Language and Equilibrium

Prashant Parikh

1 Introduction

I learned very early the difference between knowing the name of something and knowing something.
—Richard Feynman, *The Making of a Scientist*

In this book, I present a new account of meaning for natural language.

The account has three levels. Most concretely, it offers a tool to derive and compute the meanings of all possible utterances, at least in principle. More generally, it provides a method to produce variant theories of meaning and to address the many problems and puzzles that beset its study. Most abstractly, it advances a way to think about meaning and language through the lens of a broad and powerful idea and image.

At the first level the account is a theory, at the second a framework, and at the third a paradigm.

The paradigm embodies the leading idea and image of *equilibrium*—or balance among multiple interacting forces. The framework draws primarily upon game theory and situation theory. These are the best tools available at present to implement the idea of equilibrium in our context of language and meaning. The theory uses the constraints that arise from game theory and situation theory to capture the meanings of utterances. This renders their derivation a more or less straightforward computational task.

The resulting account is called *equilibrium semantics*.

1.1 Brief Background

Although the study of meaning goes back to classical times in multiple cultures,[1] there have been two broad traditions in the philosophy of

---

1. See, for example, Deutsch and Bontekoe (1997) and Raja (1977).

language that have addressed meaning in the twentieth century. One is the ideal language tradition and the other is the ordinary language tradition.[2] Frege, Russell, Whitehead, and the early Wittgenstein were among the first contributors to the former, and the later Wittgenstein, Austin, Grice, and Strawson were among the first contributors to the latter. In the second half of the twentieth century, both traditions have borrowed a great deal from each other and have partly even merged, albeit uneasily.

From all the details of both traditions, it is possible to extract two central ideas, one from each tradition. The first tradition has contributed the idea of *reference*, of language's being *about* the world (or about extralinguistic entities in particular); the second tradition has contributed the idea of *use* or *communicative activity* in a broad sense. The first tradition tried to understand reference or the aboutness of language; the second tradition tried to understand use or the communicative function of language. Both ideas, incidentally, are nicely captured in the happy phrase *the flow of information*, a composite idea that underlies and undergirds the account of meaning in this book.

Ideal language philosophy originated in the study of mathematics and the logic of mathematical statements, which led to its emphasis on the idea of reference. In the main, it did not see mathematics as a *situated activity*, partly on account of its abstract and formal nature. This led to its ignoring the dimension of use and to its focus on formal logic and especially on translating natural language utterances into logical languages.[3] In fact, in the early days, the often inconvenient facts of use were treated as a kind of defect that would be removed by *idealizing* language. While this tradition has yielded many insights of continuing relevance including, crucially, its use of mathematical methods—expressed perhaps most of all in what is called formal semantics after Montague (1974b)—its attempt to extend its ideas from mathematical languages to natural languages has led to many difficulties as elaborated in section 1.4.[4]

2. See the collection of essays in Rorty (1988).

3. Perhaps the classic example of this is Russell (1905, 1919). To be fair to these philosophers, their interest in semantics was secondary to their interest in the philosophical problems they were attempting to solve. Semantics was a means to a philosophical end.

4. Indeed, there are problems with its account of the semantics of mathematical activity as well such as the problem discussed in section 7.3.

The practitioners of ordinary language philosophy, reacting to this ar-
tificial and ersatz[5] view of natural language, started from the vision that
natural language was an inherently situated activity and that the use of
language in communication should be at the heart of its study. Unfortu-
nately, they also appeared to believe in the main that this attribute of nat-
ural language made formal methods relatively inapplicable. This tradition
too has afforded many insights into the nature of linguistic meaning, al-
though its lack of a mathematical approach has often made its arguments
imprecise and vague.

The awkward state of affairs that exists at present can be best seen in
there being the two distinct disciplines of semantics and pragmatics, each
concerned with meaning, one primarily with its referential aspect and
largely formal and conventional, the other primarily with its use-related
or communicative aspect and largely informal and contextual.[6] While
both disciplines have drawn upon each other and have developed a great
deal since their originating ideas were established, in the mainstream view,
semantic meaning is still generally identified with a sentence's conven-
tionally given and formally represented truth-conditions, and pragmatic
meaning is generally identified with some combination of contextually
inferred and informally represented Gricean implicature and Austinian
illocutionary force.[7] These two types of meaning typically coexist and
may coincide or diverge in ways best exemplified perhaps in Grice (1975,
1989) and Kripke (1977).

Kaplan (1969), building on Montague's index semantics and Lewis's
(1972) contextual coordinates, introduced context into semantics proper
via his two-dimensional notion of *character* but this was intended for
just a limited set of expressions, primarily tense and pronouns. Stal-
naker (1970, 1978, 1998, 1999b), noting the ubiquity of context-sensitive

---

5. This word is used to indicate that the notion of an "ideal" language was an *in-
ferior substitute* for natural language and is not meant pejoratively.

6. In the literature, there are many different ways of drawing the distinction be-
tween semantics and pragmatics. Two of the most important are: between the for-
mal representation of meaning and the communicative aspect of meaning, and
between literal meaning seen as conventional and rule-based and implicature and
illocutionary force seen as contextual and inferential. These two different distinc-
tions can be and are often merged. A third related distinction is that between the
truth conditions of an utterance and the felicity conditions of an utterance.

7. The origins of the idea of force go back to Frege, so the neat picture pre-
sented above should be seen as a vast schematic oversimplification for expository
purposes.

expressions in language, generalized the concept of a context from an index to a *context set*—an entire "body of information"—for literal meaning, which is essentially the underlying idea of context that is prevalent today, including roughly the one used in this book, except that Stalnaker's notion is couched in the framework of possible worlds. The question of exactly how context obtruded into the sphere of semantics, whether conventionally or inferentially, or a combination of both, was largely unaddressed.

Barwise and Perry (1983) tried to reorganize these insights in a radical rather than moderate way by extending model-theoretic or referential methods to accommodate Austin's focus on use by inventing the key idea of a *situation*, something Austin had used informally[8] and something that captures Stalnaker's idea of a body of information directly rather than circuitously via possible worlds.[9] Among other advances, this led to a blurring of the boundaries between the two traditions and between semantics and pragmatics, although not in any very precise way. One problem was that situation semantics, as their account was called, involved an overly abstract and impoverished notion of use, as did the earlier efforts

---

8. See, for example, Austin (1979a).

9. To quote from Stalnaker (1999b):

> My central assumption was that a context should be represented by a body of information that is presumed to be available to the participants in the speech situation. A *context set* is defined as the set of possible situations that are compatible with this information—with what the participants in the conversation take to be the common shared background. The contextual factors relevant to interpreting John's utterance of "I love Mary" will then be, not simply the index, but the fact that the relevant body of information includes the information that John is speaking and that the utterance is taking place on June 14, 1998.

