The Logic of Dynamic Positivism: On Making Knowledge Public

Darko Sarenac

January 15, 2008

1 Historical roots

Research put forth in this paper is mainly indebted to two fairly recent philosophical trends: extreme linguistic externalism and objectivism on the one hand, and dynamic process oriented epistemology on the other. We will argue that knowledge properly understood is a stable but evolving external, recalcitrant structure. Further, this objective structure is dynamically evolving and understanding the laws that govern this intricate dynamical process is, in our view, the most important task of epistemology.

While the roots of our view can be found in Wittgenstein's arguments against the possibility of a private language[1], the oldest explicit predecessor of the sort of epistemic objectivity that we espouse can be found in Karl Popper's seminal work *Objective Knowledge: An Evolutionary Approach*. There, Popper argues that knowledge properly conceived lies in what he terms the *third realm*. This realm is closer to that of mathematical objects and theories that that of human beliefs and is in an important sense objective. Our view differs in that the knowledge although essentially public and objective will not require postulating a separate realm of abstract objects. Logical structures we postulate will be structural properties of languages employed in communication.

The next important class of epistemic predecessors of our theory are theories that dispose the classical requirement of belief as a necessary condition to knowledge. There are several varieties here, but a good exemplar is Fred Dretske's "Precis of Knowledge and the Flow of Information"[2]. In this variety of epistemology, beliefs, and

[1] Wittgenstein Ludwig. *Philosophical Investigations*. Blackwell Publishing, 1953.

[2] In: Hilary Kornblith, ed., Naturalizing Epistemology. Cambridge: MIT Press.

the difficulties that that they introduce into the classical epistemology, are replaced with some subtler notions. Following this epistemic trend, we will try dispense with belief but also with all of its mentalistic variants all together. The final public component that we draw on are the extreme externalist views about language first introduced by Daniel Dennett in *Consciousness Explained*, and further elaborated by Andy Clark and others. On such a view knowledge/language and its evolution literally upgrade our biological hardware and its ability to interact successfully with our environment. Clark has a nifty name for this phenomenon: wideware. Wideware is our (external) language and our technological environment that makes who we are by literally reprograming us.

In a just slightly different strand, the dynamic aspects of our view are not too distant relatives of Hintikka's Socratic epistemology. The similarity lies in Hintikka's emphasis on question/answer games and the importance such games play in discovery and pursuit of knowledge. A related, but just as significant logical approach to knowledge is the *Dynamic Epistemic Logic* made popular by Alexandru Baltag, Larry Moss, Jelle Gerbrandy, Johan van Benthem, and others. Their emphasis on the role of dynamics in the development and the acquisition of knowledge was invaluable in the formation of our own view. Finally, we should emphasize the loose but important connection with mathematical structuralism. In some sense we simply extend the idea that what really makes and brakes the mathematical knowledge are structures and their properties to the rest of our epistemic pursuits. How we build these structures and how we interact with them is the main topic of this paper. Although our approach is heavily influenced by its historical predecessors, we hope that it combines the historical elements in a way sufficiently novel to be worthy of independent consideration.

2 Introduction

Most knowledge we acquire comes through the spoken or printed word, audio and video sources, maps, blueprints, and more recently websites, that is, things in the external world with mind-independent structure. Yet, when asked to locate knowledge, we are tempted to point to the contents of our heads. Here is a curious puzzle that we would like to understand. Chemistry and physics, and to some degree biology are the best examples of physical languages we have. Let us take chemistry as the example here. We could in principle learn chemistry by only ever interacting with books! In fact, it is not inconceivable that one would learn all one knows about chemistry from books and actually make substantial contributions to chemistry without ever interacting with the physical surroundings. How and why this is possible strikes us as one of the deepest epistemic puzzles. We have managed to 'translate' the physical properties into the terms of our scientific language to such degree that the actual interaction with the world can almost be dispensed with. We presumably as a species started with some sort heavily indexed, context dependent languages, and over time have evolved highly logicalized languages that are almost entirely independent from the context for their proper usage. This paper attempts to understand how this translation process from physical to logical takes place.

Our hope here is to undermine the idea that the essence of the our understanding of knowledge lies with *I know*. Instead, the crux of the understanding is in *we know*. We will argue that the former is, at best, derivative of the latter, and—contrary to the views of most epistemologists—not as interesting. The core of the paper is a novel, non-semantic view of communication. Communication on our view is a vehicle for evolving logical structures. It is these logical structures that are the real epistemic achievement.

Further and more importantly, if our view is correct, then knowledge can be liberated from belief. Given the number and significance of the difficulties and puzzle cases generated by the belief requirement in almost all epistemological theories, this represents a substantial step forward.

Though on our view knowledge itself can be found in places such as books, articles, the internet, etc., people are capable of temporarily implementing and processing small portions of it. According to our view, knowledge is an external, independent, and continually evolving logical structure. If we are correct, then the focus of epistemology should concentrate on understanding the properties and the evolution of this structure. On our view, knowledge progresses in the sense that the structure built becomes more detailed, predictive, better organized internally, and most importantly, more empirically adequate.

Since individuals implement portions of this structure, interest in the 'I know' component of knowledge still remains. However, it is through the process of communication that individuals add to the virtues of the logical structure, those virtues being empirical adequacy, consistency, and manageability of the structure. Probing the strength of the structure—through experiments and logical examination—is essential to improved sophistication of the logical network. Though there surely is measurable progress over time, its connotation is non-normative. Progress with regard to knowledge is merely the measurement of predictive power, simplicity (seen as computational intricacy of the parts of the logical structure involved), and the level of organization (how sophisticated the underlying mathematics are, and how successfully they sort terms of the science involved, as well as how simply it connects terms via laws).

It is essential to the success of this human epistemic enterprise that different individuals involved in the building process have differing capacities. The fact that different individuals most likely implement scientific (and other) terms differently enables them to bring novel insights to the table; what is essential is that they implement portions of the same logical structure. We basically are an important part of the environment of evolution of our epistemic network.[3] And the more varied the environment, the more sophisticated the network. Science cherishes diversity!

The talk of logical structures here is purely syntactic. As we will see in the example of the evolution of terms, the semantics—if there is any—is relegated to the details of the implementation. The particulars of implementation, as interesting as they may be from the point of view of our biological understanding our ourselves, are largely irrelevant both to communication (the building process itself) and the epistemic logical structure (the building outcome). What is important is that individuals have some way of partially implementing and processing parts of logical structures. Note here that we are talking about the dynamics of communication in a human natural language. There is no commitment whatsoever to the idea that this is the only way to communicate. There may be emphatic ways of drawing attention to the same processes across individuals (e.g., the mechanism of mirror neurons), chemical-fermonal interactions, or other neural mechanisms involved in the transfer of information. This other processes may even supplement every day linguistic communication, or even be essential to our ability to learn language. The claim that we make here is that in epistemic exchanges, making things explicit is essential. Thus linguistic communication takes the driver's seat in our epistemic enterprise. It is possible that we are capable of doing something that exceeds the limits of our language, but such an activity may be difficult to encode in the epistemic structure. To sum up, we are interested in explaining what we *mean* when we say *we know*. Our view is a kind of third-person empirical epistemology. How is it that this wonderful species of ours achieves all the scientific miracles that it does? The answer, if we are right, is in the evolution and communication of logical structures.

2.1 Methodology is from Venus

We propose an empirical third-person approach to communication. We want to see what kind of evidence would count as evidence of communication from a third-person perspective, and then try to see what kinds of laws and entities need to be postulated to explain such evidence. The topic of what counts as third person linguistic data is huge, and it deserves a much more thorough treatment then we can give it here. It is very difficult to see where the first person introspective information ends, and some

[3] Really, here we should be talking about epistemic structures, not a single structure. In future research we intend to pursue the issue of what makes such structures distinct.

detached third person data begins. Here we provide a first pass at what this kind of evidence may be and entail. If a scientist from Venus showed up on earth, she may notice that we communicate. That is, she might note that relatively low level "energy transfers" (i.e., noises from one human to another) are capable of effecting the behavior of a human or group of humans. She would notice, if she observes carefully, that there is a systematic relation between various noises and certain regularities in our world, although for the most part the connection is rather messy and at best statistical in nature (oftentimes having no connection whatsoever). She would also notice that these connections are negotiated in the process of usage and that some individual humans' views appear to "count more" with respect to certain terms,[4] that is, there seem to be individuals with expertise in certain areas of usage.

