Inductive Kinds

I will use the term "inductive kind" for a more general notion that includes natural kinds, and also artifactual kinds that work similarly. We start with a "quality space" that includes every (projectible) property we can ascribe to objects. Some of these will be binary, and others will admit of degrees. We locate objects in the space in terms of our beliefs about what properties they have. When we notice a clustering of objects, we take that to initiate a new inductive kind.

Functional Role

Computationally, inductive kinds are "buckets" with memory locations. Each bucket has a syntactical item (a "kind term in the language of thought") associated with it that is used for thinking about objects judged to be of that kind. All of the buckets are governed by the same functional rules (they are like proper names in this respect). They differ only in their memory addresses and the associated syntactical item (which can be viewed as a function of the memory address). From the clustering of objects, we distill statistical information about the cluster in the quality space, and encode that syntactically in the bucket using the associated kind term. We also record attributions of the kind to the individual objects in the cluster. This constitutes storing statistical beliefs about the kind, but notice that the operation of storing them is purely syntactical. We do not need a "concept" for the kind — just the syntactical item.

Given this "bucket of beliefs" about the kind, we can then employ the beliefs for further reasoning about the kind. That will lead to new attributions of the kind to additional objects, and the discovery of new generalizations. Including new objects in the extension of the kind, and adding new generalizations to the bucket of beliefs, can lead us to modify or reject previous generalizations. That in turn can have the effect of shifting the cluster in the quality space, so that some objects originally included get excluded. Thus "water" starts out including urine, but with the discovery of chemical generalizations that gets excluded and the cluster migrates.

Having introduced the new kind into our ratiocinative repertoire, we have added a dimension to the quality space. That can have the effect of altering other clusters and adding or deleting beliefs from other buckets. That is actually what happens with "water". The rise of chemical knowledge brought with it new kinds like "H_2O", which enabled new generalizations about water and its relationship to H_2O. In the expanded quality space, we get a better clustering if we exclude urine from water. I need to get a little clearer on exactly how this works.

Sometimes we start with a definition, e.g., atoms as the atomic constituents of matter, and electrons as the smallest negatively charged particles. But we don't treat these as definitions. We instead take them to initiate an inductive kind. Later discoveries then lead us to reject the initial characterizations. It looks like the definitions serve initially as reference fixers rather than real definitions. How does this fit with the clustering idea? We don't initially have a cluster, but it looks like we expect there to be a cluster because these seem like scientifically useful kinds.

Note that what kind gets associated with what bucket is an historical accident. If we learn about cows before horses, cows get assigned to an earlier memory address, but if we learn about horses before cows this is reversed.

Concepts

All of the above is syntactical and computational. It seems relatively easy to understand how we use these inductive kinds in reasoning and drawing conclusions both about what objects are instances of the kinds and what generalizations hold about the kinds. Now what about semantics? Is there a concept associated with each kind? Well, what's a concept? The traditional motivation for talking about concepts was that they were constituents of beliefs that can be shared by different people, so that we can talk about different individuals believing the same thing. It is

not clear how concepts are supposed to work in lots of respects, but one clear supposition is that concepts have determinate extensions. By this I do not mean to rule out the possibility of borderline cases. But for every object, there is a fact of the matter whether it exemplifies the concept, fails to exemplify it, or is a borderline case.

Now let us look at inductive kinds. First consider extensions. Imagine a kind initiated by the observation of a cluster in quality space. On the strength of that we adopt a number of statistical generalizations about the kind. But they may form a rather sparse set which does not resolve lots of questions. Contrast swans and horses. Suppose we learn about swans in Europe, and one of our beliefs is that all swans are white. Then we go to Australia and see things that look otherwise like swans, act like them, etc., but are black. They fall within the cluster in the quality space, so we judge them to be swans and reject our earlier belief that all swans are white. As we learn more about them, they stay there. In particular, when we learn about DNA we discover that white swans and black swans have essentially the same DNA. Next, imagine learning about horses in Europe, and then going to African and seeing zebras. It would be natural to think they are striped horses. They certainly look like horses otherwise, and act like horses. But subsequent discoveries about their DNA reveal that they are only distantly related to horses, so at that point we exclude them from the cluster. (Similarly for pigs and boars or javelinas.)

We can talk about the *judgeable extension* of a kind. This is the set of objects that the cognizer could, at any given time, reasonably judge to be instances of the kind. I mean this to include objects the cognizer does not know about but would be able to judge to be instances of the kind if he or she did know about them. We can similarly talk about the *judgeable converse extension*, which is the set of objects that the cognizer could judge not to be instances of the kind. We can make two observations. First, there can be lots of things that do not fall in either extension. They could be regarded as borderline cases, but this isn't because they are "almost like the things in both extensions". Rather, my information about the kinds just doesn't enable me to decide whether these objects are instances of it. The second observation is that the judgeable extensions change over time. For instance, the judgeable extension of *horse* originally included zebras, but zebras were subsequently excluded when we added new generalizations to the bucket.

If there is a concept associated with the kind, it must have a determinate extension. We might try identifying the extension with the judgeable extension, but the latter may change each time we obtain new information about the kind. If that also changes the extension, it must change the concept. So concepts become very transitory. If the concept changes every time our thoughts about the kind change, then concepts, and the thoughts containing them, become like Heraclitean rivers. We can hardly ever have the same thought twice. This makes concepts completely useless for their avowed purpose of enabling different individuals to have the same beliefs. If a person cannot even have the same belief as his own earlier self, it becomes extremely unlikely that two different people will every have the same belief, regardless of how we spell that notion out.