Stalnaker's (1998) context set plays a dual role: it provides both the initial body of information within which an utterance takes place as well as the final body of information that results from the utterance. So there is a single entity (a set of possible worlds) that changes through a discourse. An utterance situation and the resulting described situation, in Barwise and Perry's sense, are more general than this, as they need not be fully shared by the participants in the discourse (a point I discuss in detail in section 3.3.4 on common knowledge and section 5.7 on the indeterminacy of contexts) and the relation between them is different. As I see it, the described situation resulting from an utterance then affects the succeeding utterance situation in an ongoing discourse. I say more about this in section 2.6 and section 7.2.

Situation theory also transcends situation semantics and has been applied to areas other than language.

by Montague, Lewis, Kaplan, and Stalnaker[10]: they had no *theory* of use, just some notational "stand-ins" for broad aspects of use.[11] Nevertheless, I have found Barwise's development of a theory of information—called situation theory—as well as a few aspects of their attempt to combine the two traditions to be of great value in developing my own approach to language and meaning. Part of the reason for this is the conviction that if things are done correctly there ought to be just one unified theory of meaning rather than two uneasily juxtaposed accounts. This Hegelian *aufhebung* of the two traditions and the two disciplines is what will be attempted in *Language and Equilibrium*.[12]

The erosion of the barrier separating semantics from pragmatics has been underway from other quarters as well. Recanati (2004b, 2004c) as well as the Relevance Theorists (see Sperber and Wilson 1986a, 1986b, 2004; and Carston 2004 among others) have also been chipping away at this distinction (most strikingly with examples of so-called *free enrichment*[13]) and offering a more imbricated picture of meaning. The view that many linguistic phenomena that were previously seen as belonging to semantics in fact belong to pragmatics has come to be called *radical pragmatics* though, of course, in my view, these are all part of a *radical semantics* that I have chosen to call equilibrium semantics.

---

10. It is notable that Stalnaker (1996, 2006) has himself turned to game theory more recently.

11. Like "speaker connections," for those familiar with their account.

12. Despite the dynamic turn of Kamp (1981), Heim (1982), and Gronendijk and Stokhof (1991), most of these developments remain squarely within the tradition of Montague-inspired formal semantics where the focus is on finding appropriate meaning representations rather than on *deriving* intended and optimal meanings through use. Discourse representation theory, file change semantics, and dynamic logic are concerned more with the *results* of the communicative process than with communication itself, with the *what* rather than with the *how*. They address what Austin (1979b) called the perlocutionary act and effects of communication, not the locutionary and illocutionary acts and the securing of uptake and understanding. As such, they do not appear to question the *syntax-semantics-pragmatics* trichotomy and pipeline view of meaning bequeathed by Morris (1938) and Grice (1989) despite their undoubted technical accomplishments.

13. For example, when an utterance of a sentence such as "*Casablanca* is playing" is being interpreted, it has to be *enriched* or *completed* with a partial content such as *in New York this evening*. This enrichment is part of the literal content of the utterance. See Parikh (2006b) and section 4.4.1 for a detailed game-theoretic analysis of this phenomenon.

Typically, following Charles Morris's (1938) original trichotomy *syntax–semantics–pragmatics*, semantics is identified with what comes from the linguistic representation or with the conventional meaning of the representation and pragmatics is identified with the contributions of the ambient circumstances. Linguists especially use the term *underspecification* to describe this—semantics first underspecifies content that is later filled in by pragmatics. It is better to identify semantics with the problem of inferring the entire content,[14] regardless of what contributes to this content, the linguistic representation or the context. Indeed, it has often been assumed in the past that conventions suffice for getting at content so there is an ambiguity in the original identification of semantics with convention since it was implicit that convention would yield content. This perhaps explains the origin of the term *literal* content. That is, it is not clear whether the commitment should be to the literal, purportedly conventional source of content or to content per se. The mainstream view[15] of semantics has identified with the former, but I am urging the latter, especially since even literal content is ineluctably contextual.

A major advantage of the identification of semantics with the determination of content rather than with convention is that it allows a uniform view of all representations and symbols, whether they are linguistic, or belong to other modes such as the visual or gestural, or whether they are mental representations. The uniform view is that content of any kind is a function of two variables, the representation φ and its embedding ambience u. That is, the content can be written as $\mathcal{C}(\varphi, u)$, where φ stands for any representation, whether it is linguistic, visual, gestural, or mental. Indeed, φ can stand for any *sign* as well, including tree rings, footprints, or black clouds.

Secondly, this view of a single discipline for meaning prevents an artificial division into two subfields—semantics and pragmatics—of all the factors that should jointly contribute a unified theory of content. The former view takes the representations themselves as primary and more or less exclusive (the first variable in $\mathcal{C}(\varphi, u)$) and as the starting point for scientific inquiry, the latter view takes the flow of information and commu-

14. "Meaning" and "content" are used interchangeably in most places in this book. When "meaning" is used differently, as it will be later in certain contexts, it will be pointed out explicitly.

15. Even Austin (1979c) assumed that demonstrative *conventions* correlated utterances (statements) with historic situations. Perhaps Grice (1989) is ultimately responsible for the mainstream view.

nication and thought as primary ($\mathcal{C}$ itself) and as the starting point for scientific inquiry. If the field of language and meaning is seen as falling within the larger domain of information flow with human behavior as one central component of it, then we ought to be more inclined to the second view. Language then becomes just a *part* of all that makes meaning possible. As Dretske (1981) writes, "In the beginning there was information. The word came later." An exclusive focus on language takes hold of the wrong end of the stick and makes grammar primary, and meaning secondary and an afterthought. This leads to a parallel exclusion of context by focusing on "semantic meaning" (meaning derived almost entirely from the linguistic representation) as primary and "pragmatic meaning" (meaning arising from contextual factors) as secondary. Restoring the centrality of information and its flow enables a balanced view of the sources of meaning as such. And, as will be seen later in this chapter as well as throughout the book, the subject matter, including even syntax,[16] is best viewed not as a linear "stick" but as a *circle* instead.

Third, Austin (1975, 1979b) offered a critique of the semantics–pragmatics distinction that appears to have been largely ignored. His dialectical argument started by making a persuasive case that the meaning of at least some utterances is not a matter of truth conditions alone. While assertions require truth conditions, performative utterances require felicity conditions. Semantics would then be concerned with truth conditions and pragmatics with felicity conditions. But this argument places us on a slippery slope. Austin argued that truth conditions are themselves just part of the felicity conditions for uttering a sentence. This suggests that semantics is really a part of pragmatics or, to put the thesis in its most radical form, that there is no principled distinction between semantics and pragmatics. If illocutionary force is taken as an aspect of the content of an utterance, then once again this leads to a unified view of semantics and pragmatics.

Of course, to be convincing, the viewpoint being advanced requires a homogeneous framework that actually enables a uniform derivation of the full content of an utterance. I show it is possible to create such a framework from first principles.[17]

---

16. See section 7.1.1.

17. As will be seen later, content will be divided into locutionary and illocutionary content, but because context will play an essential role in the derivation of both types of content, this distinction is quite different from that between semantics and pragmatics.

