Foreword

The central questions confronting artificial intelligence and cognitive science revolve around the nature of meaning and of mind. Minds are presumed to be the processors of mental content, where that content has the capacity to influence our speech and other behavior. The hard part is figuring out how that is done. Mental states must exercise causal powers that sometimes have the effect of bringing about other mental states (which take the form of shifts from one mental state to another as forms of thought transition) and sometimes have the effect of bringing about physical actions (which occur within a context that includes our beliefs, motives, ethics, abilities, and capabilities). Whether our actions are successful tends to depend upon whether things are as we take them to be (whether our beliefs are true) and the extent to which we have been successful in thinking things through (whether our actions are appropriate).

This volume brings together a collection of interesting studies that address the kinds of questions raised by this general conception and attempt to answer them. One of the underlying issues is how mental states themselves can stand for other things, such as objects and their properties in the world. This has sometimes been called 'the symbol grounding problem," because it concerns the "ground," or basis, that connects mental states to objects and properties in the world. But it is better viewed as "the sign grounding problem," insofar as Charles Peirce explained the distinction between icons (as signs that resemble what they stand for), indices (as signs that are causes or effects of what they stand for) and symbols (as signs that are merely habitually associated with what they stand for). The grounding problem applies across the board for different kinds of signs, but Peirce pointed to the general direction in which its solution lies.

If some signs resemble what they stand for, others are their causes or effects, and others are merely habitually associated with that for which they stand, then the kinds of things that have minds of the most elementary kind must have the capacity to detect and utilize resemblance relations between signs and things; those of an intermediate kind must have the capacity not only to detect and utilize resemblance relations be-

tween signs and things but also cause and effect relations between things, while those of a yet higher kind must have not only the capacity to detect and utilize resemblance relations between signs and things as well as cause and effect relations between things, but also have the capacity to acquire and exercise relations between signs and things that are merely habitual. This higher grade of mentality appears to exist in species other than human beings, but has its most extensive expression in our species.

If this approach appears promising for the purpose of illuminating the nature of mind and of meaning, it does not resolve the question of whether, under suitable conditions, it might be possible to create an inanimate machine with at least some of the mental powers of human beings. The lowest level of mentality would do for a start, where iconic minds have the capacity to utilize icons. On first consideration, it seems almost irresistible to infer that digital machines which process data on the basis of its shape, size and relative location must have at least iconic capacity, which in turn implies that they have at least iconic mentality. But for that to be true, it must not only be the case that digital machines process data not as a purely causal process (in the fashion of sieves sorting gains of sand) but that they do so because those shapes, sizes and relative locations resemble that for which they stand. And what is that?

Absent "stand for" relations, processes can be causal without being mental. In terminology derivative of Peirce's theory of signs (or "semiotic"), the question is whether inanimate systems are causal systems that process signs (as "semiotic systems"). That digital machines can become more and more complex with greater and greater capacity (through ingenious design and clever programming) to mimic thought and behavior is not the question. Without any doubt, they can do this with increasing sophistication, as generations past have amply displayed. The question is whether they can overcome the gap between mere causal systems and thinking things, as causal systems of a distinctive kind. If it is possible to establish that inanimate machines can process signs (where those signs stand for something for those machines and not simply for the users of those machines), it will be the occasion to agree that artificial intelligence has been attained.

Of course, the creation of inanimate thinking things is not the only function that digital machines can perform. Throughout human history, after all, there have been three great revolutions: the agricultural revolution (in putting nature to work for us growing food), the industrial revolution (in putting machines to work for us performing physical labor) and the computer revolution (in putting machines to work for us performing tasks that in the past required human minds). The computer revolution continues to make a permanent difference in our lives. The question that remains, which these studies address, concerns the gap that remains between digital machines and thinking things. Whether or not the studies succeed in closing that gap, readers of this book are bound to agree that they provide important analyses and stimulating suggestions that assist us in assessing how close we are coming to creating artificial thinking things.

James H. Fetzer