this development of the thought experiment. Since the person inside the room doesn't understand Chinese, let's just have a computer that does the same thing (it is programmed to follow a rule-book). Suppose we then put the computer inside a robot, which interacts with the world. We add a programme for correlating visual input (through cameras in its eyes) to output – so the robot can now 'name' things in Chinese, as well as answer questions in Chinese. Is it now so obvious that the robot doesn't understand Chinese?

Searle would still say that it doesn't. Artificial intelligence isn't intelligence at all; at best, it is a simulation of intelligence. Other functionalists argue that to understand is nothing more than being able to interact with the world (including other people) in the right causal-functional way. Although we haven't been able to give a complete functional analysis of how this is done, this doesn't mean that we will not be able to give one in the future. When we can, we will be able to create genuine robotic intelligence.

Searle replies that the difference between Intentionality and a simulation of Intentionality is consciousness. Without consciousness, a series of functional

interactions remain meaningless to the robot even when they look meaningful to us. But Intentional states are meaningful ‘from the inside’ – they are meaningful to the creature that has them. Without consciousness, meaning is lacking, and therefore, so is Intentionality.