An added advantage to offering a comprehensive and detailed *mathematical* framework for meaning is that many arguments offered today for or against a theory of particular phenomena remain inconclusive because their proponents often presuppose different views of semantics and pragmatics and also give *nonuniform* accounts for different classes of phenomena. For example, in the fascinating arguments over the last century for or against Russell's (1905) theory of definite descriptions, different theorists often assume different and incompatible positions on the notion of meaning itself and then advance a very particular theory of definite descriptions that may be at odds with theories of other, even adjacent, phenomena.[18] Such arguments may be seen as offering a perspective at *two* levels simultaneously, both an implicit argument for an idea of meaning and an explicit one for a particular theory. In contrast, this book provides a *uniform* approach to the derivation of the full contents of more or less all utterances, couched within an explicit and unified framework for meaning that synthesizes semantics and pragmatics. This does not obviate the need for particular theories but it makes these accounts reasonably uniform across phenomena.

Besides combining the central ideas of reference and use stemming from ideal language and ordinary language philosophy, I also depart from both traditions in fundamental ways. I see content as *indeterminate* in a number of specific ways, a facet of meaning that has not been seriously addressed before. Finally, of course, I introduce the idea of *equilibrium* to explain its many aspects, both traditional and new, in a manner that unifies them and provides a single idea and image of the system of language and meaning.

Thus, in a simplified and abstract way, it would be accurate to say that equilibrium semantics, the account of meaning presented here, combines four distinct ideas in a single unified framework: *reference*, *use*, *indeterminacy*, and *equilibrium*.

1.2   The Origins of Symbols

Observational cosmology suggests that the universe is roughly fourteen billion years old. By contrast, it appears that a little over sixty thousand years ago the human race broke into a *symbolic* consciousness, a new kind of consciousness that allowed it for the first time to use objects and events

---

18. See, for example, the papers in Reimer and Bezuidenhout (2004). Not all views—even in the collection cited—are so fragmented of course.

to *represent* other objects and events. The experience of death or the experience of play may have been among the first events that triggered this fateful break with our "prehuman" past. Until then, presumably, man[19] was submerged in a kind of *presymbolic* awareness that allowed him just "direct" perception and "direct" actions and interactions in the world, not unlike the condition of other animals. See Terrence Deacon (1997) for one account of what is involved in man's achievement of a symbolic consciousness.[20]

It is difficult for us to imagine this presymbolic state because this new cognitive ability must have suddenly transformed the universe into a miraculously different place, one with myriad individuals, properties, and relations. Of course, these entities had been there all along, and had been perceived, reasoned about, and acted upon in direct and unmediated ways, but their richer existence as we know it today required the ability to *name* them, to form discrete mental and public symbols corresponding to them, to pluck them out of the relatively undifferentiated store of the world. Overnight, the world must have become a repository of *information*, a novel ontological space in place of the old continuum. It seems reasonable to surmise that it was this fresh and unfamiliar power to represent the world to ourselves and communicate it to others, especially through language, that made us human.[21]

19. This word and its cognates are used in their gender-neutral sense.

20. See also the book by Ian Tattersall (2002). I quote two paragraphs:

If the modern human brain, with all its potential capacities, had been born along with modern human skull structure at some time around 150 to 100 kyr ago, it could have persisted for a substantial amount of time as exaptation, even as the neural mass continued to perform in the old ways. We have much less evidence than we would like that directly bears on the origin and spread of *Homo sapiens*. However, we do know that our species originated in this general time frame, probably in Africa. And we know as well that it quite rapidly spread Old World-wide from its center of origin, wherever that was.

Further, if at some point, say around 70 to 60 kyr ago, a cultural innovation occurred in one human population or another that activated a potential for symbolic cognitive processes that had resided in the human brain all along, we can readily explain the rapid spread of symbolic behaviors by a simple mechanism of cultural diffusion.

21. The presymbolic Neanderthals were able to build beautiful stone tools, and, as far as we know, communicate like many other mammals, but they do not appear to have possessed language as we know it (see Tattersall, 2002). The dramatic experience of Helen Keller in modern times may also throw some light on this transformation, although in her case she clearly had private representational capabilities—what she discovered was public symbolic communication.

In this book, I take as my starting point this informational space of individuals, properties, relations, and other entities, and study how we use language to talk about the world and do things in the world.

I now describe in broad terms my conception of meaning and language and how they fit into the larger scheme of things.

1.3   Information and Meaning

The breakthrough transformation described above allowed *reality*, that is, the world, to be construed as a space of entities. This space is what we call *information*. It contains individuals, properties, and relations; it also contains entities involving individuals *having* properties and *standing in* relations as well as collections of such states of affairs.

Intuitively and epistemologically, it is perhaps such collections that people first learn to discriminate and identify, chunks and slices of reality called *situations*. It is from situations, from these parts of the world that agents may observe or find themselves in, that they isolate individuals standing in relations.

An ancestor who had emerged into a symbolic awareness of the world may have noticed footprints in the snow or may have found himself fashioning a tool: both are situations the ancestor encountered and identified. Equally, modern man may read a report on a company or find himself in a restaurant: again, both circumstances are situations in our special sense of the term.

Beyond these rudimentary individuations, the ancestor may have realized that such footprints *meant* that a bear had passed by recently or that the hardness of the piece of stone he was using to fashion a tool *meant* that he could use it to chip away at various rocks. Similarly, the modern man in question may also have drawn the conclusion that the company report *meant* that its stock price was about to rise or that his being in a restaurant *meant* that he could order some food. Such observations point to another type of basic entity in our informational space: links between situations that allow one situation to carry information about another. This kind of link, called a *constraint*, is the essence of *meaning*. Put metaphorically, meaning is *the flow of information*.[22]

---

22. As I wrote above, "meaning" also refers to content or, in other words, to *information* itself. Thus, it can stand for both the *flow* of information from one situation to another and for the *information* that flows. It is useful to have this dual sense for our central concept.

Smoke means fire, black clouds mean rain, footprints in the snow mean that a bear has passed by there recently. The natural presymbolic order that exists prior to man is full of constraints that enable one part of the world to be systematically linked to another and to carry information about another.[23]

Of course, if there is no one around to observe these natural regularities, they remain undiscovered and unexploited. But it was an essential part of man's survival that he was able to register these constraints and choose his actions on their basis. Many such causal constraints were instinctively recognized by presymbolic man and even other animals and lower forms, but it was the ability to mentally represent and manipulate such systematic links and communicate them to others that enabled *Homo sapiens* to succeed so spectacularly.

Thus, the informational universe contains not just individuals standing in relations and situations but also constraints. My theory of this space is called *situation theory*, first invented by Barwise (1989b), who was influenced by Dretske's (1981) account of information flow, and who in turn was inspired by the classic theory of information transmission developed by Shannon (1949). The version of situation theory presented in this book is very much my own, though it draws a great deal from Barwise and Perry (1983) and Barwise (1989b).