For instance, the Venutian scientist notices that mother is an expert user and her child recognizes this. The physics teacher is an expert user and the student recognizes this. Our scientist from Venus also notices that we humans need to learn and (over a long period of time) construct the very structures that we use to communicate. She may also notice that new humans implement a (tiny) portion of whatever structure is already there. They mirror whatever is known by their societies. Given these facts, a reasonable hypothesis, we argue, is that humans are engaging in a joint project of building an epistemic structure by communicating the building materials. These building materials succeed or fail at carrying a message in virtue of some of their properties. We postulate here (as an empirical hypothesis) that the properties that evolve and carry the communication are *logical properties* of the entities involved. Further, we claim that what is really interesting about the evolving language with respect to communication is the fact that language evolves more complex logical structures. Furthermore, we will argue, for the purposes of studying our epistemological success, these logical structures are best seen as independent of anyone's head, but exhibited in books, computer architectures, artifacts, etc. The heads are simply the fertile ground where the structures mutate, mate, and mature, but the final product—a mature logical structure—is a public entity, and is not counted as a part of our overall epistemic structure until it is made public. To sum up, we hypothesize that the evolution of language really carries with it an evolution of *logical structures.* These structures enable us in turn to evolve an epistemic (logical) structure. It is the building of this structure that distinguishes ours from most other earthly species.[5] That is, logic makes us special.

[4] As a first approximation at what might see as 'counting more' behavior for our Venutian scientist, perhaps something like this: the noises produced by some individual humans elicit greater behavioral change in greater numbers of humans than other individuals producing those very same noises. Also the usage of these 'experts' appears to be yardstick for the usage of other humans.

[5] It is a genuine empirical question whether we are the only earthly species able to communicate in

2.2 Briefly About Logical Structures

Our theory postulates logical structures as the basic theoretical entities. Our main claim is that these structures explain not only why humans with such diverse brains/minds are capable of communication, but also underlie the very possibility of human knowledge, everything from building tools and houses, to advanced branches of mathematics. In fact, one can explain our success as a species as the success of the interaction of the logical structures we built with the physical environment. Although we use the term *structure* to suggest that what is communicated are structural properties—properties learned via a game-theoretic equilibrium—the way we think of the structures is purely syntactic. They are simply symbols, or better still, noises, related (via introduction and elimination rules) to the physical world and other such structures. The most obvious example of large scale logical structures that we have built and that enable us to interact with our surroundings include mathematical theories such as geometry, arithmetic, analysis, statistics, computer programs, etc. The less obvious ones include books, articles, stories, medical databases, and just plain sentences of a natural language[6]. All of these items communicate to us and carry knowledge in virtue of the logical structure that they implement. The main reason we choose to focus on the logical properties of language, is that these are indeed the very properties that we know to be able of supporting communication. They have been used not only in computer protocols and telephone devices, but even radios and other sorts of technology that at the base does not seem to be *logical* depends on formal and logical relations and properties to encode the information to be transfered.

2.3 On Differences and Similarities

Differences in linguistic and other mental abilities abound. Some of us have larger vocabularies and process language more efficiently than others. Some of us are better at processing arithmetic, others process geometrical notions more readily. The idea that any human is just as capable as any other human of learning any aspect of language is extremely ill supported. In fact, there is little neurological reason to suppose that

this sense. Perhaps chemical interaction in ants do this very thing, build logical structures. As R.C. Jones pointed out to us, the observations that the Venutian scientist make regarding the role of 'noises' and the subsequent dynamics of logical structure of human communication apply to the humans-as-Venutians/ants-as-subject model as well. Scientists who study ants (or porpoises or chimpanzees or ravens) do exactly what we have our Venutian friend doing in the science-fiction example. This connection further emphasizes the difficulties that the Venutian scientist would have gathering data and deciding what to count as noise, etc. Our picture here is somewhat simplified to make our point. What we are really staying away from is counting introspective data in our theorizing about language.

[6] An interesting question is whether music is this kind of structure.

various concepts are implemented in the same or even similar fashion across brains. It seems to us rather unlikely that all such differences will be explainable in terms of environmental variables, and at the very least, at present, there is no reason beyond wishful thinking to suppose that they will. It seems unlikely that the ability to add or multiply natural numbers, for example, is implemented by the same algorithm across brains, or even that concepts such as *set* or *flower* will be instantiated in the same way across brains. The examples don't even need to be so technical. Here is a garden variety list of neurological differences pertinent to communication:

Some of us are:

–incapable of experiencing pain.

–able to tolerate pain better than others.

–able to perceive a wider range of colors than others while others are colorblind.

–able to (or claim to be able to) interact directly with mathematical structures in the way that the rest of us do with physical ones.

–highly susceptible to opiate addiction and capable of experiencing the related cravings.

–depressed or have other mental differences.

–female, male, or other of a number of gender-related variations.

–brain damaged.

–religious.

–more perceptive to the needs of others.

–angry and violent.

–"psychopathic". [7]

And the list continues. When the differences are drastic, we talk of mental disabilities. But in most cases, we would argue, the differences are present, significant, but not easily behaviorally noticed. What needs to be explained is how it is that we communicate in the face of such neuro-chemical and sociobiological differences, and when (and how) exactly it is that communication breaks down. How is it that a person who does not experience pain can still communicate about it? Or is it that they just *think* they do?

[A possible implementation task: to have two radically different program architectures, say a neural net and a classical AI program, learn color vocabulary from one another or by interacting. For further variation they could have differing input or perceptual strategies.]

[7] I strongly urge the reader to consider views of humans with neurological differences. A good accessible source is: http://www.wrongplanet.net/.

2.4 How public is language and its logic?

Here we argue for a radical take on this question: language is public all the way. Language uses us (that is, humans as a species) as a carrier, and we are blessed to be able to implement it. The relation is more like biological symbiosis. Language needs us for its evolution.[8] There is no reason why other species capable of supporting some versions of language could not have evolved or be designed, as portions of it have already been implemented in various computational devices. What language evolves, its "genotypes" so to speak, are species of logical structures. These then are independent of humans and come in different kinds. The simplest kind of a logical structure is a *term*. We now look at the way terms are introduced, learned, and the way they evolve in some detail.

2.5 An example of a term: 'red'[9]

It is often pondered in introductory philosophy classes whether one knows that one's neighbor sees red in the same way they do. Maybe they see red the way we see blue? Or green? We argue that as important as the answer to this question seems, for the purposes of communication it does not matter at all. What matters is that:

1. Your neighbor has the appropriate introduction rule for the term red, that is, *red* is only used in the appropriate circumstances.

2. Your neighbor has the appropriate elimination rules for *red*, that is, when someone else uses *red*, she reacts appropriately.

What this will require, in the abstract, is that the term 'red' is invoked by (roughly) the correct class of surfaces[10], and just as importantly that 'red' plays the correct logical role with respect to other terms. For example: it is surfaces that are red, if a patch is red, it is not green, red and soft are not comparable, etc.

Our move opens communication up to other species and kinds of entities. If they are capable of relating to the physical world in a logical way, that is, if they are capable of 'translating' physical patterns into logical structures, they can communicate. We

[8]Our language here should be remniscient of Dawkins's theory of *memes*. This is not accidental, though, our view is substantially different. We eventually hope to have a contribution to make to the theory of cultural evolution. It appears that we are a lot more interested in the technical details of the process.

