August 25, September 20, 21, 2010; March 8, 2011
Substantial revisions, June 11, 2017

Dense and Sparse Meaning Spaces:
When Referential Stability Fails and Succeeds

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

The instability of meaning under theory change, construed as instability of reference, requires a dense meaning space. That is, the space of referents is densely populated so that slight change in theory can shift the referent of a term to another nearby in the space. If, however, the meaning space is sparse, there are no suitable referents nearby in the space. The referents of terms can remain unchanged, even with substantial changes in the theory, simply because there are no better referents to which to attach the terms. Dense and sparse meaning spaces are associated with antirealist and realist inclinations, respectively.

1. Introduction

The thesis of the incommensurability of theories was introduced into our modern tradition by Thomas Kuhn (1970, Ch. XII) and Paul Feyerabend (1962). The thesis is rich and complex. It enthralls some and enrages others; and it had engendered an expansive and enduring debate. Any comprehensive discussion of it lies well beyond this note. See Oberheim and

[1] This paper developed from a commentary on Norsen (this volume). His study can be consulted with profit for further details of the case of temperature in the context of the instability of meaning. My thanks for helpful discussion to Travis Norsen, Nora Boyd, P. D. Magnus and participants in the Workshop on Concepts, Induction, and the Growth of Knowledge, Pittsburgh, September 17, 2010.

Hoyningen-Huene (2016) for a convenient overview that includes reports of how these authors' views on incommensuarability evolved after these initial expositions.[2]

My concern in this note is limited to one component of the thesis. It is the idea that theoretical terms like "temperature," "mass" and "planet" do not retain their meanings when theories invoking them change. Thus the two theories are using the same words to talk about different things. This difference serves, for Feyerabend, to separate the content of an old theory from that of its new replacement, so that reduction relations fail to obtain between the theories. For Kuhn, this difference makes communication between proponents of different paradigms impossible.

Even after restriction to this narrower aspect of the incommensurability thesis, the pertinent literature remains expansive and bewildering. Sankey (2008, Ch.4) has provided an illuminating narrative of efforts to blunt the instability by calling upon philosophical theories of reference. In it, Scheffler (1967, Ch.3) used the Fregean distinction between sense and reference to constrain the instability: while the sense of a term may vary from theory to theory, its referent may be more stable. The non-descriptivist, causal theory of reference, as exemplified in Putnam's (1973) "twin earth" thought experiment, seeks to defeat the instability in another way. According to it, the reference of terms is fixed rigidly by a baptism so they are immune to changes of theory. Sankey finds neither approach satisfactory and recounts his own approach, which employs a synthesis of descriptivist and causal theories of reference.

My goal in this paper is provide a different way to understand how notions of reference interact with the claims of referential stability and instability. It is independent of these standard accounts of reference and can be developed with both or none. Rather it draws attention to the space formed by possible referents. While these spaces will sometimes be called "meaning spaces," the word "meaning" in this context is narrowly restricted to the referents, that is, the entities designated by terms in the theory. Broader notions of connotation and sense are excluded from the space. To accommodate both realist and antirealist sentiments, we have to be liberal in the entities we allow in this space. They may be actual physical states; or possible physical

[2] Similarly, the work of Kuhn and Feyerabend supports an almost insurmountable secondary literature. For more, see Bird (2000) and Preston et al. (2000).

states; or abstract entities like sets; or even fictions whose properties are stipulated to conform with assertions in some theory.

The principal claim is that judgments of instability or stability of reference under theory change will depend on the structure one attributes to this space. One might judge the space to be dense, as are the rational numbers on the number line. Then one will be drawn to the instability of reference, since slight changes of theory may be all that is needed to displace the referent of a term to a neighboring referent. Or one might judge the space to be sparse, as are counting numbers distributed over the number line. Then one will be drawn to the stability of reference, since slight or even more significant changes of theory may not be sufficient to dislodge the referent of a term and move it to a well-separated neighbor.