Another way to describe the causal links that are part of the natural order is to say that smoke is a *sign* of fire, black clouds a sign of rain, and footprints a sign of a bear's presence. The term ''sign'' will be used to refer to constraints that do not involve human agency in a basic way. Once man broke into a symbolic consciousness, a new type of entity arose that I will call a *symbol*. Symbols are *artificial* constructs that involve human intention and agency in a basic way and that are at least partly social. Our modern man's company report is a collection of symbols. Symbols are organized in systematic ways and such structures are called *symbol systems*. The system of traffic lights is one example of a symbol system, but the major symbol systems are those of language. Once again, see Terrence Deacon (1997) for one account of the distinction.

The object of a theory of meaning should be symbols and symbol systems and how they are used by agents to bring about a flow of

---

23. All these constraints are causal and causality is therefore the basic cement of the universe. If causality implies regular succession of causally related events, then this statement would need to be modified to take account of the nondeterministic implications of quantum mechanics. See chapter 1 of Dretske (1981) for an illuminating discussion of the connection between causality and information.

information. For an entity to be a symbol, it must *stand for*, *be about*, *refer to*, or *represent* some other entity in the world and this relation must owe its existence ultimately to human intention and agency. Both signs and symbols involve "aboutness," but the requirement of human intention and agency is what distinguishes symbols from signs. The relation of a symbol to its referent is the relation of *reference* or *representation*.[24] This relation can be expressed as a constraint between two situations and thus enables one situation to carry information about another, the hallmark of meaning. A red light *means* you have to stop and a green light *means* you can go. The universe of symbols and symbol systems is very wide because it includes not only verbal languages, but also images, gestures, and other symbol systems. Peirce (1867–1913/1991) and Saussure (1916/1972) were perhaps the first figures to build explicit semiotic theories of this generality, but this kind of attempt has since fragmented into the separate study of each symbol system with little underlying unity.[25]

While the relation of reference or representation may be said to be the central aspect of meaning, it is also in some ways its most obvious attribute.[26] A less obvious aspect, indeed one that still eludes many, is the equally central relation of use. This is the aspect connected with the requirement of human intention and agency. It would perhaps not be an exaggeration to say that the subtlety of the relation of use is what makes semantics (or what many call pragmatics today) difficult. This is not to diminish the great strides that made referential semantics possible, but once the basic ideas of Tarskian model theory were in place, the rest has

24. Usually, the term "reference" is reserved for the relation between a symbol and an individual and "representation" is used more widely. I will use "reference" more or less interchangeably with "representation."

25. Interestingly, some semblance of unity prevails in the so-called Continental tradition in the work of figures such as Jakobson, Barthes, Derrida, and Eco because they all draw upon Saussure. See, for example, Colapietro (1993). While my book is entirely in the analytic tradition, readers more inclined to Continental semiology may also find it of interest as problems of common concern to both traditions are addressed. An explicit connection with the Continental tradition and Derrida is made in section 5.7.

26. Some philosophers deny this referentiality altogether while some others—like most Continental thinkers—assert that the relation of meaning is one between a symbol and a mental entity. I will simply take reference as the starting point of our study. This is not to deny that language has mental significance of course.

been a matter of working out details, however innovative they may be. A similar revolution has yet to occur in the domain of use, although part of the difficulty is that its subtlety makes many philosophers and linguists deny its importance and sometimes even its existence. The main reason for this skepticism is that there is as yet no systematic theory and mathematical apparatus to model use; this is a lacuna I hope to fill in this book in a compelling way.

In addition to the two central aspects of reference and use, a third equally fundamental attribute of natural language and many other symbol systems is the indeterminacy of meaning. Except for the relatively copious literature on vagueness, this property has also remained largely unexplored in its other dimensions. It is primarily this attribute of meaning that has allowed writers such as Derrida (1988) to make some amazingly outlandish claims about language and meaning. But if approached systematically and mathematically, this vital but amorphous attribute becomes easier to grasp and allows one to understand some rather commonsensical facts about language that have been ignored by many. It also makes clear why, along with the relation of use, this property of meaning is responsible for many of the difficulties faced by computational linguists.

Finally, the fourth entirely new feature of natural language that appears to have gone almost completely unnoticed is that of equilibrium. While Lewis (1969) was a precursor, Parikh (1987b, 2001) may have been the first to bring this aspect *squarely* into the realm of meaning. The *generative* idea in philosophy, linguistics, and artificial intelligence, the idea of starting with a stock of simple objects and combining them according to formal rules to derive more complex objects, was enormously fruitful, but perhaps too much has been attempted with this single idea. What semantics (and language more widely) needs in addition is the equally powerful idea of equilibrium. Essentially, equilibrium allows one to consider the *interactions* of objects at multiple levels, something that generativity precludes. Earlier, I considered *Language and Interaction* for the title of this book. There was a deliberate ambiguity in this title: ''interaction'' was meant to refer not only to the interactions between agents involved in the flow of information, but also to the interactions among various entities at multiple levels in the system of language and meaning.

Equilibrium semantics rests on the four fundamental ideas of reference, use, indeterminacy, and equilibrium because these features *inhere* in meaning; they are not imposed on it by the framework. The account I

construct is like any other empirical theory in the sciences; in addition, a science of meaning is a *social* science.[27]

Situated games of partial information play a central role in capturing all four of these ideas in a unified mathematical framework.

1.4   Language

The focus in this book is on language although the methods developed will also be applicable to other symbol systems. Language is possibly our most sophisticated symbol system and is certainly the most intricately structured. Meaning is also almost completely social: the relation between a word and its referent is in the main not fixed by a "natural" relation such as resemblance. The relation is, in a specific sense, *arbitrary*. "Table" could have meant "chair" and vice versa if English had evolved differently.[28]

I now discuss the four key ideas introduced above in some more detail.

1.4.1   Reference

The concept of reference came to be better appreciated and more precisely understood in modern times. Since this happened via the work of logicians such as Frege, Russell, and Tarski working with formal languages, and since these methods were then extended to natural languages, it seems best to start with a long quote from one of the more elegant modern texts on formal semantics by L. T. F. Gamut (1991a, 1991b).

The semantics of standard logic can be seen as a referential theory of meaning (and thus as a correspondence theory of meaning). When defining a model for predicate logic, the first thing we do is choose some set of entities[29] as our domain. The set is independent of the expressions which collectively form a predicate-logical language. We then specify a relation between the predicate-logical language in question and the domain. By means of an interpretation function, the constant symbols are assigned individual domain elements, and the predicate symbols are assigned sets of domain elements (or sets of ordered se-

27. It is a social science for two reasons: because both speaker and addressee are integrally involved in communication and meaning and because the entire society plays a part in contributing to meaning. The first point is discussed throughout the book, the second briefly in section 1.5 and section 7.1.2.