[9]The main reason we use 'term' here instead of say 'concept' or a similar more common philosophical expression is that we wish to be liberated from the semantic baggage that such expressions carry. Ideally, we would want to 'term' to be purely syntactic and physical.

[10]The story is further complicated by the fact that the term 'red' could be invoked by a hallucination or by imagining the color red in one's mind. The idea below, as we'll see is that the introduction and elimination rules cover all such cases.

have chosen 'red' as the running example. But we could have easily chosen 'electron', 'large cardinal', or 'good'. We have chosen 'red' for its relative simplicity as well for the fact that it relates a logical structure to physical properties. There will be several kinds of terms:

1. Terms that pick out and 'logicalize'[11] some regularities in the natural world. Examples here are red, electron, species, mammal, sister, GDP, etc.

2. Purely logical terms. These kinds of terms have been 'logicalized' to such an extent that their usage is entirely context independent. They cannot be forced to be altered by a new discovery about the world. Such terms have only logical properties and do not directly pick out any regularity in the physical world. Examples include pure mathematics, metatheory of logic, other kinds of mathematics which try to systematize the existing formalisms such as set theory, category theory, lambda calculus, etc. Although the pure logical terms do not play a direct role in capturing the patterns in the physical world, they provide a language or a "toolkit" for logicalizing terms of kind (1) above. While terms of kind (1) logicalize patterns in the world, terms of kind (2) logicalize patterns of such logicalization. They involve a kind of second-order logicalization.

3. Terms that 'pretend' to pick out regularities in the natural world, but no amount of new information is likely to alter them. Examples here are most fictional entities. These kinds of terms play an important role in our understanding of the world as they exhibit the power that our languages have in freely reforming and recombining to capture much wider classes of patterns than those immediately available to our sensory apparatuses. Also, they emphasize the freedom that we have in modeling the non actual possibilities as a kind of combinatorial property.

It seems rather plausible that the species of terms of kinds (2) and (3) have evolved from terms of kind (1). It also seems plausible that science began when we started reflecting on terms of kind (1) and evolving terms of kind (2). The most famous example here is the evolution of geometry from the practice of physical measurements and land division.

The list here is not intended to be exhaustive. What all of these kinds of terms have in common is that to learn them one has to learn how to introduce them in communication as well as how to eliminate them, that is, 'understand' them (to use a folk-theoretic term here).

[11]We use the term 'logicalize' in a somewhat non-standard way. To logicalize a pattern is to find a vocabulary for expressing that pattern in the natural language.

3 Evolving Logical Structures: Simple Terms

Learning to communicate is a never-ending dynamic process. We use 'experts' throughout our lives not only to learn new terms, but also to improve our usage of terms already in use. We also as experts (if we are lucky) offer new terms to the public for communication, whether as scientist, mathematician, or pop-culture columnist. The *Oxford English Dictionary* requires three public uses for a term to qualify for entry into the *OED*. In the next section we try and capture some of the dynamics of this kind of process.

3.1 Introduction and Elimination *Rules*

All terms have introduction and elimination rules. Paradigm here is borrowed from the systems of natural deduction in proof theory. The idea is that to learn a new lexicographic item, all you need to understand are the circumstances under which this new item is introduced as well as the way in which it can be eliminated. The easiest and the simplest example from Natural Deduction are the introduction and elimination rules for AND.

Introduction Rule for AND

If you are in a situation where sentence 'A' holds, and sentence 'B' holds,

you can then introduce a new sentence 'A AND B'.

Elimination rule for AND

If you are in a situation where 'A AND B' holds then,

you can eliminate AND to obtain sentences A, B.[12]

Occasionally, as in the example above, introduction and elimination rules for a term can be stated explicitly. In fact, it is one of the goals of science to make such introduction and elimination rules (as) explicit (as possible). Yet, in most non-trivial natural language cases we don't have an insight into how precisely to explicate these rules. We simply play a game, and once the game reaches an equilibrium, we are initiated as competent users of a particular term, while for the most part, having no idea what we are doing or how we are doing it. It is important to note here once again that we are

[12] For the ease of explanation, we use situations here. In proof theory, Elimination and Introduction rules are purely symbolic or syntactic.

not trying to account for linguistic intuitions. Thus the fact that we have 'semantic intuitions' and that, to some, they seem as a good explanation is not of direct relevance. We learn the logic of a term, but we don't learn what that logic is. While some terms like 'bachelor' can be explicitly introduced as 'unmarried male', others like our main example 'red' have no such easy rules.[13] Even in the case of 'bachelor' one only postpones the real work to eliminating 'unmarried male', or put slightly differently, if it wasn't for the existence of a synonym the introduction and elimination rules would be at least as complex as those for 'red'.[14]

Looking at the case of 'red' more closely reveals that there is nothing simple about this 'simple' observation term. Although theorizing about colors goes at least as far back as Plato and Aristotle, it has been only in the last 50 years or so that, through the science of vision, we have started understanding what the conditions for introducing and eliminating 'red' are, and the details are by no means either semantic or straightforward. One can say that the term 'red' picks out a stable but complex pattern. Lets get to the details of how this "game" plays out.

As a general picture of the process, let's say that a speaker S plays an non-ending communication game against varied communities of *experts*. The goal of the game to learn new terms by approximating the experts' usage as closely as possible. Obviously, S's success will depend not only on her processing capacities, but also on the external circumstances such as her exposure to experts and her choice of groups to align with, and thus this process will essentially be both selected and selective.

The study of which experts we align with belongs mostly to human primatology as the root of these alignments themselves is, for the most part, biological. Still, a couple of interesting philosophical observations about the alignment procedure can be made (see below). Our dynamics will take a shape of a game. The idea here is that your usage of a term t is currently stable if you can answer a certain variety of questions pertaining to t. Or, to put it differently, the function that produces your uses of t is stable if it is appropriate for all the circumstances that you encounter. Once you encounter a circumstance where your function produces an inappropriate usage, you alter the function to account for the new circumstances. A counterexample forces you

[13] We invite the suspicious reader to explicate the logic of say *between*, *left* and *right*. For a more standard example of the difference between using logic and understanding logic, consider the difference between arithmetic and its formalization. While we teach arithmetic in grade school and consider it fairly trivial, we teach the logic of arithmetic in grad school, and consider it one of human kind's finer scientific achievements.

[14] The work of R.E. Jennings on the evolution of logical uses of disjunction and conjunction shows, decisively in our opinion, how little we understand our usage of terms. Even the seemingly simple ones like 'or' and 'and' turn out to be a lot more elusive that they first appear. An interesting empirical question is what gives us the semantic illusion of logical competence.

to move towards a new equilibrium.

[example, learning that black is not a color but absence of color/light is an interesting update]

3.2 Introduction Rules

The introduction rules are temporally posterior to elimination rules in that we mostly learn how to understand terms before we learn how to use them. However, a term must have been introduced before anyone can eliminate it[15]. The intro rules are thus logically prior, and we introduce them first.

It will be useful to abbreviate some terminology. S is a speaker, the person uttering a sentence. A is the audience, the individual or individuals processing the uttered sentence. A is nonempty, and presumably finite, but there are no other restrictions on it. It is even possible, as in the important case of internal dialogue below, that both S and A are the same person! E_t^S is the set of individuals that S considers experts with regards to the term t. The set E_t^A is defined via the obvious symmetry. The sets E_t^S and E_t^A are dynamic, that is, they change over time, they are not explicitly defined and are most likely not extensional. It is not so much that we have a list of experts that we carry around, but rather we have a capacity to recognize experts; we know an expert when we encounter one. The function f_t is the introduction function for the term t and similarly g_t is the elimination function for t. These functions are at the crux of our view, but essentially they are what enables us to use t and understand t respectively. We will say a lot more about these functions below. U is an utterance of the form $xxxREDxxx$ that contains the term 'red'. As 'red' is the only term we use explicitly in the definitions, we will use U', U'' etc. as abbreviations for sentences containing 'red'. For the purposes of readability we are keeping the definitions to the particular term. Generalizing the definition is obvious enough.
Here is the *update* algorithm for term 'red':
We begin with the speaker S who as a somewhat competent speaker of English has a some variety of terms at her disposal. Among them she has the introduction rule f for 'red'.