The next section will collect canonical texts by Kuhn and Feyerabend that define the instability thesis. Section 3 will introduce the notions of dense and sparse spaces. In Section 4, I will offer the principal argument of the paper: that instability of reference is associated with dense meaning spaces and stability with sparse spaces. While one might associate instability and stability of reference with descriptivist and causal theories of meaning respectively, I will argue in Section 5 that either theory of meaning can be associated with instability or stability. Section 6 will trace an association between sparse meaning spaces and inclinations to scientific realism; and between dense spaces and inclinations to antirealism. Section 7 will show how the notions of dense and sparse meaning spaces can defuse a puzzle with Putnam's twin earth example. Section 8 offers brief conclusions.

2. Instability of Reference

2.1 Kuhn

The incommensurability thesis plays a central role in Thomas Kuhn's (1970) *Structure of Scientific Revolutions*, a work that has been described euphorically as "one of the most cited academic books of all time." (Bird, 2013) A clear statement of the associated instability of reference for dynamical terms like "mass" appears within the discussion of the recovery of statements $N_1, N_2, \ldots, N_m$ of a Newtonian form from the relativistic statement $E_1, E_2, \ldots, E_m$ by restricting them to velocities slow in relation to the speed of light. Kuhn continues (1970, pp. 101-102):

> Apparently Newtonian dynamics has been derived from Einsteinian, subject to a few limiting conditions.
>
> Yet the derivation is spurious, at least to this point. Though the N_1's are a special case of the laws of relativistic mechanics, they are not Newton's Laws. Or at least they are not unless those laws are reinterpreted in a way that would have been impossible until after Einstein's work. The variables and parameters that in the Einsteinian E_1's represented spatial position, time, mass, etc., still occur in the N_1's; and they there still represent Einsteinian space, time, and mass. But the physical referents of these Einsteinian concepts are by no means identical with those of the Newtonian concepts that bear the same name. (Newtonian mass is conserved; Einsteinian is convertible with energy. Only at low relative velocities may the two be measured in the same way, and even then they must not be conceived to be the same.) Unless we change the definitions of the variables in the N_1's, the statements we have derived are not Newtonian.

A more mundane example, provided later in the text, pertains to the notion of a moving earth (1970, pp. 149-50)[3]

> Consider, for another example, the men who called Copernicus mad because he proclaimed that the earth moved. They were not either just wrong or quite wrong. Part of what they meant by 'earth' was fixed position. Their earth, at least, could not be moved. Correspondingly, Copernicus' innovation was not simply to move the earth. Rather, it was a whole new way of regarding the problems of physics and astronomy, one that necessarily changed the meaning of both 'earth' and 'motion.'[4] Without those changes the concept of a moving earth was mad. On the other hand,

---

[3] The two foonote references "4" and "5" are: "4. T. S. Kuhn, *The Copernican Revolution* (Cambridge, Mass., 1957), chaps. iii, iv, and vii. The extent to which heliocentrism was more than a strictly astronomical issue is a major theme of the entire book." and "5. Max Jammer, *Concepts of Space* (Cambridge, Mass., 1954), pp. 118-24."

> once they had been made and understood, both Descartes and Huyghens could realize that the earth's motion was a question with no content for science.[5]

This passage reflects more clearly what seems to have been Kuhn's principal concern: the instability of meaning precludes direct communication by proponents of different paradigms. In his 1970 *Postscript* to the 1962 *Structure*, he summarizes the point as (1970, p. 202):[4]

> Briefly put, what the participants in a communication breakdown can do is recognize each other as members of different language communities and then become translators.[17]

Earlier in the *Postscript*, Kuhn had given examples of words involved in these communication breakdowns (1970, p. 201)

> Because the words about which difficulties cluster have been learned in part from direct application to exemplars, the participants in a communication breakdown cannot say, "I use the word 'element' (or 'mixture,' or 'planet,' or 'unconstrained motion') in ways determined by the following criteria." They cannot, that is, resort to a neutral language which both use in the same way and which is adequate to the statement of both their theories or even of both those theories' empirical consequences.