28. This arbitrariness of conventional meanings is dealt with partially in section 7.1.2.

29. Here "entity" stands for "individual" in our sense. For us, "entity" is used to refer to any object in the ontology introduced in section 1.3.

quences of n domain elements in the case of n-ary predicate letters) as their references. With this as a basis, we are in a position to define the reference relative to this model of all sentences in our language (that is, their truth-values), in the so-called truth definition.

The semantics of predicate logic is indifferent to the kinds of things we choose to put in the domains of our models. And whatever the domain may be, the theory of meaning is always a referential one: the meanings of the symbols are always their references.

One important characteristic of the semantic interpretation process, a characteristic which also happens to be shared by the nonstandard systems we shall meet up with, is that a strict parallelism is maintained between the syntactic constructions and their semantic interpretations. The truth definition mirrors the syntactic definition of the formulas of the language in question. There is a methodological consideration underlying this practice, one which can be traced back to Frege. This German logician and mathematician gave the first satisfactory analysis of sentences with relational predicates and multiple quantification in 1879, in his *Begriffsschrift*. Now the fundamental insight behind his solution to these age-old problems is that every sentence, no matter how complex, is the result of a systematic syntactic construction process which builds it up step by step, and in which every step can receive a semantic interpretation. This is the well-known *principle of semantic compositionality*.

This extract explains clearly how reference is conceptualized and set up for formal languages. The framework of formal semantics for natural language has largely taken over this conceptualization and added to it more complex entities to handle the more complex devices of natural language. Montague Grammar and its derivatives such as Discourse Representation Theory[30] represent in some sense the pinnacle of this approach to meaning.

But some aspects of the underlying conceptualization that formal semantics shares with the semantics of predicate logic are problematic for natural language for the following reasons:

· Restriction of the domain to individuals
· Holism of truth values
· Reference as assignment
· Compositionality
· Extensionality and intensionality

As we have already seen in section 1.3, and will see in greater detail in the next chapter, there are a plurality of entities in the informational space. It is this significantly richer space that will be seen to be required for the semantics of natural language because natural language is much

30. See Kamp and Reyle (1993).

richer than (first-order) formal languages. Individuals and sets of (sequences of) individuals may be adequate for predicate logic but they are far too impoverished to handle the complexities of natural language. Some of these inadequacies have been addressed in formal semantics by bringing in properties and relations as entities in their own right (as opposed to modeling them as sets), but what is needed is a thoroughgoing revision of the ontology. This is provided by situation theory, both in its original form and especially in the version presented here.

The second assumption of the holism of truth values is a particular instance of the previous assumption. Instead of supplying appropriate structured entities to play the role of the contents of utterances, formal semantics and philosophy have continued to deal with truth values as their (referential) "meanings." Barwise and Perry (1975, 1983) have criticized this holism in very thorough ways and there is little point in repeating this criticism here. Unfortunately, situation theory and situation semantics have fallen out of favor[31] and so their solution to the problems posed by this holism have been largely ignored.

Third, reference has been treated simply as *assignment*, a move that is perfectly legitimate for formal languages, but leaves much to be desired for natural languages. This is one reason for the split in the study of meaning: formal semanticists have contented themselves with simply addressing the problem of representing meanings and have left the messy facts of use that lie at the core of reference to pragmatics and the philosophy of language. Unfortunately, by and large, these latter disciplines have simply replaced assignment by *convention*, that is, (literal) contents

---

31. Partee (2005) identifies two reasons why this may have happened: Barwise avoided the notion of possible situations and Barwise and Perry never gave a satisfactory account of quantified noun phrases such as "every student." While this may be true from the viewpoint of the Montague grammar community, I believe the reason for their falling out of favor is more fundamental, one that plagues even the formal semanticists: as I said above in section 1.1, they had no *theory* of use, and so failed to deliver the breakthrough of incorporating Austin into model theory bruited by situation semantics. They could not develop such a theory partly because of fundamental gaps in situation theory, something I rectify in chapter 2, and partly because they had no access to the ideas of game theory. This led to an inability to carry the entire program forward, not just to a lack of solutions to the particular problems posed by possibility and quantifiers. I address noun phrases in chapter 6 and I have no objection to possible situations when required as long as they do not lead to a modal realism. Indeed, they arise in a natural way in the situation-theoretic construction of games described in chapter 3 and sections A.2 and A.4 of Appendix A.

are generally taken as conventionally given. What is required is both an adequate ontology *and* an account of reference in its full complexity that meshes with these representations. I attempt to do this via a combination of game theory and situation theory that allows one to actually construct a formal definition of reference.[32]

The fourth shared assumption is Frege's venerable *principle of compositionality*.[33] Again, this is perfectly valid for formal languages because it is always assumed that such languages are perfectly precise and unambiguous. But natural languages are notoriously ambiguous and vague and then Fregean compositionality breaks down. This is because the meaning of one word (and phrase) in an utterance of a sentence can affect and be affected by the meanings of the other words (and phrases) in the sentence.[34] When no ambiguity or vagueness are present, these interdependencies and interactions of meaning are otiose and superfluous. But when they are present, as they almost always are in natural language, the simple generative idea of Fregean compositionality falls short. As I said earlier, generativity prohibits the *interactions* of various objects. Equilibrium semantics offers a generalization of the *Fregean principle of compositionality*, called the *fixed point principle* or *fixed point compositionality*, that is able to accommodate these pervasive attributes of natural language; when they are absent, the *fixed point principle* reduces to the special case of Fregean compositionality.

Finally, Frege's (1980) classic paper appeared to make it clear that reference could not be *direct*, that there had to be some intervening layer such as that of his "sense." These issues have been hotly contested after Kripke's (1972/1980) dramatic work on direct reference for a subclass of words. I will side with Frege and offer a picture of word meaning that is a generalization and refinement of the traditional observation that the ordinary word "meaning" is ambiguous and needs to be split into two parts

32. I do not actually give this definition of reference in the book because it is messier than similar definitions. It follows the general line of definitions discussed in chapter 6 of Parikh (2001) with appropriate modifications to accommodate the more general framework for meaning presented here. It is analogous to Definition 2 in section 5.10 of this book.

33. Frege himself may or may not have put forward this principle, but it is widely known as such and I will stick to this appellation.

34. In an utterance of "The waiter is rude," all four words are ambiguous as will be seen in chapter 6, and their intended meanings affect and reinforce one another. Indeed, it will be seen in section 7.1.1 that meaning and grammar are also interdependent.

that have variously been called connotation and denotation, sense and reference, and intension and extension. I call these two tiers conventional meaning and referential meaning.[35] In this book, I do not offer any argument for this position, as it would mean a long detour and require addressing a large literature including, most prominently, Kripke (1972/1980) himself. I hope to do this on another occasion but some inkling of my views may be gleaned from chapter 6 on noun phrases where I briefly address proper names and where I offer detailed counterarguments to Kripke's (1977) critique of Donnellan (1966) based on my theory of definite descriptions. In any case, I hope the theory will be immune to the kinds of criticisms Kripke (1972/1980) and Putnam (1975) (among many others) have made. A somewhat surprising consequence of my account is that while every word in a natural language has at least one conventional meaning, phrases and sentences have no conventional meanings; on the other hand, when used in an utterance, words, phrases, and sentences all receive referential meanings or contents. Each word in an utterance of a sentence such as "The waiter is rude" has at least one conventional meaning, but the various phrases and the entire sentence have no conventional meaning. On the other hand, all words, phrases, and sentences acquire contents when uttered. This again is a failure of compositionality at the level of sense or intension or conventional meaning.