1. S utters U to an audience A.
2. If the audience A is capable of interpreting the utterance U without oddity, keep the introduction function f, end subgame, move to the next round of the game: the next utterance is made, action takes place, or whatever.

[15]Though if we had to make a guess, it would be that the temporal precedence here is not as obvious as it first appears and that the dynamics plays its role in this case too.

3. If (someone in) the audience finds the utterance U inappropriate for the circumstances[16], then first ensure the term 'red' is the cause of the oddity, and that A is an expert on the usage of 'red', that is, $A \subseteq E^S_{RED}$.[17]

4. If $A \subseteq E^S_{RED}$, then f needs to be updated to account for the discrepancy. Move to 5. Else, if $A \not\subseteq E^S_{RED}$, end subgame, move to the next round.

5. S now needs to guess new introduction function f' for 'red'. There is of course no guarantee that the guess will be successful (though we are quite good at this sort of update guessing), but a successful guess has the following properties:

   - It accounts for all the previous data concerning 'red' consistent with the new usage.
   - It accounts for the new usage.

6. Following 5, adjust the elimination rule g for 'red' to account for the new information. This is done by playing an 'internal' round of elimination game as both S and A, that is, the elimination function is updated silently via the internal monologue (see the next section for the elimination game).

7. Having made the guess, end the subgame. Move to the next round of the game.

3.3 Elimination Rules

The elimination part of the game is essentially integrated with the introduction part and once again, the game is played indefinitely against a community of experts. Here is an example of a round of the elimination game:

We again begin with the speaker S, and a singleton audience A. A–as a reasonably competent speaker of English–has a wide variety terms at her disposal. Among them she may or may not have the elimination rule g for 'red'.

1. A speaker S utters U.

[16] In our opinion, there is no simple definition of what makes an utterance odd to a person. It could be incoherent or inconsistent with their own usage. This we think is the crux of the oddity, but there are other issues as well. Is the term appropriate for the social setting? Is the term giving the sentence inappropriate emphasis? We are currently working on a full paper on these issues, and a brief definition or a paragraph cannot serve the intricate issue justice.

[17] This, in the simple cases amounts to checking that the audience is not incapacitated in some way, drunk or visually impaired, cognitively challenged, speaks English, etc. When we move to scientific cases, the story gets fairly involved.

2. If A has an elimination rule g for 'red', go to 3, else go to 7.

3. A computes $g(U)$, that is, she computes the utterance U using g as well as other relevant elimination rules. If $g(U)$ computes without oddity, move to the next round of the game and keep g intact. Else, if g produces some computational oddity, move to 4.

4. Check that the term 'red' is the culprit, $S \in E^A_{RED}$. If $S \notin E^A_{RED}$, end elimination subgame, move to the next round of the game. Else, go to 5.

5. Make a guess of a new elimination function g' for 'red'. Successful guess has the following properties:

   - It accounts for all the previous data concerning 'red' consistent with the new usage.
   - It accounts for the new usage.

6. Adjust the introduction rule f for 'red' to account for the new information. This is done by playing an 'internal' round of introduction game as both S and A in your head, that is, as internal monologue.

   End subgame, move to the next round of the game.

7. If A does not have an elimination rule for 'red', she plays the following (Question/Answer) game until her usage of 'red' stabilizes, that is, A obtains an introduction and an elimination rules which allow her to compute usages of 'red' and, as far as she know, align with S's own usage.

   (a) QUESTION: Guess an introduction rule f for 'red', and an utterance U' containing 'red'. Utter U'. Wait for the response from S.[18]

   (b) If S does not find A's utterance odd, A guesses an elimination rule g' that computes the original utterance U together with the new utterance U'. If the computation is successful, the round ends, the players move to the next round.

   (c) If S disapproves, then there are again two cases.

[18] Asking the right kinds of questions that provide us with the maximal amount of information about the term in question is an art that we tend to be good at. The games then usually last only one or two rounds before the usage is stabilized. The questioner gets the minimal information that they need to make the guess by asking the maximally informative questions.

i. S simply denies A's utterance U'.

ii. S produces an ANSWER utterance U'' containing 'red'.

(d) If (i), go to (a).

(e) If (ii), guess an elimination rule g''. If $g''(U)$ and $g''(U'')$ both compute, stop. Else, go to (a)

3.3.1 Some observations

There is of course no guarantee that the question/answer game with respect to some term t will terminate with stable introduction/elimination rules. You may simply be unable to learn the rules, not have enough time or background, or simply not care enough to pursue the game for long enough. Furthermore, it could be that the speaker herself does not have a stable term usage. In this case, it is all but hopeless for you to learn a stable usage of t. For a clear example of this, one needs only to look at the case of an incompetent parent trying to teach his child basic algebra.

Furthermore, whenever a player is forced to make a guess, she may simply be unable to produce the right kind of guess. This may formally correspond to the informal idea of the inability to learn. It should be further obvious that one's understanding of a term depends on the number of rounds of the game played with experts as well as on the quality of experts. In other words, when it comes to language competence, *education matters*. But another aspect of learning is emphasized in the question/answer game as well. The ability to ask the right kinds of questions is essential to one's learning process. A good question is the one that given your informational background and needs will entice a maximally informative answer. On this view, the process of learning is not unlike querring a search engines like google and yahoo: ask a wrong question and you will not get the information you need to update your usage. In one or two rounds, with the right kinds of questions you can fine tune the understanding of a term.

A possible implementation:

```
   Use genetic algorithms to devise a guessing procedure for
introduction and elimination rules for some class of terms, say,
color terms. One can think of the rules as either programs or
artificial neural networks. New data forces further evolution
of such program/network. Temporal constraint can measure how
successful such attempts are where an obvious measure of success
is whether there is an equilibrium between the usage of the
expert and the program.
```

3.4 Equilibria

3.4.1 Definition of Introduction Equilibrium

An equilibrium, Eq, is always relative to a class of experts, E. An intro equilibrium is a fixed point defined as follows: *A speaker S has reached the* introduction equilibrium *with respect to term 'red' at some time, t, iff for any utterance, U = "xxxx red xxxx", containing 'red' that S's introduction function, f, produces, and any expert $Y \in E^S_{RED}$, Y approves the utterance U.*

That is at least at this time, given what the community of experts in E^S_{RED} knows about red, there is no need to update the function, f. This of course could change as either S aligns with a different set of experts, or experts in E^S_{RED} learn more about red.

3.4.2 Definition of Elimination Equilibrium

An elimination equilibrium is again a fixed point. *S has reached the 'red'* elimination equilibrium *at a time, t, iff for every utterance U = "xxxx red xxxx", containing 'red' uttered by an expert $Y \in E^S_{RED}$, S's elimination function, g, computes U.*

The intro equilibrium corresponds to the competence in language production. The elimination equilibrium roughly corresponds to the competence in processing spoken or written language.

It is important to note that once S's introduction (elimination) function has reached a fixed point, it is not necessarily identical to that of all or any of the experts. It just covers the finite range of foreseeable cases in the 'correct' way. And this is where things get interesting since in some currently unforeseeable circumstance, the functions can behave differently. This feature is essential for the evolution of logical structures. More on this later.

3.5 Definition of an Expert

There is a fair amount to be said about what qualifies one as an expert with regards to a term. We will avoid some of the discussion (for now) by defining the sufficient conditions for being an expert:

A speaker, S, is an expert with respect to 'red', if there is an S', S' is an expert, and $S \in E^{S'}_{RED}$, that is, there is another expert that recognizes S as an expert.

