August 25, September 20, 21, 2010; March 8, 2011

Dense and Sparse Meaning Spaces

Comments on Travis Norsen, "Scientific Cumulativity and Conceptual Change: The Case of Temperature"

John D. Norton[1]
Department of History and Philosophy of Science
Center for Philosophy of Science
University of Pittsburgh
www.pitt.edu/~jdnorton

Prepared for Richard M. Burian and Allan Gotthelf, eds.,
Concepts, Induction, and the Growth of Scientific Knowledge,
forthcoming.

Contrary to the incommensurability thesis, I argue that the referents of theoretical terms can remain stable under theory change, if they are associated with "sparse meaning spaces." In them, reference is error tolerant, for there are no alternatives in the neighborhood to which terms in altered descriptions can shift their reference.

1. Introduction

Professor Norsen has written a fine study of the historical changes in the concept of temperature. It is distinctive for the care with which he has traced what is truly essential in the historical development of the theoretical notion of temperature. Although the evolution of the

[1] My thanks for helpful discussion to Travis Norsen, P. D. Magnus and participants in the Workshop on Concepts, Induction, and the Growth of Knowledge, Pittsburgh, September 17, 2010.

concept of temperature is a long and tangled story, Professor Norsen keeps his focus commendably on just those parts needed for the philosophical work he wishes to undertake.

In this regard, a major goal of Professor Norsen's analysis is to refute a claim by earlier writers like Kuhn and, more explicitly, Feyerabend, on the meaning of terms like temperature. Feyerabend points out that the background assumptions change when we move from a purely thermodynamic analysis of temperature to a statistical one. In the first, heat only passes from hot to cold. In the second, when statistical fluctuations arise, heat can also pass from cold to hot, in violation of the Second Law of Thermodynamics. Therefore, Feyerabend concludes, the two terms cannot have the same meaning. A lot is at stake with this claim. It has been taken in the wider literature to underwrite a further claim of the incommensurability of theories under theory change.

Exactly how we are to understand "meaning" in this claim is not entirely clear. If we are to get some purchase on Feyerabend's claim, we need some characterization what meaning is. In the following, I will take the meaning of a term to be its semantic content—that to which it refers. As a result, meaning and reference will come out to be the same thing. There will be many possibilities for the meanings, the referents, of some particular term and, at this stage, I am leaving open just what sorts of things they may be. They may be abstract entities like sets; or possible physical states; or actual physical states. Whatever they may be, in each case there will, in general, be multiple possibilities for the meaning of particular term. These possibilities form a space of meanings. One identifies the meaning of a term by identifying the map from the term to some particular meaning in this space.

Professor Norsen dissents from Feyerabend's analysis and urges what I believe is the correct conclusion. A term can retain its meaning when background assumptions change. My remarks here will serve only to sharpen the problem he addresses and explain why I think his conclusion is correct. In particular, I will seek to diagnose why our philosophical community is divided by the question. Some readily agree with Feyerabend that changing background assumptions alters meaning; others do not. I will conclude that the two groups are divided by their assumptions about the character of the spaces possible meanings can form. The first group assumes a dense meaning space so that small changes in background assumptions enable a term to attach to a new meaning. The second group assumes a sparse meaning space, so that there are no nearby meanings to which a term might attach, when there are changes in background assumptions. This affords some stability of meaning, since small errors in description can be discounted and even corrected.

2. Energy as a Surrogate for Temperature

Before proceeding to this analysis, I will provide a version of Feyerabend's objection that enables us to see a little more clearly how differences in background theory might disrupt reference.[2] Ordinary thermodynamics tells us that the energy E of some sample of an ideal gas at temperature T is related to the amount of the gas, measured by n, the number of moles of gas, and the temperature T as:

E is proportional to n.T

This relationship holds no matter how small the sample of gas. This means that we can use the energy E as a surrogate for temperature. It is true for n=1, a single mole of gas, such as 2 grams of hydrogen. It is true when n gets to be very small, even when n is of the order of 10^{-25}.