2.2 Feyerabend

For Feyerabend, the instability of meaning served to undermine the notion of theory reduction. His extended example of instability of meaning (1962, pp. 76-81) arises in the context of his critique of Nagel's account of theory reduction. The example pertains to Nagel's claim that the statistical physics ("kinetic theory") of a molecular ideal gas stands in a reduction relation with ordinary phenomenological Boyle-Charles law. The difficulty is that a state of thermal equilibrium in phenomenological thermodynamics is static. The corresponding state in molecular theory is dynamical, with fluctuations allowing heat to move spontaneously from cold to hot. While such spontaneous motions are exceedingly improbable at macroscopic scales, they do

[4] Kuhn's lengthy footnote 17 refers to and critiques Quine's work on translation.

represent a violation of the second law of thermodynamics. The law is demoted to a result of very high probability at non-molecular scales. Of this, Feyerabend (1962, p. 78)[5] [6]writes

> Now whatever procedure is adopted, the kinetic theory does not give us such a concept [of a strict, non-statistical second law]. First of all, there does not exist any dynamical concept that possesses the required property.[95] The statistical account, on the other hand, allows for fluctuations of heat back and forth between two levels of temperature and, therefore, again contradicts one of the laws implicit in the "established usage" of the thermodynamic temperature. The relation between the thermodynamic concept of temperature and what can be defined in the kinetic theory, therefore, can be seen to conform to the pattern that has been described at the beginning of the present section: we are again dealing with two incommensurable concepts. The same applies to the relation between the purely thermodynamic entropy and its statistical counterpart; whereas the latter admits of very general application, the former can be measured by infinitely slow reversible processes only.

Feyerabend (pp. 78-79) then quotes Nagel at length, ending with a remark by Nagel similar to one of Kuhn's quoted above.

> Unless this hypothesis [of the sameness of the two temperatures] is adopted, it is not the Boyle-Charles law which can be derived from the assumptions of the kinetic theory of gases. What is derivable without the hypothesis is a sentence similar in syntactical structure to the standard formulation of the law, but possessing a sense that is unmistakably different from what the law asserts

Feyerabend "admit[s] the correctness of the last assertion" and proceeds to the more general claim:

---

[5] Feyerabend's footnote 95 reads: "I shall not discuss, in the present paper, the somewhat different situation with respect to the first law."

[6] Feyerabend (1983, p. 175) makes the same point more succinctly elsewhere:
"As is well known, the Brownian particle seen from a microscopic point of view is a perpetual motion machine of the second kind, and its existence refutes the phenomenological second law."

> … it has been my contention all through this paper that extension of knowledge leads to a decisive modification of the previous theories both as regards the quantitative assertions made and as regards the meanings of the main descriptive terms used. Applying this to the present case I shall therefore at once admit that incorporation into the context of the statistical theory is bound to change the meanings of the main descriptive terms of the phenomenological theory.

Feyerabend then turns to the example of classical and relativistic physics. The mass of classical physics, he notes, is an intrinsic property, whereas the mass of relativity theory is relational. (I understand that the latter means that it is dependent on the state of motion of the frame of reference.) He then concludes (pp. 80-81):

> It is also impossible to define the exact classical concepts in relativistic terms or to relate them with the help of an empirical generalization. Any such procedure would imply the false assertion that the velocity of light is infinitely large. It is therefore again necessary to abandon completely the classical conceptual scheme once the theory of relativity has been introduced;…

The idea central to this last claim is then rendered general with the remark (p. 81):

> After all, the demand for meaning invariance implies the demand that the laws of later theories be compatible with the principles of the context of which the earlier theories are part,…

It is then summarized in a general argument (p, 82):

> Our argument against meaning invariance is simple and clear. It proceeds from the fact that usually some of the principles involved in the determination of the meanings of older theories or points of view are inconsistent with the new, and better, theories. It points out that it is natural to resolve this contradiction by eliminating the troublesome and unsatisfactory older principles and to replace them by principles, or theorems, of the new and better theory. And it concludes by showing that such a procedure will also lead to the elimination of the old meanings and thereby to the violation of meaning invariance.