More interestingly, someone counts *objectively* as an expert if they are able to win a certain number of question/answer games against any opponent. An expert has the most sophisticated and encompassing introduction and elimination functions with respect to the terms of their expertise. This objective expertise will have to be defined in such a way that it allows both for multiple experts, and disagreement on introduction and elimination rules among experts.

We can also define a *disagreement among experts* over a term t if there is an utterance U containing t such that one expert accepts U and the other rejects U. The most fruitful kind of disagreement is the one that can be settled by either logical analysis or by experiment. Examples here include not only empirical claims, but also hypotheses in mathematics, as well as the fruitfulness of programs in both science and mathematics.

4 Kinds of Equilibria

It was Descartes who claimed (perhaps jokingly), in the opening sentence of the *Discourse on Method*, that good sense is equally well distributed:

> Le bon sens est la chose du monde la mieux partagée; car chacun pense en être si bien pourvu, que ceux même qui sont les plus difficiles à contenter en toute autre chose n'ont point coutume d'en désirer plus qu'ils en ont.[19]

However, contra Descartes, not everyone will be able to achieve an equilibrium regarding any term, t. For instance, achieving equilibrium regarding, say, a concept of a large cardinal will take more time, memory, and 'processing power' to produce the right kinds of guesses than most of us posses or are willing to invest. Perhaps what is at the heart of this Cartesian illusion is the fact that most terms have *seemingly* simple introduction and elimination rules. Our physiology is just well-suited for learning such terms. Or, to put it in the terms of the evolution of our species, we have evolved to learn those terms in the same way birds have evolved to fly.

Even if it does turn out that much of language is to us as flying is to birds, there at least seems to be much greater diversity in processing capacities and kinds of maneuvers that one can learn. This stems largely from the fact that much of language—especially language suitable for science—has evolved after the evolution of the brain. Thus language evolved with the kinds of hardware already determined. The problem that scientific languages had to solve is what kinds of terms can be evolved by parallel processing in the available hardware. Different kinds of hardware (brains) were used for evolving terms in different domains (e.g., arithmetic, geometry, poetry).

4.1 Local and partial equilibria

There is another important insight in Descartes' quote above. If your equilibrium with respect to a term is stable—you have no evidence that needs to be accounted for—then

[19] From *Discours de la méthode* by René Descartes at Project Gutenberg.

to you, it seems that there is nothing more to learn. In fact, it would come as a surprise (and a major breakthrough) that there is more to t than you had originally surmised. Thus, not everyone's equilibrium with respect to a term, t, will be the same. One can achieve a minimal or a low-level equilibrium by being exposed to the minimal set of data regarding t. One's equilibrium can evolve from a low-level naive equilibrium to a more sophisticated higher-level equilibria. This is one of the aspects of education, where one's usage of concepts is being deepened. The other of course is broadening, where one acquires new terms with their introduction and elimination rules.

The most obvious example of an unstable equilibrium is the internally inconsistent usage. For instance, if you have the propensity to use term t and its synonym t' differently, while recognizing that they are synonymous, both t, and t' are unstable. For a more concrete example, if you claim that it is morally permissible to perform a particular medical procedure on a patient if it causes death in just 20% of cases, but not morally permissible if the same procedure kills 1 in 5 of the recipients of the procedure, and you also recognize that '20%' and '1 in 5' are synonymous, then both terms are unstable. Internal monologue often identifies instability in an equilibrium. Another important class is instability is with respect to some class of accessible observations. Thus if you believe that the earth is flat in the 21st century, your usage of 'earth' is empirically unstable.

A quasi historical example of the evolution of an equilibrium is in order here. The example is meant to demonstrate how one could have progressively sophisticated conceptions of 'geometrical object', each one as stable and logically as impeccable as the previous one. In fact the same individual can hold and apply different conception geared towards different applications and purposes.

4.1.1 From topology to metric space: evolution of an equilibrium

History of the nineteen-century geometry will provide the sufficiently detailed backdrop. Although the standard metric Eucledean geometry had a status of an excepted dogma for purposes of mathematical and physical sciences at the time, a great diversity of alternate geometrical structures were developed in the nineteen century. They ranged in nature from non-Euclidean metric geometries of Lobatchevsky, and differential geometry of Riemann and Gauss, the projective and affine geometries, all the way up to topology. Each one of these theories provides a conception of geometrical object. What interests us here are the formal relationships among the specimens of this widely diverse set of theories of space and spatial objects. The relation was investigated by the influential nineteen-century mathematician Felix Kline. He found that far from being disparate incompatible systems, the geometries of the nineteen-century can be categorized by invariance under groups of spatial transformations. Further, and more importantly, the groups of relevant invariances can be straightforwardly ordered by inclusion, and they fully specified the geometries in question and classes of objects they can distinguish. The more spatial structure the geometry in question has, the more invariances are required to preserve the essential properties of its objects. That is the weaker the spatial theory, the easier it is to be and remain the object of that kind under changes. Or to put it yet another way, the weaker the theory the more tolerant of change it is.

Listing invariances from the weakest to strongest, homeomorphism is the invariance that captures topological properties, projectivities capture properties of projective geometry, linear transformations capture affine geometry, and rigid transformations capture Euclidean metric geometry. It obviously follows from the strict inclusion that for instance rigid transformations preserve projective, affine, and topological properties, linear transformations preserve projective and topological properties, etc. What this means is that we can look at the four theories of space–topology, projective, affine, Euclidean geometries–as progressively refined sets of lenses for viewing the physical space. In our terminology, the theories progressively logicalize further detail of the physical space. So what is a spatial object according to these four sets of lenses? In the roughest of outlines, in topology something counts as an object if it has boundaries. Nothing destroys the object save a new tear, or filling or repair of an existing tear. In other words, topological objects are extremely malleable. You can bend them and stretch them without any danger. An object remains the same in the projective setting if there is an angle and a distance from which the new object looks exactly like the same as the object one started with. So the size and shape can vary to some extent, but one cannot bend and stretch however one pleases. The affine case preserves object's shape in as far as it can be done while ignoring angles size and location, but preserving parallel lines. FInally, Euclidean objects have not only determinate size and shape, but also determinate coordinate position. To come to our point explicitly, one can think of these four theories of space as providing logical structure for the term 'spatial object' at various levels of detail. Thus topology provides a perfectly stable equilibrium for the term, though, there is lot more more to be said about *spatial objects*. We can think of this stable equilibrium as the ordinary usage of the term, that is, its dictionary definition. There remaining three theories can be thought of as the progress of the refinement of the usage of the term that comes with increased scientific understanding of space. As more spatial patterns get noticed and encoded, the theory of spatial objects becomes more sophisticated.

The story here is meant to illustrate our picture of the evolution, logicalization, and sophistication of terms. It is oversimplified in more than one direction. FIrst, there is a sense in which the topological properties of space are more sophisticated than the standard Euclidean ones. This is true definitely in the sense that being more abstract they are further removed from the physical pattern recognition. The fact is further confirmed by the fact that topology is the youngest of the above four theories. SImilar observations can be made for projective and affine geometries. Second, the picture of increased sophistication of terms along a clean linear ordering itself is too simple for most purposes. The new scientific term most likely only partially overlaps with the old scientific term or with the folk usage. In addition, the ordering, if there is any, is more likely to be a messy partial order than a neat linear inclusion. For a well-known example, the scientific term'fish' does not simply take over the folk term and give us more detail. Rather the theoretical connections force us to clean the term up, exclude whales and dolphins, and include other species not previously thought to be fish.

5 On Guessing the Rules

It looks like we are putting a lot of stress on the notion of guessing. It seems at least intuitive that we learn new vocabulary rather readily. Sometimes no more than one round of the game suffices. Our goal for future developments and implementations is to construct a formal and structured notion of guessing. At this point, we want only to make several remarks about it.