Despite these departures from the many assumptions shared by the standard semantics of formal and natural languages, equilibrium semantics does share its foundational assumption that language requires a referential theory of meaning. Indeed, my account requires that every word, phrase, and sentence in an utterance have a reference, even apparently syncategorematic words such as THE and OR.

1.4.2 Use
There are at least two reasons why the concept of use is subtle and has resisted analysis. One is that it involves a number of other concepts that are often poorly understood. The other is that it is difficult to develop a mathematical apparatus that can accommodate all these concepts and that does justice to interactions between agents. With formal languages, it is possible to abstract from use and pretend we are dealing just with inert symbols rather than with their use. With natural languages, this

---

35. When the context is clear, I will use just the term "meaning" instead of the longer "referential meaning." This sense of "meaning" is identical with that of "content."

becomes impossible as we will see below. Those who have tried to ignore the relation of use as central to semantics have had to resort to many awkward contortions such as positing all kinds of entities at multiple layers of sentential representation, whose connection with empirical reality becomes increasingly tenuous and ad hoc, reminiscent of the epicycles of pre-Copernican astronomy.[36]

   Minimally, the following concepts are intimately related to, if not included in, the concept of use:

- Belief, desire, intention, and agency
- Sentence and utterance
- The situatedness of language
- The efficiency of language
- Ambiguity
- Communication and information flow

36. A single example should suffice: Recanati (2004b) cogently confutes the "representationalist" analysis of an utterance of "It is raining." I quote:

John Perry (1986) and many others after him have argued as follows. Even though nothing in the sentence "It is raining" stands for a place, nevertheless it does not express a complete proposition unless a place is contextually provided. The verb "to rain," Perry says, denotes a dyadic relation—a relation between times and places. In a given place, it doesn't just rain or not, it rains at some times while not raining at others; similarly, at a given time, it rains in some places while not raining in others. To evaluate a statement of rain as true or false, Perry says, we need both a time and a place. Since the statement "It is raining" explicitly gives us only the two-place relation (supplied by the verb) and the temporal argument (indexically supplied by the present tense), the relevant locational argument must be contextually supplied for the utterance to express a complete proposition. If Perry is right, the contextual provision of the place concerned by the rain is an instance of saturation like the assignment of a contextual value to the present tense: both the place and the time are constituents of what is said, even though, unlike the time, the place remains unarticulated in surface syntax.

   But is Perry right? If really the contextual provision of a place was mandatory, hence an instance of saturation, *every* token of "It is raining" would be unevaluable unless a place were contextually specified. Yet I have no difficulty imagining a counterexample, that is, a context in which "It is raining" is evaluable even though no particular place is contextually singled out.

   Then Recanati (2004b, 9–10) goes on to provide an irrefutable counterexample and thereby argue against the representationalist view. I should add that the main thrust of Perry's (1986) paper is antirepresentationalist, so even his failing to avoid this tendency is especially revealing of the pervasiveness of the mainstream view that can ultimately be traced back to logicism.

It is astonishing that mainstream linguistics in the twenty-first century has no theoretically grounded conception of agency.[37] Whatever our innate endowment may be, language (its dimension of meaning in particular) is surely a social institution and as such supervenes on use and human agency. Ever since Grice, the philosophy of language has had recourse to the concept of rational agency, but it has remained informal. The only framework today that has an apparatus with a mathematically formulated and philosophically sound conception of agency is that of game and decision theory. This conception has undergone exciting changes since the work of Tversky and Kahneman (Kahneman, Slovic, and Tversky 1982) and is still evolving. In this sense, my earlier work (Parikh 1987b, 1992, 2000, 2001; Fehling and Parikh 1991) and others' subsequent contributions (e.g., the volume edited by Benz et al. 2006 and the volume edited by Pietarinen 2007) to the now burgeoning field of game-theoretic semantics and pragmatics have been the only approaches that involve the concept of agency in a full-blooded way. Belief, desire, and intention are integral to action and it is the singular virtue of decision and game theory that they offer a way to integrate these component factors that result in action. Any systematic approach to questions of use must draw upon a theory of action and especially interaction that includes its constituents of belief, desire, and intention. Indeed, we will see in the paragraphs that follow that all the other elements of use listed above can be addressed adequately only because we have recourse to game theory and situation theory.

Perhaps equally astonishing is the insistence of many philosophers and linguists on dealing with sentences rather than utterances despite the contributions of ordinary language philosophy. A sign or symbol seldom carries information by itself. It is only when we take account of the circumstances in which the sign or symbol occurs that we can infer a referential meaning.[38] Likewise, a sentence by itself does not have a meaning. It is only an utterance, an act involving the production of a sentence (or other symbol) in some situation, that carries information. While words and sentences appear to mean things in the abstract, only a moment's reflection is required to see that a name such as HARVEY or common noun

---

37. Even Optimality Theory does not, just as it does not have many of the other concepts listed above. See Blutner and Zeevat (2004).

38. This observation is in fact a little subtler for images and similar symbols than for words.

such as BANK can only carry their conventional meanings[39] when abstracted from their circumstances of use. Without an embedding utterance, a name cannot possibly refer (which one of countless Harveys is being referred to?) nor can a noun, verb, preposition, or article.[40] Identifying utterances with sentences may be permissible only for formal languages. The *circumstances* of utterance simply cannot be ignored in the case of natural language. Put differently, a sentence and its component parts cannot ever connect with reality, cannot ever be about anything without being ensconced in an utterance. I suspect the reason for the reluctance to deal squarely with utterances is that, as hinted above, there simply appears to be no mathematical or otherwise solid apparatus to deal with the messiness and unruliness of *contexts*. Sentences are well-behaved, rule-governed objects and so we feel more comfortable with them and have ways of manipulating them. I hope that equilibrium semantics, with its use of game theory and situation theory, will dispel these doubts.

I have already referred to the situatedness of language when talking about contexts and circumstances, just alternative words for what will technically be called the *utterance situation*. Indeed, agents are always in situations of one sort of another, and not just our utterances, but all our actions as well as their constituents—beliefs, desires, and intentions—are situated. In the case of utterances, as with all actions, this situatedness implies that meaning is a result of both the sentence uttered and the utterance situation. In fact, as already mentioned earlier, we will write $\mathcal{C}(\varphi, u) = \mathcal{C}_u(\varphi)$ for the content of a sentence φ uttered in situation u. The context makes many contributions to the meaning of an utterance in general unlike the case of formal languages where one can effectively write $\mathcal{C}_u(\varphi) \equiv \mathcal{C}(\varphi)$ for all situations u. This situatedness occurs more widely than just with natural language. When someone waves his hand to extend a greeting, there is a situation in which he performs the action and thereby conveys a greeting. In a different situation, the same action could have meant a goodbye.

