Cognitive Selflessness

Abstract

This essay explores the nature of selflessness from a cognitive perspective. The philosophical doctrines of self, personal identity, and selflessness have a long history, but their inclusion into modern scientific paradigms has been somewhat neglected. The integration of the notions of self and selflessness into modern psychology can be achieved by drawing analogies between several philosophical and psychological frameworks. In particular, System 1 (an intuitive aspect of mind) and System 2 (a rational aspect of mind) correspond to selfless (non-dualistic or non-conceptual) and self/other-based (dualistic or conceptual) aspects of the mind. Details of this analogy are presented which yield useful insights into both domains.

Introduction

From a psychological perspective, mental selves correspond to concepts. Concepts often correspond to continuous objects, but more generally they are concrete or abstract things that can be named. Accordingly, cognitive or mental selflessness entails cognition which is selfless or non-conceptual, and which does not involve naming. Systems which operate in intuitive and rational/conceptual modes have been thoroughly studied in psychology for a long time under different names, such as conceptual and non-conceptual mind. This essay refers to these two mental processes using modern terms popularized by Daniel Kahneman: System 1 and System 2.

According to both traditional wisdom and neuroscience, the mind can and does operate non-rationally when it operates intuitively via System 2; it operates in a massively parallel way through the use of associations. For System 2, it is not possible to activate a single concept without activating many others. In other words, it operates in a way which we might identify as sub-symbolic (Churchland) or non-conceptual (Nyima). Although the rational mind is also capable of selfless thoughts, its inherently serial perspective-taking render it slower than intuition to adopt all of the multiple perspectives that are necessary for a selfless view that is fast enough to enable action.

Our claim with respect to dualism and non-dualism is that System 2 and System 1 are respectively characterized by those modes of knowing. This does not mean that System 2

always operates altruistically, but since it operates non-conceptually, there is some sense in which it operates selflessly.

System 2

System 2 is a system in the brain which operates rationally (linguistically), and processes information serially. According to Daniel Kahneman, "System 2 allocates attention to the effortful mental activities that demand it, including complex computations. The operations of System 2 are often associated with the subjective experience of agency, choice, and concentration". We present System 2 first, because even though it is an older system in terms of human evolution, humans tend to be (verbally) more familiar with it.

System 2 operates with (abstract) concepts. It is responsible for logic, language, reasoning ability, and many of the things we associate with human intelligence. It operates serially, however, and so it is ridiculously slow compared to System 1. It is also biased in some ways that we might not recognize. Although the logic behind rational arguments can be checked, we cannot check why we put forward one rational argument as opposed to another: the source of rationality is not rationality itself. In some sense, System 1 gives rise to System 2, and its influence on System 2 cannot be eluded.

With respect to selflessness, System 2 traffics in generalities, the characteristic selves of a set of events. It isolates a particular narrative from the data that is presented to it, and that single story line is the one that is understood: it is that story line that we become aware of, and in so doing, we become unaware of the other story lines. Although it is not inherently selfish, it is inherently dualistic, and therefore it is incompatible with true selflessness.

Concepts

If we are going to understand the conceptual mind, it is essential to carefully define the term "concept".

Concepts are closely related to names. Although there is considerable debate about the exact relationship of concepts to names, here we will consider concepts and names to be capable of a one-to-one correspondence[1] (in practice, some concepts may not have names). So it may clarify things to say that concepts are things which can be named by the mind. They also share the criteria for selfhood mentioned in the overview (i.e. spatio-temporal independence and singularity).

Concepts are generalities: they are formed by generalizing over experience (that experience can be described as a set of events). The *extension* of any concept is the set of events over which it generalizes; the concrete concept corresponding to "New York City" has an extension which is the set of all events that have ever, or will ever, take place in New York. Concepts can also be abstract, such as "cities", which corresponds to a spatiotemporally discontinuous set of events. There is some debate with respect to the representation of that knowing as it is stored in out brain, such as whether a prototype city is stored or the connections between properties of each city, but that is not relevant here. The important part to understand is that however concepts are represented, the concept of a city is only a part of our knowing: in virtue of connoting grey and buildings, it does not connote color and open fields. The full set of all events has not been increased by the knowing of "city", but restricted.

To express this slightly differently, concepts are generalizations, and generalizations are formed by forgetting particulars. Therefore, concepts are formed by an act of forgetting: by forgetting everything not shared by instances of the concept, we restrict our full knowing of the world to knowing the part of the world referenced by that concept. The fact that concepts are formed by forgetting may initially be quite surprising, since most of us conceive of concepts as bearers of information. While in some sense concepts do bear information, if the intuitive mind is without concept (non-conceptual), it is not forgetting, and thus all knowing is present to it simultaneously[2].