In a paper written shortly afterwards, Feyerabend (1965, p. 268) gives a more succinct statement:

… a diagnosis of stability of meaning involves two elements. First, reference is made to rules according to which objects or events are collected into classes. We may say that such rules determine concepts or kinds of objects. Secondly, it is found that the changes brought about by a new point of view occur within the extension of these classes and, therefore, leave the concepts unchanged. Conversely, we shall diagnose a change of meaning either if a new theory entails that all concepts of the preceding theory have extension zero or if it introduces rules which cannot be interpreted as attributing, specific properties to objects within already existing classes, but which change the system of classes itself.

3. Dense and Sparse Spaces

The texts quoted in the last section are part of the classic, originating canon that defines the instability of meaning in the context of the incommensurability thesis. The notion has polarized the philosophical community. One side finds it a liberation from the confines of a reductionist picture of science in which, as theories change, fact accumulates upon fact like layers of sediment. The other side finds the claimed instability a flimsy and implausible excuse for what is really a groundless and even haughty skepticism about claimed successes of science. While here I cannot plumb the psyches of these combatants to discern what motivates them, I can observe that success or failure of the claimed instability is closely tied with a presumption about reference that has not, so far as I know, been made explicit.

The referents of terms in a science form a space, whose entities can vary from actual or possible physical states to fictional constructs. Here I will identify two possible structures for that space and, in a later section, show how the two structures coordinate with two views, for and against, the instability of reference.

3.1 Dense spaces

In a dense space, each entity has others arbitrarily close to it. The clearest and most familiar examples are mathematical. Rational numbers are distributed densely on the number line. What results is extreme sensitivity to details in the means used to pick out such numbers. Consider approximate formulae for π. The most familiar is 22/7=3.142857…. Improved approximations are provided by other formulae such as 355/113= 3.1415929… or even just the truncated decimal 3.14159265. Since π is irrational, no rational approximation such as these is without error. The small error in each formula leads each formula to designate a different number.

A richer example is 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 descriptions contradict each other. A function on numbers is simply a set of ordered pairs of numbers. Two functions are the same if they comprise the same sets. These two descriptions pick out different sets. That difference stems directly from the fact that the two descriptions ascribe different properties to the function. Setting aside familiar philosopher's tricks,[7] this will generally be the case; if the descriptions of functions differ, then they do not pick out to the same function.

The linear and sine functions share an important property. If we consider just values of x very close to zero, then the two functions become almost the same. However, aside from their agreement at the origin, being almost the same is not identity, No matter how close the value of $x>0$ comes to 0, the two functions remain distinct.

3.2 Sparse Spaces

In a sparse space, each entity is separated by some space from its nearest neighbors. The counting numbers on the real line are an example. This sparseness can make it easier for us identify different counting numbers. Imagine form example that we seek to determine the number of coins in a bag by weighing the coins in one group. Assume that each coin weighs close to 5g with some small error. If their combined weight is 51g or 49g or 52g, we immediately

---

[7] 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 D_1 and D_2 are made contradictory by taking any contingent proposition X and forming new definitions $D_1\&X$ and $D_2\&(\text{not-}X)$.

correct for the deviation from 50g = 10x5g and infer than there are 10 coins in the bag. Fractional coins—10.2 or 9.8 coins—are not in the sparse space. There is nothing better than 10 coins in the space to associate with these weights. That is, our designation of coin number through weighings is not confounded by small errors. A sufficiently large shift in the weight will eventually change matters. Were we to weigh 55g or 56g, we should shift the number to 11 coins.

A more striking example is provided by things designated by 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 still pick out the same thing. 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;

"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. There is no other city in the neighborhood that can do justice to the bulk of each religious groups descriptions.

3.3 Error Tolerance

The essential difference between the two types of spaces lies in how they accommodate slight differences such as errors in the means used to designate entities in each space. A sparse space is tolerant of such errors, whereas a dense space is not. For example:

- 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.

This differing tolerance translates directly into instability or stability of reference when the spaces are meaning spaces.

4. The Instability of the Instability of Reference

If we assume that the space of meanings is sparse, then the canonical examples of the instability of reference fail. This failure results directly from the error tolerance of the sparse space.