Intimately related to this situatedness is the efficiency of language. The fact that the same sentence can be used to convey different contents in different circumstances is precisely what makes language efficient. Partly,

39. That is, the suitably indexicalized properties and relations corresponding to what one might find in a dictionary for a common noun as explained in section 7.1.2; the case of names is a little different and is dealt with in sections 3.2 and 6.3.

40. In the extended sense of reference I am using.

efficiency is purely linguistic, that is, it resides in the lexical ambiguity of words and structural ambiguity of sentences, thereby allowing one word or one sentence to carry multiple semantic values via linguistic mechanisms, but by and large it is contextual. Context is central to efficiency as it provides the ambient information that gets *added* to the purely linguistic information a sentence provides to produce meaning. In these two ways, efficiency is an entirely observable and empirical fact about language and meaning. However, when we begin to build theories of language and meaning, then another dimension of efficiency enters the picture. *Agency* is central to contexts and agency itself is *more or less* efficient because it is more or less *rational*. While few agents (including human beings) are always perfectly rational in ways sometimes required by game and decision theory, they are not entirely irrational either. In any case, whatever our variable degree of rationality, it is arguable that it lies at the center of the efficiency of language, not only in its contextual aspect, but even in its purely linguistic aspect. Indeed, it is at the center of all human institutions, let alone just the institutions of language and meaning. (That is why game and decision theory have such wide applicability.) Until it is known better and more empirically how we make choices, one aspect of efficiency then is that it provides a way of *idealizing* language and meaning via the idea of rationality, just as earlier the predicate calculus had provided a way to regiment the messiness of natural language. However, this time it is not language itself that is being idealized but its use. Moreover, because game and decision theory are increasingly open to all kinds of *behavioral* modes of choice, not just perfectly rational ones, the degree of idealization involved can be "tuned" for greater or lesser realism in our theory-building. These two dimensions of efficiency, one purely empirical (involving both language and context), the other more or less theoretical (involving agency), are both what make semantics difficult and what make language an extraordinarily rich symbol system.

I have alluded to ambiguity in language a few times in the foregoing. I see ambiguity in an extended sense as an essential part of language. Generally, linguists and philosophers think of ambiguity as being either lexical or structural; I use the term very widely to cover any and all cases where more than one possible interpretation of an utterance may exist. This includes not just lexical and structural ambiguity, but all those cases that result from the efficiency of language. It turns out, as shown especially in Parikh (2001, 2006b) and Clark and Parikh (2007), that more or less the same game-theoretic models can be applied to disambiguating all

multiple possible interpretations. It is because of this uniformity of application that I feel that ''ambiguity'' should be broadened to encompass all these varied cases.[41] In section 7.1.1, the notion of ambiguity will be generalized even further.

The essence of use lies in communication and information flow between agents. Grice (1957, 1969) was the first person to introduce and explore the related concept of nonnatural meaning. Strawson (1964) and Schiffer (1972) refined Grice's attempts at definition. Parikh (2001) expanded this exploration into an infinite lattice of concepts of communication and information flow as well as related infinite lattices of concepts associated with speaker meaning and addressee interpretation, and gave rigorous game-theoretic definitions of these concepts. We will see in section 5.10 that these game-theoretic definitions need to be generalized to account fully for the indeterminacy of meaning.

The six interlocking notions above are all integral parts of the concept of use. It is the complexity of these constituent notions that makes it difficult to incorporate the concept of use into semantics in a smooth way and this is why it has been consigned to the ''dustbin of pragmatics.'' It is fortunate that the situated game theory I will present is ideally suited to tackle each of these component notions in a thorough way that allows a seamless unification of semantics and pragmatics.

To summarize, by and large, meaning involves utterances as actions and is situated. Different situations with the same signs or symbols typically carry different information. In one situation, smoke may mean fire, in another, a cigarette smoker. Or, as remarked above, a hand wave may mean hello or goodbye. To paraphrase Wittgenstein (1953/1968), an utterance of ''This chair is brown'' could mean a variety of things in different circumstances. This situatedness of meaning goes hand-in-hand with the efficiency of language whose flip side is a pervasive ambiguity in signs and symbols, particularly those of language. All of these together make communication and the flow of information between agents possible.

1.4.3 Indeterminacy

A consequence of the fact that meaning depends on use is that it may be indeterminate in a number of ways.[42] Three of these are:

41. This extended applicability of my games of partial information has sometimes been misunderstood as applying only to the narrower, traditional concept of ambiguity.

42. An extended discussion of this feature of meaning can be found in Parikh (2001, 2006b).

1. Content is not always fully intended.
2. Content is not always deterministically given.
3. Content is not always the same for speaker and addressee.

Overwhelmingly in the literature of the last century, both in linguistics and in the philosophy of language, the negations of these statements have been the shared beliefs. Even Wittgenstein, Austin, and Grice, three philosophers who were very sensitive to aspects of the use of language, and their many followers, appear to have made these assumptions.

At least part of the meaning of an utterance is explicitly intended by speakers because they are engaged in purposive activity. Indeed, there are corresponding intentions on the side of the hearer as well.[43] But intentions are themselves *situated* so that speakers do not need to explicitly intend everything. For example, did Dostoevsky explicitly intend the entire literal and implicated content of *Crime and Punishment*? It is perhaps even physically impossible to do so because, even minimally, there would be so much implicit information being conveyed. The way to think about content is to allow for some of it not to be explicitly intended and to inquire if the speaker might assent to the unintended part of the content. This will be shown in detail once our game-theoretic models are in place. However, since in practice we are seldom in a position to ask such questions of speakers and addressees about their utterances and interpretations, we have to reconcile ourselves to saying that content is partly indeterminate in this sense, since it may have partly not been explicitly intended and we cannot know whether it has been correctly inferred. Of course, since intentions are invisible, it is not possible to be sure even about the intended parts of the content. Perhaps Grice (1989), who introduced intention-based derivations of content, expected too much to be resolved by intentions.

Grice was also probably the first philosopher to allow that content may be partly indeterminate in a direct sense, but he restricted this indeterminacy just to implicature, assuming that literal content was always determinately given in communication. It is natural to extend this type of indeterminacy to all aspects of meaning. The principal insight here is that linguistic communication and information flow can be *probabilistic*, something that does not seem to have been noted before at this general level.[44]

---

43. See Strawson (1964) and Parikh (2001).

44. Except in Parikh (2001, 2006a, 2006b). More recently, probabilistic approaches to particular phenomena such as conditionals have appeared in Kaufmann (2005). See Cohen (2003) as well. Probabilistic methods have, of course, been common in computational linguistics.