The problem of how the guesses are made is as difficult as it is interesting. We are going to try to produce a sketch of how the story may go. We will use an analogy which we will try to unpack in some detail later. But for now, let's begin with rock climbing as a running example. What happens when a rock climber attempts a difficult section of a climb, say between two bolts, A and B? And let's say, for the sake of the example, that there is a finite number of both hand holds and foot holds. What the climber needs to do is to 'guess' the sequence of moves, that is, the sequence of machings of hands to handholds, feet to footholds, and body positions, just by looking at the rock. This initial plan—which is analogous to the already existing introduction rules used to form sentences—does not have to be complete, and more importantly, it is essentially revisable as new information comes along (e.g., the rock is steeper than it looks, the holds put the climber out-of-balance, etc.) Now, if all goes well, there is no need to do any revision. However, if the sequence does not succeed and the climber does not reach point B, it is back to the drawing board where one must first try and identify the problem. (What was it? Which part of the sequence? Was it handholds? Footholds? Body positioning?) Once the problem has been identified, a new guess of a sequence can be submitted to the test. Similarly, if the utterance fails the expert's muster, one first identifies the source of the problem, say usage of the term 'red', and then ventures a new guess for testing. The guess depends on many variables, most obviously previous guesses, but also on the sort of counterexample at hand, and the empirical adequacy of the new utterance.

The latter is of course crucial in the most interesting empirical claims/terms.

Our analogy here between sequencing physical movement and sequencing utterances is not incidental. It seems that feedback loops used in learning and producing language are not all that different from those used in producing complex motions like throwing a ball, catching a ball, or, for that matter, rock climbing. The main difference is that in addition to physical feedback, we also essentially have social-expert-based feedback, that is, feedback that depends on the social as well as physical data.

As food for thought we offer some additional similarities. Just as there are often many ways of climbing a segment of a rock climb, so are there many different ways of saying the same thing. Different people will do it in different ways; different people excel at different kinds of climbing (crack, face, slab), just as different people process different classes of terms more readily (a physiologically based division of labour). The more experienced one is in the usage of a term, t, the better one's chances of producing a workable guess, just as better climbers are more likely to climb difficult sequences. The finer the repertoire of moves, the more elegant the climbing, just as the finely tuned scientific vocabulary readily produces novel and elegant results. Also, in both cases, if one has been in a similar situation before one is more likely to produce the right sequence.

Our purpose in using this analogy is not only to link language learning and production to another, more physical process, but also to show that the ingenuity involved does not seem to be different in kind to that of other neural processes. The difference is more that of application than that of kind.

6 Functions f and g

The introduction and elimination functions f and g are at the core of our approach. The obvious question to ask is why we did not introduce some more explicit parameters of these functions. For instance, we could have defined a set of contexts or possible worlds $\mathbb{W}$ and a set of possible utterances $\mathbb{U}$, and insisted that say $f : \mathbb{W} \to \mathbb{U}$. In other words, we could have smuggled the semantics through the back door. The reason that we do not define f and g in any such explicit way is that, on our view, such a definition simplifies matters to an unproductive degree. In our opinion, the crux of understanding language in a deep way is in the understanding how these functions work. The devil is in the details, and details ought not be swept under the rug. We can give a rough sketch of what we take such functions to likely end up looking like. Let's take the introduction function f for the term 'red'. We take, as a first approximation, such a function to be a (natural) neural network, the weights of which are determined by a multitude of input variables and the structure of which may very well be determined genetically. The variables are of at least 4 different kinds: variables carrying a physical value (what frequencies hit the retina), the processing filters in the brain that normalize the input to the light conditions of the situation, logical or relational input (how input compares to other kinds of input, what kind of property it is, what it rules out, etc.), and grammatical properties that tell us how to construct sentences using 'red'. Trying to specify such a function purely logically or semantically in our view misses the target, and until the full physical details have been logicalized, that is, until we have a complete story of how vision works, such a specification is unlikely to be correct or informative. Thus, although introduction and elimination rules introduce logical properties into our epistemic structure, they themselves are strictly not purely logical. The semantic story is akin to Molliere's rendering of Aristotelean science by which the issue of why opium puts one to sleep is resolved by the "discovery" that opium has "dormative properties". The explanation that we require is more along the lines of the explanation that would be provided by neurobiology and chemistry. There may be kinds of terms, say mathematical terms, that are fully analyzable semantically, that is, all the input variables for f are semantic. If there really are such terms, it would certainly be productive to find a useful formal specification of which terms fall into that class.

We submit some of the predictions that this $f - g$ functional approach to introduction and elimination rules suggests.

First, most terms will have multiple disjunctive uses. As one opens a dictionary one notices that a term has a set of greatly varied meanings. When terms enter dictionaries in this way, the difference in meanings has been recognized, but in most cases disjunctivitis in the usage of a term will go unnoticed. This is what makes it nearly impossible to specify f and g in terms of necessary and sufficient semantic conditions. Thus the prediction is that ambiguities will be inherent in usages. Further, since usages are fine-tuned based on a finite number of interactions, a qualified user will still have a number of cases where the function underdetermines the usage. In other words, vagueness is also inherent in many terms. These kinds of 'deficiencies' of usage, though perhaps unavoidable, will of course be unsuitable for serious science, hence, the projects of Leibniz and Frege to make a special-use scientific language: logic.

Second, over time, the number of usages of a term will increase as a sufficiently similar pattern will be 'attracted' by the term. This is a kind of dynamic attraction that our brains seem to be good at. For instance, according to the neurobiologist

Ramachandran, we associate noises like 'bubu' with something soft and rounded, and 'kiki' with something spiky and sharp [20].

Thus some new soft feature of the world may 'migrate' towards 'bubu' and a sharp feature may migrate towards 'kiki'. This kind of dynamic migration seems to be responsible for migrations like the use of 'gravity' moving from the latin 'gravitas' to mean grave or serious to its Newtonian physical usage.

Third, most terms will pick out complex patterns that will be rather unsuitable to picking out regularities. When you have a heavily disjunctive usage, it is rather difficult to find a regularity that will fit all uses. Explicating and specifying usage is thus an important goal of science. When one disambiguates a term, one makes its equilibrium more sophisticated and more likely to be predictive.

The term 'belief' is a good example of a hopelessly disjunctive term. There is a great variety of circumstances in which it is appropriate to introduce this term, and it seems unlikely that any single theory will usefully account for all such usages. For a few instances, it is appropriate to say that S believes that P, if S said P, but sometimes it is appropriate to say that S really believes P even when they assert not P. Also, one can say that S believes that P if it seems that S is in the right kind of position to utter P, or S has been exposed to a situation that should or could have make him utter P. Or everyone should believe that P, hence S does, as in the cases of simple mathematical truths like $2 + 2 = 4$, etc. In other words, the usages are too disjunctive to be of any use in forming laws.

7 How to think with introduction and elimination rules

It is essential that the introduction function, f_t, and the elimination function, g_t, be different. f_t tells you how and when to construct a program or action recipe U using the term t. g_t on the other hand tells you, given such an utterance or program U containing t, and given elimination rules contained for terms other than t, how to execute or compute U.[21] This suggests that we can conceive of the internal processing of introduction and elimination functions as thinking: the familiar internal monologue. Here

[20]For a live experiment see work of Vilayanur Ramachandran, especially his TED talk in 2007, www.TED.com.

[21]For a simple example of this difference, we urge the reader to look at the introduction and elimination rules for $\vee$ (or) in any natural deduction system for classical propositional logic. The rules are ostensively different and yet they both play an essential role in proving completeness of the system, that is, they are both necessary.

is how it is done in detail.

Suppose you want to find out more about a term, t. Do as follows. Using introduction rules and, in particular, an introduction rule, f_t, guess a novel sentence, U. Process U using g_t. If the processing produces an oddity, that is, it doesn't compute, update your introduction and elimination rules to eliminate this oddity. Proceed to form another hypothesis, U', containing t. Stop when tired. This kind of process really gives life to the philosopher's claim that the conclusion is when one gets tired of thinking! Changing the elimination rules in this way does not have a direct influence on the evolution of the public language, but changing the introduction rules may. If new utterances are made containing t, those utterances are now in the public domain and can influence the structures in which t plays a role. Further, different introduction and elimination rules across individuals explain why different individuals make different contributions to the study of a particular subject.