To avoid this, we must insist that the actual extension of the concept is different from the judgeable extension. But what could possibly determine a different extension? All there is to inductive kinds is the bucket of beliefs we have about them. There seems to be no way to pick out a different extension.

Consider a different problem. The way in which a cluster in quality space migrates can be partly an historical accident. Millikan (1847?) suspended charged oil drops in an electric field whose intensity he could vary, and observed that the charges all came in integral multiples of a small value. He hypothesized that there was a smallest unit of negative charge, and he called its instances "electrons". This was well before the rise of atomic theory. With atomic theory people formed many more beliefs about electrons. Eventually, atomic theory led to the discovery of quarks, and the discovery that some quarks have a negative charge 1/3 that of electrons. Imagine the history of all this compressed into the lifetime of a single physicist. His inductive kind *electron* started out having a judgeable extension consisting entirely of quarks. The *judged* extension was quite different of course, but that turned on a mistake. Eventually, however, the judgeable

extension came to consist of electrons. Furthermore, by the end of the story the physicist has a newer inductive kind — *quark* — whose judgeable extension is the same as the original judgeable extension of *electron*.

Now imagine a different scenario. Suppose that after measuring oil drops and hypothesizing the existence of electrons, Millikan went on to develop more sophisticated detectors that were able to detect quarks. He measured their charges, and concluded that they had 1/3 the charge he observed on the oil drops. He might have concluded, "I was right in thinking that there are such things as electrons, but it was just dumb luck. What I was measuring on the oil drops was not electrons at all." We can imagine that physics went on to discover that the charges on the oil drops also represented particles, and those particles play an important role in atomic physics. Of course, they would not have been called "electrons", because that term was already taken for talking about quarks.

The starting points and endpoints of these two stories are the same, except for the words employed in talking about quarks and electrons. It is tempting to say that in the first (real) case, Millikan's concept of an electron changed, because initially Millikan was thinking about quarks, and in the end he was thinking about electrons; but in the second case the concept did not change. But there seems to be no way to make sense of this. The judgeable extension of *electron* changes in both stories. In the second story, the (imaginary) physics of his day would not have enabled Millikan to identify most quarks as quarks. At the end of the story he is able to identify many more quarks. So if his concept of an electron remains unchanged, a mere change in judgeable extension cannot be sufficient to change the concept. But then what reason do we have for thinking that the concept changed in the first case? The only difference between the start and the end of the story is that the judgeable extension is different.

It seems to me that there is no way to make sense of the idea that there is a single concept that persists through changes in the judgeable extension and or beliefs about electrons. If these changes eventually lead to changes in concepts, then it seems that they must immediately lead to such a change. But that makes the concept of a concept pretty useless here.

There is one thing that remains unchanged through all this. That is the syntactical item that the cognizer uses for thinking about the kind. Might we just identify the concept with the syntactical item? That would make it well-defined, but it would not serve most of the purposes concepts are supposed to serve. For example, concepts would not have determinate extensions, and on this construal there is no sense to be made of two people having the same concept. Even if we could somehow identify memory addresses across cognizers, your kind *cow* may have the same memory address as my kind *horse*.

So my conclusion is that it makes no sense to talk about concepts in connection with inductive kinds. They play a purely syntactical/computational role in thought, and there is both no way to make sense of concepts associated with them and no need for concepts.

Anti-individualism

Tyler Burge has urged that a concept like *arthritis* cannot be understood just by talking about the cognition of a single cognizer. For its use, it is essential that cognizers are part of a community of cognizers. He calls this "anti-individualism". I want to deny his conclusion. First, I don't think there is any such thing as the *concept* arthritis. This is an inductive kind, and as such it is purely syntactical. How then do we explain Burge's "arthritis" example?

Imagine a person who has heard the term "arthritis", but knows almost nothing about it except that it involves pain. He forms the belief that he has arthritis in his thigh, and goes to his doctor with this complaint. But his doctor informs him that you can only have arthritis in a joint. Our sufferer accepts this, and concludes that he does not have arthritis after all.

How can we understand the rational dynamics of this? Among the beliefs that our cognizer has in his *arthritis* bucket are the belief that arthritis is called "arthritis", and that his doctor knows a lot more about arthritis than he does. On the strength of the first belief, he can communicate

with his doctor about arthritis and on the basis of the second belief he can use his doctor's pronouncements as a source of information about arthritis. So by thinking of arthritis as an inductive kind, his reasoning becomes readily explicable on a purely individualistic basis. There is nothing special about the inductive kind *arthritis* here.

Public Language and Inductive Kinds

We are led to the conclusion that different people cannot have the same beliefs about inductive kinds. A single individual can have the same thought at different times as long as that is understood syntactically, but there is no way to compare syntax across individuals. But then how can people talk to each other about inductive kinds?

In broad terms, the answer is that language conveys less than the whole belief. For successful communication the audience must "receive" thoughts that are appropriately related to the thoughts the speaker is "sending", but they need not be (and in fact, cannot be) the same thoughts. What exactly is the connection? It is sometimes supposed that the only requirement is that denotations and extensions be the same. But that does not work for inductive kinds, which do not have well-defined extensions. And it is clearly too much to require that the judgeable extensions are the same. In the arthritis example, the judgeable extension of *arthritis* for the patient is quite different from that of the doctor, but they can still communicate. The communication seems to be mediated by the belief that arthritis is called "arthritis". This suggests that we can construct an account of the rational dynamics of communication without being able to define success. This needs to be explored.

Truth

If inductive kinds do not have determinate extensions, then the truth of beliefs about them is not determinate either. This does not mean that we cannot talk about truth, but its use is purely formal, via the Tarski T-schema, and its purpose is to give us a way of quantifying over beliefs.