Take Kuhn's example above of the instability of the term "mass" as we move from Newtonian to a relativistic theory. There is by supposition something in the space of meanings to which the term "mass" in each theory seeks to refer. It is loosely speaking that property of bodies that measures their resistance to acceleration. Comparison of the two theories is only viable in domains where both might apply, that is, when we consider bodies moving at speeds much less than that of light. In low speed domains, both Newtonian and relativistic theories attribute the same mass numerically, for example, to the moon; and that mass enters into formally identical laws, up to tiny errors that are scarcely measurable. This agreement is enough for both theories to be designating the same referent, that is, the actual physical mass property of our moon. That they differ in some properties attributed to the mass, such as conservation or convertibility to energy, is insufficient to confound the reference.

Feyerabend's example of temperature fares similarly poorly in a sparse meaning space. Once again we can only compare phenomenological thermodynamics and statistical physics in domains in which they both apply, that is, the domain of macroscopic bodies. Temperature is the single parameter that determines whether two systems can enter without change into thermal equilibrium. Their thermal states will remain unchanged when the systems are allowed to interact if and only if their temperatures are the same.

In phenomenological thermodynamics, the temperature T of a body is a property of state; and the energy of a body in equilibrium at temperature T will remain fixed (unless of course of the state properties are changing). In statistical physics, a system in thermal equilibrium at temperature T will fluctuate between different energies E according to the probability distribution[8]

$$p(\mathrm{E}) \text{ is proportional to } \exp(-E/kT)$$

where k is Boltzmann's constant. Two thermal systems with energies distributed according to this distribution can enter into equilibrium with each other just in case the value of the parameter T is the same.

Both theories will assign the same temperatures and other related formal properties to macroscopic bodies. They will differ in some properties. The probability distribution of statistical physics will allow the energy of a body at a fixed temperature to fluctuate. However the magnitude of the fluctuations on ordinary time scales will be so slight for macroscopic bodies as be indiscernible. The mean energy around which these tiny fluctuations arise will coincide with the energy attributed to the body by phenomenological thermodynamics.

In a sparse meaning space, both theories designate the same referent with the term "temperature." There is nothing else in the neighborhood to which they might refer. The neglect by phenomenological thermodynamics of the indiscernible fluctuations in energy are slight errors insufficient to disrupt the agreement.

All this changes if we adopt the assumption that the meaning spaces are dense. For now the essential toleration of error is lost and the meanings of terms becomes unstable under theory change.

For example, the slight difference in properties attributed to Newtonian and relativistic mass in the slow speed domain is sufficient to shift the referent of the term "mass" so something nearby but different. The same shift arises from the slight differences in properties attributed to temperature and thermal states of macroscopic systems by phenomenological thermodynamics and statistical physics. We can only prevent this shift if the two theories agree whenever they

---

[8] This definition of the temperature of a system in statistical physics is very general and holds in both classical and quantum statistical physics. It is far preferable to the association of temperature with the mean kinetic energy, as in unfortunately quite standard in the reduction and incommensurability literature. That association holds only for special cases. It cannot be applied, for example, to the temperature of heat radiation.

attribute properties to the same systems, since a slight disagreement is sufficient to change the referent. Hence we saw Feyerabend above insisting that invariance of meaning "implies the demand that the laws of later theories be compatible with the principles of the context of which the earlier theories are part."

Feyerabend (1965, p. 267) allows that some slight changes of theory do not alter meaning. His example is that slight changes in gravitational potential are insufficient to change the meaning of terms in the theory that hosts the potentials. Similarily "the existence of rubber bands of different strength does not indicate that there are various concepts of 'rubber band.'" While these are reasonable expectations, it is not clear how Feyerabend can maintain them while also insisting that meanings derive from the theory in which the terms appear. If the latter insistence is true, then Feyerabend's mere declaration to the contrary is insufficient to preclude the possibility that even small changes in the theory might well issue in small changes in meaning.

5. Independence from Theories of Reference

One might imagine that the association of stability and instability of meaning with sparse and dense meaning spaces will require in turn a particular connection in each case with specific theories of meaning. That is, one might expect that instability of reference will be associated with descriptivist theories, where the referent of a term is determined by its description. For slight changes in the description would yield corresponding changes in the referent. Similarly, the causal theory of reference attaches meaning to terms by rigid designation and thus naturally leads to stability of reference.