Oftentimes the process of 'thinking' starts with an external question. This question guides the kinds of guesses that one makes using f_t and g_t. If the question is ill defined or sloppy, the answers produced are confused and difficult to compute. For instance, if one thinks of chemical substances as spirits, one will ask questions like 'What does fire have in its right hand?', a question that will be unlikely to produce sharpening in the usage of the involved terms. Thus, it is essential to the success of thinking not only that the terms be carefully chosen, but also that the questions head in the right direction. Furthermore, given the limited power of our processing units, an ability to decompose questions into simpler component questions is essential. We can also explain planning as a sequence of guesses (or "action recipes"), some of which are eliminated in computation as new approximations are made.

A possible implementation:

```
Devise a thinking interaction. Two programs F, an introduction program, and G, an elimination program, interact. F guesses an answer, A, to a question, Q. G attempts to compute A. If it does not compute, G returns no as well as the log of the problem. F produces another guess, A', based on the log, etc. Guesses themselves are programs, so we can think of the search for an answer as the evolution of an appropriate class of programs. The elimination program provides us with an environment in which the evolution takes place.
```

8 Introducing a new term into a language

One of the most interesting aspects of the evolution of language, and thus the evolution of logical structures, is the introduction of new terms. The dynamics of this process, often overlooked in the relevant literature on development of theories, we will argue, play an important role in the advancement of our understanding of our environment. Our natural neural networks are genetically wired to pick out patterns in our environment as well as to complete incomplete patterns. The pattern recognition algorithms alone, as we argued above, do not constitute knowledge in the interesting public sense of the term. Essentially what needs to happen in order for these patterns to become knowledge is that they need to be logicalized. That is, one needs to develop the appropriate terminology for expressing the patterns in the public language.

One needs to build the basic logical building blocks, namely *terms*. What this amounts to literally is introduction of new vocabulary, or at the very least serious tinkering with the structure of some 'old' vocabulary. The logical structure inherent in this new vocabulary then enables the scientist to express the regularities she has discovered. For example, we talk about the discovery of atoms and molecules. What this amounts to, on our view, is:

The scientist has discovered new patterns perhaps about new entities. But we essentially want to be ontologically uncommittal here.

The second step, expressing the patterns in language is what is epistemically significant. It may be, on a realist picture, that I am recognizing patterns using molecules and atoms in my environment. But until the vocabulary was devised, and patterns concerning atoms and molecules were recognized in this language of atoms and molecules, there was no knowledge of atomic physics and chemistry. That is, until the physical is logicalized, there is no knowledge.

The process of logicalization basically enables us to learn chemistry from books, or to train our brains to interact with (the parts of) the public logical structure that we call 'chemistry'.[22] It suffices to say that the 'right' kinds of terms for capturing the chemical or any other patterns are hard to come by. A quick glance at the history

[22] The issue here seems to us independent of the realism/antirealism debate. One can be realist or not, and still hold that the universe does not have a language handed down to us. Rather we have to *make* such language up, and this process is what we are after in this project. The miracle that language says anything however imprecise, statistical and vague about the physical world seems to us to be the miracle worth looking into. This goes back to the old question of how is it that mathematics is so successful in talking about physics.

of chemistry uncovers all kinds of 'unsuccesful' enitites, from phlogiston and the four Aristotelean elements, to spirits inherent in various substances. What seems really wrong with the unsuccessful scientific vocabulary is that it does not lend itself to recognizing patterns via generalities or laws. Of course whether a term is successful is not an issue that can be painted straightforwardly in black and white. Oftentimes the best terminology for describing patterns in some domain performs just slightly better than random; we continue to seek better terminology, but at the time, those terms are all we have. Whereas 'numbers', 'atoms', 'molecules', and 'species' are clearly successful (if not always entirely easy to explicate), terms like 'agent', 'will', 'action' are much harder to asses. Do they form any useful testable generalities? After all, many terms in any natural language do not, for example terms used to pick out Greek (and other) gods and angels.

A rich class of successful terminology more easily lends itself to introduction of new successful terms. The strategies of modern science and mathematics are, as far as we can tell, the most fertile grounds for breeding new species and kinds of terms. Empirically or purely logically contentful terms oftentimes give raise to new kinds of contentful terminology. There is a sense in which most kinds of terms of natural langauge pick out some pattern, although most such patterns are rather "noisy" and hopeless to explicate. Ethical vocabulary comes to mind here.

It is important to note here that we do not only recognize physical patterns, but we also recognize meta patterns, that is, similarities among various kinds of patterns. Thus while the former is in the domain of the empirical sciences (whether they be purely physical, or social, economic etc.), the latter form formal sciences. Recognizing the abilities and shortfalls of the terminology we have enables us to open up space for the introduction of new logical structures that then can be further applied to empirical sciences.

Thus, by whatever means one has come to recognize a new pattern,and by whatever analogy (or filling of the logical space) one has devised a new term, t, the introduction of the term into the public language takes the form of the introduction/elimination (or question/answer) game. This game oftentimes further shapes the ability of the term to successfully recognize and generalize patterns. The history of the term 'set' is a good example of this. Cantor's original conception was severely flawed, but making it public enabled others to devise a stable usage of the term. The stable equilibrium for the term 'set' nowdays is mainly derived from the Zermelo-Fraenkel formalization of the term 'set', though there are certainly other stable uses arrived at via different sets of axioms (e.g., predicative set theory, constructive set theory, etc.).

A speaker (scientist) S attempts to introduce the term in a sequence of utterances $U_1, ..., U_n$ that contain t with most of the remaining terms in $U_1, ..., U_n$ having already established equilibria. This explains why rich languages are more amenable to introduction of new vocabulary: the more diverse a logical structures one has, the more new terms one can introduce. Occasionally, or perhaps in most cases, more than one term is introduced at a time. The terms form logical dependencies and so finding the right kind of introduction and elimination rules for one term depends on finding the right kind of introduction and elimination rules for all the others.

After a term, t, has been introduced via $U_1, ..., U_n$, the community of competent language users, A, then attempts to devise an elimination rule for t. If the elimination rule is not forthcoming, the elimination rule can be further shaped by guessing an introduction rule for t and producing a sequence of utterances $U_{n+1}, ..., U_{n+m}$ which S either accepts or rejects. If either $U_1, ..., U_n$ are inconsistent in the sense that the audience A is unable to devise the appropriate rules, or for some U_{n+k} which S accepts, $U_1, ..., U_n$ together with $U_n + k$ is inconsistent, then the term has failed.

In certain ideal cases we can give the introduction and elimination rules explicitly via a set of axioms. In most cases, however, the introduction and elimination rules form a complex pattern.

If a term has failed, S can keep the new term by giving up some of $U_1, ..., U_n$ and perhaps also adding further $U_k, ..., U_{k+l}$, that is, the term can be further clarified.

Good examples of terms that do not have explicit linguistic introduction rules are observation terms like 'red', 'green', 'salty' etc. In cases such as these, the elimination is ostensive, i.e., it is a complex but rather particular cognitive relation between the speaker and a class of objects/surfaces. It still remains, importantly, that however red seems to you, and however you process light frequencies, as long as you can learn introduction and elimination rules by some means, you can communicate using the logical structure captured by the term 'red'.

A possible implementation:

Devise a general class of models to show that richer successful vocabularies lend themselves more readily to introduction of further successful terms. That is, if you have a vocabulary that successfully captures some patterns in the environment, you can, based on this vocabulary, devise further vocabularies to pick out additional patterns. Historical examples that come to mind: integers with division lend themselves to introduction of rational numbers, rational numbers with the function of taking the square root lend themselves to introduction of real numbers, etc.