This relation no longer holds exactly in statistical mechanics. The energy of a quantity of an ideal gas will fluctuate slightly as its molecules lose or gain energy in exchanges with the thermal environment. For one mole of a gas, the fluctuations are an imperceptible fraction of the total energy. It is not so for 10^{-23} mole, for we now typically have only one or so molecules present. The energy of each of those molecules will vary with time and will be distributed probabilistically over a wide range of energies. While the average, the expected energy, will continue to conform to the above formula, the energy E of the sample will at any moment be very different from the expectation value. Hence the energy E of a tiny sample at any moment will be a poor surrogate for its temperature.

This failure of surrogacy of energy for temperature can disrupt reference. In ordinary thermodynamics, doubling the energy of a small sample of an ideal gas refers to a process in which the temperature doubles. In statistical physics, the doubling of the energy of a sample of

---

[2] The idea that fluctuations automatically produce violations of the Second Law of Thermodynamics is one that is easy to believe. However the proposal is the starting point of a ponderous literature on Maxwell's Demon that seeks to establish that the violation cannot be secured. Our present purpose does not require us get embroiled in this mess.

ideal gas might merely refer to a thermal fluctuation in the energy of the few molecules comprising the sample, while the temperature remains constant.

3. Two Cases

Does the temperature of thermodynamics and statistical mechanics have the same meaning in the two theories, in spite of these differences? Since I am construing meaning as reference, one might seek to answer this question by calling up theories of reference already in the philosophical literature. I do not think, however, that these theories alone can decide the question. However, we can see how the question can be decided by looking at two extreme cases of unstable and stable reference. They happen to coincide with motivating examples from descriptivist and causal theories of reference, respectively. But that does not mean that we must attach instability of reference to the descriptivist theory and stability to causal theory. For present purposes, that can be left as an open question.

4.1 Mathematical Functions

First, consider the case of functions in mathematics. Take the function y = f(x) that passes through the origin (x,y) = (0,0) with unit slope and:

has everywhere zero second derivative; or

has a second derivative equal to itself, negated.

The first is just the linear function y = x. The second is a sine function, y = sin x. These are contradictory descriptions. A function on numbers is simply a set of ordered pairs of numbers. Plainly, the two descriptions refer to different sets. That difference stems directly from the fact that the two descriptions ascribe contradictory properties to the function. Setting aside familiar philosopher's tricks,[3] this will generally be the case; if the descriptions of functions contradict, then they do not refer to the same function. In this case, this difference of referents does not arise

[3] By "philosopher's tricks," I mean maneuvers like taking two compatible descriptions and rendering them contradictory by appending irrelevant but contradictory statements. For example, logically compatible definitions D1 and D2 are made contradictory by taking any contingent proposition X and forming new definitions D1&X and D2&(not-X).

from confusing the description with the referent. There are very many different, logically compatible descriptions one can give that refer to the same function.[4]

The case of the linear and sine functions has been chosen to mimic an important feature of the relationship of thermodynamics and statistical mechanics. These two theories agree arbitrarily closely when we look at large samples of matter over short times. While energy E is a surrogate for temperature T in the ideal gases of thermodynamics, it is not a perfect surrogate in the statistical analysis of ideal gases. It fails completely for very small samples of an ideal gas, but comes closer and closer to surrogacy as we consider larger samples.

One might want to say that the approach to agreement is enough to establish that the term temperature in the two cases has the same meaning. Arbitrarily small differences eventually just do not matter to reference. However, if one takes the case of mathematical functions as one's model, that conclusion is blocked. Differences, no matter how small, always matter. By considering smaller and smaller neighborhoods around the origin $y=x=0$, one can bring the linear and sine functions arbitrarily close to one another. However, no matter how close they come in these smaller neighborhoods, they can never come close enough to be the same. The two functions are different over any domain, no matter how small, excepting the trivial domain with one point, $x=0$.[5]

If the case of temperature is like that of functional descriptions, then it does not matter how closely the thermodynamic and statistical properties of temperature approach in selected domains. They are always referring to different things.