These natural expectations do not survive scrutiny. In a sparse meaning space, a descriptivist theory of meaning can sustain stability of reference. The error tolerance characteristic of sparse meaning spaces protects the stability of reference from confounding by small changes in description.

Conversely, a causal theory of reference may have serious trouble maintaining stability of reference if the meaning space is dense. The difficulty is most apparent if we try to apply the causal theory to general theoretical terms. For then the baptism that rigidly designates the meaning of the term cannot be effected by simple ostention. This is the familiar "qua" problem, described in Sankey (2008, p. 67). To say "cat" and point at a furry object leaves open whether the baptism of the general term "cat" is to all living things, to furry things, to four-legged things, and so on. A descriptivist supplement is needed to fix the reference. The problems of a descriptivist theory of reference in a dense meaning space reappear in the supplemented causal theory.

The problem persists in physical theories. 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.[9] 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. Once again, a descriptive supplement is needed.

Even baptism of proper names can be troublesome in dense meaning spaces. Unambiguous baptism may only be possible if the meaning space is sparse. We can baptize "the old city of Jerusalem" with the wave of a hand and a loud proclamation since the intended referent is not likely to be confused. The old city is demarcated by a prominent and impressive city wall. It is in a sparse meaning space. But it would be a great deal harder if we were dividing up undifferentiated territory into states, for then the meaning space is dense. Just how are we to fix the referent of "Freedonia"? We might do it with lines on a map; or instructions on how to place border markers by proceeding in particular direction from some designated starting point; and so on. These artifices introduce descriptive intermediates whose referents may be difficult to discern uniquely in a dense meaning space. Is a wobble in a line on a map carving off territory to Freedonia? Or is it a drafting error?

6. Realism and Antirealism

How do we decide whether the meaning space is dense or sparse? There is no simple recipe. Rather, whether one finds the meaning space of theories to be dense or sparse will depend on one's position toward realism and antirealism concerning scientific theories.

---

[9] See Norsen (this volume) for a review of the long history of efforts needed to make the notion of temperature definite.

6.1 Realism

The situation is simplest if one harbors realist inclinations in philosophy of science. The sparseness of the meaning space follows almost automatically. Realists see in the history of science the evolution of a sequence of theories that progressively come closer to the true realities. Realists do not need to believe that the evolution has reached its final stage. They may believe us to be still remote from it. However an essential characteristic of the theories in the sequence is already apparent: if each is read literally, it portrays a sparse meaning space. In both phenomenological thermodynamics and statistical physics, there is no magnitude adjacent to temperature but distinct from it that could serve equally well as temperature. So, realists should conclude, in so far as our best theories are guides to reality, that reality provides a sparse meaning space.

Correspondingly, some sort of antirealism is required if one is to judge that space dense. One need not be a skeptic about the existence of the external world. One merely needs to regard the referents of terms in scientific theories as somewhat removed from direct attachment to the external world. This world, in large measure or even entirely, remains beyond the reach of our theories. In such a view, the portrait of things provided by the theory is more a creation or construction of the theory. As a result, slight changes in theory will produce corresponding changes in the constructions. The resulting space of meanings will be dense, for we can create new entities in it merely by making suitable adjustments to our theories.

6.2 Kuhnian Antirealism

While we may never quite be able to discern Kuhn's positive view, he was definite enough in his remarks for us to see that he held some form of constructive antirealism. His 1970 *Postscript* to the 192 *Structure* includes these remarks (1970, p. 206):

> One often hears that successive theories grow ever closer to, or approximate more and more closely to, the truth. Apparently generalizations like that refer not to the puzzle-solutions and the concrete predictions derived from a theory but rather to its ontology, to the match, that is, between the entities with which the theory populates nature and what is "really there."
>
> Perhaps there is some other way of salvaging the notion of 'truth' for application to whole theories, but this one will not do. There is, I think, no theory-

> independent way to reconstruct phrases like 'really there'; the notion of a match between the ontology of a theory and its "real" counterpart in nature now seems to me illusive in principle.