8.1 Evolution and Introduction of Grammatical Structures

As important as the basic logical blocks for building our public epistemic structure are, there is nothing special about their evolution. They just form a conveniently simple model of evolution of kinds of logical structures. Another important class of logical structures that enables us to capture more complex patterns in our environment are *linguistic constructions.* They enable us to build more complex logicalizations from the simple ones, and our hypothesis is that they evolve in the essentially same fashion as terms. One starts with languages that presumably only have a simple sentence structure and limited variety of terms. The complex syntactic structure co-evolves as the complexity of the recognized patterns, and complexity of corresponding terms, increases. It is of course possible that these structures are evolved in a biological fashion, that is, that every human already has a set of predetermined grammatical categories, but this does not strike us as the most profitable starting point.

It is useful to think of grammatical structures as programs. When the terms are supplied into a program, the program computes. The program will require certain kinds of terms and specific relations among terms to be able to compute. For example, Chomsky's famous example, "Colorless green ideas sleep furiously" does not compute, and neither does "Snow is red" but the kinds of errors produced are different. The programs can result in action (imperatives), in query (questions) or in an implementation of a logical structure that can be further evaluated and processed (statements). Thus, one can think of an utterance as an input code for the speaker, and an execution of the program for the audience. We say that a program produces an 'oddity' if it does not compute, that is, for some reason or another, the set of instructions is not implementable.

9 Metaphysics of Communication

Terms like "truth", "meaning", "reference", "information", *et al.* play no special role in our theory of communication. They are not theoretical entities but simply more terms in need of explanation. We feel no particular obligation to give these terms any special privilege. If our theory succeeds, their usage and evolution will be explained along with that of other terms. Thus, if someone were to ask, for example, whether it is true that snow is white, we would have to answer that indeed, it is true, but that simply is how the terms involved are used. Nothing deeper is claimed. The peculiar usage of philosophers should, if the theory has sufficient explanatory power, be explained alongside other kinds of usage. Our theory furthermore applies to itself. We claim that "evolution of terms", "introduction/elimination rules", and "equilibria" are theoretical terms the usefulness of which is measured in the same way as that of terms of any empirical theory. If the theory finds fertile ground, their usage may evolve further. They may acquire more detailed logical structure.

While we do not want to go into the metaphysical details of what logical structures are, several things need to be pointed out. First, since we build logical structures in communication, for every logical structure there is a definite time when it didn't exist. This we mean literally in the same sense in which for any actual physical object, for example, a building, a house, or a bridge say, there was a time before it existed. Further, although the logical structures constitute our knowledge, we have no privileged insight into all the logical consequences of the structures that we devise. Quantum physics, arithmetic, and set theory are presumably well-defined logical structures and constitute some of our finer epistemic achievements, but it does not follow from this that, as a species, we have the processing power to derive all of the consequences of those theories. We have finite processing power and we can only work and interact with small finite parts of the structures that we have helped evolve. This is the sense in which the structures are objective. Once they have evolved, they are there for us to explore in the same way that biological species are there for us to explore. The main difference is that the process of inquiry itself may influence further evolution of a species of term. Looking at a logical structure may force a change in that structure.

Logical structures are thus abstract, and yet temporal and always physically implemented. They are about as metaphysically mysterious as actual computer programs. This is as much as we are prepared to say at the moment.

10 Homage to Positivism

What then is the objective of science on this view? There are in fact at least two. Obviously, the main one is to devise the right kinds of languages that enable us as a species to logicalize as many patterns in our environment as we are capable of recognizing, to move as many patterns as we can from the realm of intuition, to the realm of public knowledge. The second goal is more subtle. It is to make the languages we use as systematic and explicit as possible. For example, although at present the introduction and elimination rules for 'red' cannot be made explicit, the hope is that there is a scientific successor of 'red' couched in some explicitly formal theory that enables us to fully logicalize 'red', or put differently, to make the rules obviously syntactic, however complex they may be. Thus there are three cognitive levels: the first is simply recognizing a pattern in the outside world; the second is logicalizing that pattern, that is, couching it in the public language and sharing it; the third is reflection on the logicalization itself and systematization of that logcalization. As a logical structure has evolved, we, at some point, recognize the logical properties of that structure. This then enables us to devise further logicalizations, and the dynamics play on.

There is yet another, more destructive goal of the logicalization of our environment, and that is to make the environment as hostile as possible to empirically and formally contentless terms making the system of public exchanges rigorous in a way that species of useless terms find survival increasingly difficult. Every term, at least in the realm of epistemic inquiry, is put to the harshest test of fitness. Its survival depends on picking out some formal or empirical pattern. The issues here are predictably subtle and difficult. For instance, look at the endless debate on whether 'god' or 'good' picks out any such pattern. Yet, at the very least, our species of dynamic positivism recognizes that just because you can say it does not mean that it makes sense.

References

[1] M. Aiello and J. van Benthem, *A modal walk through space*, Journal of Applied Non-Classical Logics **12** (2002), no. 3/4, 319–363.

[2] M. Aiello, J. van Benthem, and G. Bezhanishvili, *Reasoning about space: the modal way*, Journal of Logic and Computation **13** (2003), no. 6, 889–920.

[3] J. Barwise. *Three views of common knowledge*. Proceedings of TARK, 1988, 365-379.

[4] J. van Benthem and D. Sarenac. *The geometry of knowledge*, to appear.

[5] J. van Benthem. *Information as correlation and information as range*, ILLC manuscript, 2003.

[6] J. van Benthem. *Exploring logical dynamics*, CSLI publications, 1997.

[7] P. Blackburn, M. de Rijke, and Y. Venema. *Modal logic.* Cambridge tracts in theoretical computer science, Vol. 53. CUP, Cambridge, 2001.

[8] D. M. Gabbay and V. B. Shehtman, *Products of modal logics. I*, Log. J. IGPL **6** (1998), no. 1, 73–146.

[9] D.M. Gabbay, A. Kurucz, F. Wolter and M. Zakharyaschev. *Many-dimensional modal logics: theory and applications.* Studies in Logic and the Foundations of Mathematics, Volume 148. Elsevier, 2003.

[10] D. Gabelaia, A. Kurucz, F. Wolter and M. Zakharyashev. *Products of 'transitive' modal logics.* Submitted for publication, 2004.

[11] R. Goldblatt. *Diodorean modality in Minkowski spacetime,* Studia Logica, **39**(1980), 219–237.

[12] B. Lowe and D. Sarenac, *Cardinal spaces and topological representations of bimodal logics*, to appear.

[13] K. Kuratowski and A. Mostowski, *Set theory*, Studies in Logic and the Foundations of Mathematics, Vol. 86. North-Holland Publishing Co., Amsterdam-New York-Oxford; PWN—Polish Scientific Publishers, Warsaw, 1976.

[14] J. C. C. McKinsey and A. Tarski, *The algebra of topology*, Ann. of Math. (2) **45** (1944), 141–191.

[15] G. Mints, *A completeness proof for propositional s4 in cantor space*, Ch. 6, In Logic at Work, Kluwer, Publishing, 1998.

[16] D. Sarenac. *Modal logic and topological products*, PhD Thesis, Stanford University, forthcoming in 2005.

[17] V.B. Shehtman, *Two-dimensional modal logics*, Mathematical notices of the USSR Academy of Sciences **23** (1978), 417–424. (Translated from Russian.)

[18] E. Spaan, *Complexity of modal logics*, PhD thesis, University of Amsterdam, Institute for Logic, Language and Computation, 1993.

Fred Dretske1985. "Precis of Knowledge and the Flow of Information." In: Hilary Kornblith, ed., Naturalizing Epistemology. Cambridge: MIT Press.

Jaakko Hintikka. *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning.* Cambridge University Press, Cambridge, 2007.

R.E. Jennings "The Genealogy of Disjunction" OUP, Oxford, 1999.

Karl Popper. Objective Knowledge: An Evolutionary Approach. Clarendon Press, Oxford, 1972.