4.2 Proper Names

Matters go very differently with the referents of proper names, the motivating case for Kripke and Putnam's causal theory of reference. It is a familiar occurrence that contradictory descriptions of entities with proper names can have the same referent. Here's a simple example. The city that we otherwise know as Jerusalem is described variously as:

"the city in which the Temple was built" by Jews;

---

[4] I think this one also works for the linear function $y=x$: that strictly increasing function that is its own inverse.

[5] This sole point of agreement is eliminated by considering functions over domains with $x>0$.

"the city in which Jesus, son of God, was crucified" by Christians; and

"the city from which Mohammed, God's true prophet, ascended to heaven" by Muslims. Each of the three religious groups harbor contradictory background assumptions. While the contradictions are not immediately apparent, they are there. Christians, for example, would contradict the Muslim's description of Mohammed; and Muslims would return the favor with Jesus. Yet they all refer to the same city.

If temperature is like these proper names, then the term can still refer to the same thing, even though the term may appear in two different theories, thermodynamics and statistical mechanics, with contradictory background assumptions.

5. Error Tolerance

Which of the two cases is temperature like? Prima facie, temperature is unlike a proper name. The latter denotes a particular thing that can be picked out by ostension. Temperature denotes a theoretical property of equilibrium thermal systems; and it cannot be picked out by ostension. Merely pointing to a hot oven, no matter how artfully, falls far short. Temperature can only be picked out by engaging in a theoretical discourse. One might point to the level of Mercury in a thermometer and then explain the theory that assures us that it measures the property, temperature, defined within the theory. In this regard, temperature is akin to the description of functions; in both cases, the theoretical background in which the terms appear is essential for determining to what they refer.

While temperature is akin in this regard to mathematical functions, that similarity is not the one that decides the stability or instability of the meaning of the term. What decides stability is the extent of error tolerance in the specification of meaning. The crucial difference between the two cases of mathematical functions and proper names lies in this:

- The definition of mathematical functions is highly intolerant of errors. The slightest change in the definition can lead it to pick out a different function (or none at all if the new definition is self-contradictory). We cannot refer to the linear function by describing a function, whose second derivative is everywhere zero, except in the interval $101<x<113$, where is it something else. That is just a description of a different function.

- The designation of the referents of proper name terms exhibits far greater tolerance for error. One can have contradictory descriptions, so that at least one and possibly both descriptions are erroneous. However they can both still refer to the same thing. Christians and Muslims disagree on many facts about Jesus and Mohammed, so that at least one is error-ridden. Yet, they agree on enough of the use of maps to locate the city in which the two died and to determine that they are referring to the same place.

6. Density and Sparseness of Meaning Spaces

What matters to stability or instability of reference is error tolerance. This tolerance derives from the structure of the space of meanings. In the case of mathematical functions, the referent of terms like the linear function and sine function is a set of ordered pairs. The space of all such sets form the space of meanings of function terms. This space is dense in that the slightest change in a definition of a function can reattach the function name to a different set of ordered pairs. There is such a richness of possibilities that any error in the description of a function is likely to issue in a changed meaning.

One might think of throwing a dart at a dartboard. The slightest change in the throw will lead the dart to hit a different point on the board, which is dense in points.

In the case of proper names, the meaning space is a sparse set. In the case of cities, there are no competitors for major cities in the space that might be just like Jerusalem in almost all aspects. As a result, when Christians and Muslims make contradictory claims about the city, the reference is unaffected. In principle, reading uncharitably, each might say of the other that they have specified no city at all when their descriptions make false claims of the intended city. In practice, however, we select out enough of the facts upon which the two agree to see that they each have the same referent.

The situation is more like the game of quoits. In it, ring-like hoops are thrown over an upright peg. To score, one need not throw the quoit so its exact center falls on the peg. Any throw that places the peg somewhere in the interior of the quoit succeeds. Slight errors in tosses do not compromise scoring.

7. The Stability and Instability of Meaning of Temperature

What is the meaning space of terms like temperature? Here one's prior assumptions about the world are decisive. Is one a realist or a constructivist?

7.1 Stability for Realists

If one has realist inclinations as Travis Norsen and I do, we are seeking to attach terms like temperature to things in the world. Precisely how the attachment goes is an issue for analysis elsewhere. We might attach temperature to a property of things; or we might conceive of properties extensionally: temperature partitions things in the world into equivalence classes of things at the same temperature; and these equivalence classes admit a transitive numerical ordering.