System 1

System 1 is a system in the brain which processes information via association, in parallel. According to Daniel Kahneman, "System 1 operates automatically and quickly, with little or no effort and no sense of voluntary control". Because the development of System 1 preceded System 2, many people devalue System 1, linking it with instinctual behavior. This point of view is clearly incorrect: System 1 is much faster than System 2, it is responsible for intuitive judgement, and it learns throughout our lives. Although it can be subject to bias, it is arguably no more biased than System 2, and we could not conduct our daily activities without it. In fact, because the information that it contains is essentially a superset of the information in System 2, one might even argue that it is smarter.

We made the claim in the abstract that System 1 is nondualistic, as opposed to System 2 (which is conceptual and therefore dualistic). But what does it mean for knowledge to be nondualistic? In particular, if System 1 does not know "self" and "other", then in what sense

is it knowing? System 1 can be likened to a mass of unprocessed data, or at least data which has not been reduced to concepts. If we mine the data in particular ways, and apply particular functions to it, we can retrieve concepts or generalities. However, that does not mean that those concepts are represented *as concepts* in the dataset. As an example, imagine a set of photographs of the pages of a dictionary. Those photographs do not have any semantic content until we perform character recognition on the photographs.

Studies of episodic memory (which is photograph-like) and semantic memory (which is verbal or dualistic) show that System 1 and System 2 destructively interfere with one another. This is not surprising if they share their respective physical implementations overlap: if the meaning of concepts (in System 1) is derived by forgetting irrelevant details in a knowledge network that represents everything we know (System 2), then we would expect exactly this result. This organization seems likely for several reasons, especially because meaning is probably non-trivial in terms of its requirements for neuronal representation.

Associations

If non-conceptual knowing does not store concepts explicitly, then how might it store information? According to Platonism, which is a theory of universals, what is ultimately real are the concepts themselves, or the generalities. If we understand Platonism as a theory of mind, then System 1 need only store the associations between activated concepts. The degree of association between that concept and all other concepts is the meaning of that concept.

The neuronal implementation of this theory is known as associationism. In humans, it has been demonstrated that the representation of a concept is distributed among many nodes (or neurons). This finding adds support to the hypothesis that neuronal processing is sub-symbolic and parallel, but it leaves open the question of how symbolic processing occurs, since processing seems to be discrete (rather than continuous, which is how sub-symbolic systems operate). To ask a previous question in a slightly different way: if we are storing concepts, then in what sense is System 1 non-conceptual? Remember, however, that we are not storing concepts (at least when System 2 is not operating); System 1 stores conceptual information without ever becoming explicitly conceptual. In order for a concept to become explicit, it must be isolated from other concepts, so unrelated conceptual information must be actively suppressed. Again, the term non-conceptual does not mean unknowing: it means non-dualistic, and not explicitly conceptual (or not isolating of any particular concept).

Conclusion

To return to the issue of selflessness, being selfish entails not being otherish, and being otherish entails not being selfish. So any non-conceptual system such as System 1 does not suppress any information, and it is therefore incapable of "not being otherish" or "not being selfish". Therefore, System 1 is non-dualistic in the sense we have been talking about it.

Unfortunately, the non-dualism of System 1 does not immediately entail kindness to others, because System 1 has a first-person perspective and limited knowing. It does not inherently have an empathic response to others, since it does no have inherent knowing of the pain and pleasure of others. The reward and punishment systems of an organized society only go so far with respect to that reward and punishment for ethical behavior, so System 1 often falls short of complete altruism. System 2, on the other hand, knows a lot about others. Because System 2 learns in a completely different way (i.e. through communication rather than association), it has a breadth of third-person information available to it. But the third-person information which has been learned from System 2 does not immediately integrate with the information from System 1.

The integration of System 1 and System 2 requires that System 1 operates as if it had direct experience of the pain and pleasure of others. If the pleasure-seeking System 1 had this experience from the knowledge-seeking System 2, it would behave altruistically. Conversely, behaving altruistically (implicitly) teaches System 1 the knowing of System 2. The integration of the non-dualistic, self-knowing System 1 and the dualistic, other-knowing System 2 is only possible when the knowing of System 2 is "lived". In other words, the ethical universality (in the Kantian sense) of System 2 must be made non-dual by forming corresponding associations in System 1. That is only possible with ethical training; in particular, it is only possible by experiencing the associations that are a result of treating one's neighbor as oneself.

---

[1] Mathematically, we will consider the relation between names and concepts to be a bijection.

[2] This understanding of conceptuality is related to the Apoha theory of Dignaga.