This antirealism may not seem so far fetched if it applied to more abstruse theoretical terms like "entropy" or "quark." However, if we assume that it extends even to quite prosaic terms in science, then we can make sense of what otherwise seems an odd and even bizarre pronouncement.

We saw Kuhn's remark above that Copernicus changed the meaning of the term "earth" by affirming that the earth moved; and the result was a breakdown of communication between geocentrists and heliocentrists. To realists who presume that the term "earth" resides in a sparse meaning space, the term "earth" can only refer to the thing on which we stand. Successful reference can tolerate the error committed by at least one of the geocentrists or heliocentrists in their differing attributions of states of motion to the earth. That Kuhn would think otherwise is readily explicable if the "really there" of his *Postscript* remarks above include the earth. For then the earth becomes a construct within various astronomical theories and is to some lesser or greater extent removed from the true realities of nature. Then attributing different properties to earth changes the construct. Since they can be arbitrarily small changes, the referent of the term earth lies within a dense meaning space. It follows that geocentrists and heliocentrists are talking about different things.

6.3 Feyerabendian Antirealism

Feyerabend's antirealism is expressed more forthrightly, but is correspondingly more difficult to elaborate. Feyerabend puts it this way in a 1965 journal article. He considers the distance "*AB*" between two simultaneous events in space, according to Euclidean geometry (theory *T*) and general relativity (theory *T*'). He writes (1965, pp. 270-71; Feyerabend's emphasis):

> In traditional philosophical terminology: $(AB)_T$ and $(AB)_{T'}$ are *constituted* by the basic principles of *T* and *T*', *respectively*. These entities cannot be described, not even in part, by means that are independent of either theory at the time of the advent of *T*'. In earlier papers I have expressed the fact by saying that "$(AB)_T$" and "$(AB)_{T'}$" are *incommensurable notions*.

Since the term “constituted” is widely used in philosophy with rather ordinary meaning, we can only speculate on the import of the mention of “traditional philosophical terminology.” Presumably it refers to the Kantian distinction between constitutive and regulative principles. It is not easy to extract a clear meaning for “constitutive” without developing the fuller Kantian framework, whose details remain opaque to me. However Everett’s (2014, p.4) gloss seems to convey sufficient for our purposes:

> In Kant’s philosophy constitutive principles governed the application of the faculty of understanding to the manifold of intuition. In effect they are rules for the construction of the phenomenal world.

The best reading I can then make of Feyerabend’s claim is this. The two terms $(AB)_T$ and $(AB)_{T'}$ do not refer to some distance in space that is independent of the two theories. What they designate is determined—“constituted”—by the relevant theory. So, if the two theories differ, then the referents of the terms differ.

Writing several decades later, Feyerabend is clearer about the dependence of meaning on theory. He writes (1983, p. 180):[10]

> After all, the meaning of every term we use depends upon the theoretical context in which it occurs. Words do not “mean” something in isolation; they obtain their meanings by being part of a theoretical system.[137]

One might imagine this dependence to apply to theoretical terms only. Feyerabend (1983, pp. 181-82) proceeds to assure us otherwise:[11]

> What we have said applies, of course, also to singular statements of observation. Statements that are empirically adequate and are the result of observation (such as “here is a table”) may have to be reinterpreted, not because it has been found that they do not adequately express what is seen, heard, felt, but because of changes in sometimes very remote parts of the conceptual scheme to which they belong.[140]

Similar sentiments appear later in the paper (p. 213):

---

[10] Feyerabend’s footnote credits Wittgenstein as “one of the most ardent defenders of this principle.”

[11] Feyerabend’s footnote elaborates, beginning with “Direct observation, therefore, cannot at all teach us ‘what we see.’”

> According to the point of view I am advocating, the meaning of observation sentences is determined by the theories with which they are connected. Theories are meaningful independent of observations; observational statements are not meaningful unless they have been connected with theories.

We arrive once again at meaning spaces determined not by the world but largely or entirely by our theories. Since entities in such a space are created at our wish by even small changes in theory, the resulting space is dense.