Under this realist view, the space of meanings is sparse. There are other properties in the world that divide things up in superficially similar ways. Among the familiar ones, the properties of mass or volume or average density divide things into equivalence classes that also admit transitive numerical orderings. However these other properties are unlike temperature in so many ways that quite radical changes would be needed in a theory of temperature before the referent of the term could come out as mass, for example.

This sparseness provides a great degree of error tolerance. All the facts mentioning temperature in ordinary thermodynamics cannot be true of the temperature of a world in which statistical mechanics gives the correct account. Thermodynamics and statistical physics will disagree markedly on some facts in the realm of the very small, such as whether energy is an exact surrogate for the temperature of ideal gases. However they will agree closely enough on macroscopic systems for us to discount the errors of thermodynamics. If its term "temperature" refers at all, we can recognize that it refers to the same thing as the term does in statistical mechanics. There is no other candidate in the meaning space in the vicinity. Thermodynamics errs, however, in some of the facts it ascribes to temperature. However these errors are not sufficiently great to force the terms of the two theories to have different referents.

Error tolerance is not unlimited for realists even with sparse meaning spaces. It fails, for example, in the case of the term "Mercury," as it might be used by a modern chemist and by the alchemist, Paracelsus. Superficially their referents may seem to agree on the familiar liquid metal we find in modern thermometers. However the divergence of referents become clear when we realize that a surrogate for Mercury for Paracelsus is that it is one of the three basic principles that form all substances; and it is the one whose presence gives materials their fusibility and volatility. Paracelsus would likely not allow that the modern chemist's pure sample of Mercury is a pure instance of his Mercuric principle. From the perspective of modern chemistry, Paracelsus' use of the term goes beyond the limits of error tolerance. He is trying to refer to something that does not exist.

7.2 Instability for Constructivists

Matters can turn out differently if one is not a realist. One might adopt a constructivist position and assert that theoretical descriptions create their referents. One might say, for example, that the space of meanings is created by the theories as abstract entities, in the way that functional descriptions attach to the abstract entities, sets of ordered pairs. Such a space could conceivably be so dense that the slightest change in the description of a term in a theory may create a different referent. That circumstance would be compatible with Feyerabend's claim that a change in background assumptions generates a change in meaning.

In sum, then, whether one arrives at a conclusion of stability or instability of meaning derives from ones prior assumptions concerning issues like realism or constructivism. This means that the assertion by either group of stability or instability of meaning cannot be used to mount further arguments for realism or constructivism, on pain of circularity. Finally, if instability of meaning is sufficient to establish the incommensurability of theories, then this pathway to incommensurability is open only for constructivists. Realists need not be troubled by it.

8. Twin Earth: An Illustration of Sparseness

As closing illustration, the notion of sparseness of the meaning space helps explain the bewilderment felt by many including me when we are shown Putnam's celebrated "Twin Earth" thought experiment.[6] We are to image a twin of our earth on which everything appears just as on our earth. There is a substance that has all the appearance of water in ordinary circumstances, but it is not H_20. It is, Putnam tells us, "a different liquid whose formula is very long and complicated. I shall abbreviate this chemical formula simply as *XYZ*." and that on Twin Earth the word "water" means XYZ. For anyone with even a meager background in chemistry, it is unimaginable that there could be such a substance. The quantum theory underpinning chemistry only admits a small roster of elements and the readily determinable physical, chemical, thermal and electrical properties of water are so extensive as to admit no other chemical combination than H_20. To state this another way, the meaning space for liquids invoked by the thought experiment is sparse and the only referents to which the liquid of Earth and Twin Earth could attach is H_20. The thought experiment collapses into a fantasy so remote from reality that we should despair of using it to discern how reference works in the real world.

---

[6] I thank Jim Woodward for suggesting this application. See Hilary Putnam, "The meaning of 'meaning" in *Philosophical Papers, Vol. 2: Mind, Language and Reality*. Cambridge University Press, 1985. Ch. 12.