7. Twin Earth Again

Here is a closing illustration of the use of the notions of denseness and sparseness of the meaning space. They will help explain the bewilderment felt by many, including me, when we read Putnam's (1973) celebrated "Twin Earth" thought experiment.[12] In it, 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 chemical formula is very long and complicated. I shall abbreviate this chemical formula simply as *XYZ*." On Twin Earth the word "water" is used to designate the substance XYZ. Putnam proceeds to argue that the terms "water" on earth and on Twin Earth have different meanings.

The thought experiment is much celebrated in the literature on reference. However, for anyone with even a meager background in chemistry, it is unimaginable that there could be such a substance XYZ. 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 chemical combination other than H_20.

How is it possible that the one thought experiment can receive such divergent receptions? The answer lies in the structure of the meaning spaces assumed. Putnam is tacitly supposing a dense meaning space for the terms in his twin earth narrative. In it, a neighbor to ordinary earth's H_20 is another entity close in all properties except for the small detail of its chemical composition, XYZ. Those who find the thought experiment baffling employ a meaning space

---

[12] I thank Jim Woodward for suggesting this application.

provided by modern chemistry. That space is sparse and there can be no such near neighbor to H_20 in it. People in this latter group, if they are forced to proceed nonetheless, have no meaning space in which to understand the terms. The thought experiment has degenerated to something like a fantasy story about unicorns or goblins. The story's author can stipulate any property, mundane, sublime or absurd, for its characters. The resulting meaning space is so remote from reality that there is no prospect of using it to discern how reference works in the real world.

8. Conclusion

What determines whether one favors the instability or stability of reference under theory change, I have argued, is whether one holds that the space of meanings is dense or sparse, respectively. This last commitment, I have further argued, depends on whether one is inclined towards realism or antirealism concerning science. A sparse meaning space is associated with realist inclinations and stability of reference. A dense meaning space is associated with antirealist inclinations and an instability of reference. Since one's position on realism and antirealism is antecedent to judgments of stability or instability of reference, it follows that claims about the instability of reference will be convincing to antirealists and unconvincing to realists. As long as the two sides differ in these fundamental presumptions, we can expect no resolution of the debate over the instability of reference.

References

Bird, Alexander (2013) "Thomas Kuhn," *The Stanford Encyclopedia of Philosophy* (Fall 2013 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/fall2013/entries/thomas-kuhn/.

Bird, Alexander (2000) *Thomas Kuhn*. Chesham: Acumen.

Everett, Jonathan J. (2014) *Constitutive or regulative principles? The Kantian legacy for contemporary philosophy of science*. Doctoral thesis, University College London.

Feyerabend, Paul K. (1962) "Explanation, Reduction and Empiricism," pp. 28–97 in H. Feigl and G. Maxwell (ed.), *Scientific Explanation, Space, and Time*. (Minnesota Studies in the Philosophy of Science, Volume III), Minneapolis: University of Minneapolis Press.

Feyerabend, Paul K. (1965) “On the ‘Meaning’ of Terms,” *Journal of Philosophy*, **672**, pp. 266-74.

Feyerabend, Paul K. (1983) “Problems of Empiricism,” pp. 145-260 in R. G. Colodny, ed., *Beyond the Edge of Certainty: Essays in Contemporary Science and Philosophy*. Lanthan: University Press of America.

Kuhn, Thomas S. (1970) *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press.

Norsen, Travis (this volume) “Scientific Cumulativity and Conceptual Change: The Case of Temperature.”

Oberheim, Eric and Hoyningen-Huene, Paul, (2016) “The Incommensurability of Scientific Theories,” *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/win2016/entries/incommensurability/.

Preston, John, Munévar, Gonzalo, and Lamb, David, eds (2000). *The Worst Enemy of Science? Essays in Memory of Paul Feyerabend*. Oxford: Oxford University Press.

Putnam, Hilary (1973) “Meaning and Reference,” *Journal of Philosophy*, **70**, pp. 699-711.

Sankey, Howard (2008) *Scientific Realism and the Rationality of Science*. Aldershot: Ashgate.

Scheffler, Israel (1967) *Science and Subjectivity*. Indianapolis: Bobbs-Merrill.