Probabilistic communication is important in the determination of both literal content and implicature as well as illocutionary force. Once we take note of this fact, we begin to see it everywhere, even in the simplest information flows. This makes the need for probabilistic approaches to interpretation such as that used in game and decision theory even more apparent. That is, all aspects of content may be probabilistic and indeterminate, and an addressee may not infer just propositions from an utterance, but also the probabilities with which they are being conveyed.

Finally, in practically all the relevant literature, it has been assumed that communication and information flow consist in just an identity between a proposition conveyed and a proposition grasped. Closer examination of this assumption shows that this is the simplest case, and in a sense that will become clear later, it is an ideal case that is seldom realized in practice. Essentially, contents will usually be different for different participants in a communicative interaction. This fact introduces yet another aspect of indeterminacy into meaning.

Other sources of indeterminacy, some of a more technical kind, will be introduced in chapter 5.

1.4.4 Equilibrium

Our final and central unifying idea and image is that of equilibrium. This idea has a long history: it goes back to Newton's mechanics, and then, especially via Adam Smith in the eighteenth century and the mathematician Cournot in the nineteenth, finds its way into economics and later into game theory in the twentieth. Today, the idea of equilibrium occurs in practically every natural and social science. Indeed, it is arguable that Saussure (1916/1972) had an inkling of this notion in his idea of phonology (and semiology, generally) as a *system of differences*. While I initially saw equilibrium in language and meaning via the application of game theory, it is possible to abstract from the game theory and see equilibrium as an inherent and empirical part of this system.

Equilibrium simply means balance among multiple interacting elements. In each context where the idea of equilibrium is applicable, we have to determine what these elements are and what the nature of their interactions is. Once we have these, all that remains is to describe the conditions under which these multiple interactions "balance." In some contexts, the only observable states of a system are those in equilibrium; in other contexts, disequilibrium may also be observable. Equilibrium may also be dynamic and evolve over time.

For language, equilibrium enters essentially through the element of *choice*: the speaker must choose his utterance and the addressee must choose her interpretation and these choices must be in *balance*. As we will see, the speaker and addressee participate in multiple games at multiple levels in a single utterance—at the level of words, phrases, and the sentence itself—and so there are multiple equilibria that occur in communication. Not only does each equilibrium involve a balance among the choices and strategies available to the speaker and addressee in each game, but the multiple equilibria are themselves in balance: an equilibrium of equilibria!

Saussure (1916/1972) viewed language as a system of opposing elements that underlay a process of choice. "What idea or phonological material there is in a sign matters less than what there is around it in the other signs."[45] Saussure's position seems to be that a sign has significance not simply because of the association between sound and meaning but because this association is embedded in a larger system of sound-meaning associations.[46] The significance of this notion becomes apparent in the context of making a linguistic choice; it is worth quoting Saussure (1916/1972) at length:

Our memory holds in reserve all kinds of more or less complex phrases, regardless of their type or length so that, when we employ them, associated groupings can be called upon to fix our choice. When someone says *marchons!* ("Let's walk!"), he thinks unconsciously of a divers group of associations in the midst of which the phrase *marchons!* finds itself. On the one hand, it is part of the series *marche!* ("Walk!" [familiar]) *marchez!* ("Walk!" [formal, plural]), and it is the opposition of *marchons!* with these forms that determines the choice; on the other hand, *marchons!* evokes the series *montons!* ("let's go up/get aboard!") *mangeons!* ("let's eat!") among which it is chosen by the same process. For each series, we know what must be varied to obtain the correct contrast with the desired unit. If we change the idea to be expressed, then other oppositions will be needed to make the correct value appear; we say, for example *marchez!* or perhaps *montons!* ...

This principle applies to phrases and to sentences of all kinds, even the most complex. At the point where we say the sentence «que *vous* dit-il?» ("What does he say to you?"), we vary an element in a latent form, for example: «que *te* dit-il?» ("What does he say to you [familiar]?")—«que *nous* dit-il?» ("What does he say to us?"), etc., and it is in this way that our choice is fixed upon the pronoun *vous*. So in this operation, which consists in mentally eliminating anything that

45. "*Ce qu'il y a d'idée ou de matière phonique dans un signe importe moins que ce qu'il y a autour de lui dans les autres signes.*" (Chapter VI, section 4) Note that all translations in this section are by Robin Clark.

46. Thus, changing one part of a system should ramify throughout the system, causing all the other associations to change as well.

does not lead to the desired differentiation at the desired point, associative group-ings and phrasal patterns are both in play.[47]

Accordingly, Saussure's insight is that a language is defined by a system of choices that can bear meaning, and the fact of choosing one sign in-stead of another is, in itself, a critical act that is an inherent part of the linguistic system. The content of this insight has been far from clear, how-ever, and most theoretical linguistics has shied away from using an action such as choice as part of its theoretical machinery in favor of the study of linguistic representations, a static object that seems more amenable to for-mal investigation.

Philosophers influenced by the ordinary language tradition, on the other hand, have always had recourse to the idea of an *utterance* and therefore of an *action* and even implicitly of an *interaction*. While Grice (1989) may have been the first to introduce the idea of interaction into the philosophy of language, it remained relatively implicit in his analysis of speaker meaning. Strawson (1964) was perhaps the first to explicitly consider the interaction between a speaker and addressee by noting that there were reciprocal intentions and actions on the addressee's side in communication, but his insight has been ignored in subsequent work by philosophers and linguists. By and large, the addressee remains a ghost in mainstream semantics, at most a passive recipient of the speaker's actions.

---

47. "*Notre mémoire tient en réserve tous les types de syntagmes plus ou moins com-plexes, de quelque espèce ou étendue qu'ils puissent être, et au moment de les em-ployer, nous faisons intervenir les groupes associatifs pour fixer notre choix. Quand quelqu'un dit* marchons!*, il pense inconsciemment à divers groupes d'associations à l'intersection desquels se trouve le syntagme* marchons! *Celui-ci figure d'une part dans la s'erie* marche! marchez!*, et c'est l'opposition de* marchons! *avec ces formes qui détermine le choix; d'autre part,* marchons! *évoque la série* montons! man-geons! *etc., au sein de laquelle il est choisi par le même procédé; dans chaque série on sait ce qu'il faut faire varier pour obtenis la différenciation propre à l'unité cherchée. Qu'on change l'idée à exprimer, et d'autres oppositions seront nécessaires pour faire apparaître une autre valeur; on dira par exemple* marchez!*, ou bien* montons!. . .*

"*Ce principe s'applique au syntagmes et aux phrases de tous les types, mêmes les plus complexes. Au moment où nous pronon cons la phrase:* «que vous *dit-il?»*, nous faisons varier un élément dans un type syntagmatique latent, par exemple* «que te *dit-il?»—«que nous *dit-il?»*, etc., et c'est par là que notre choix se fixe sur le pronom* vous. *Ainsi dans cette opération, qui consiste à éliminer mentalement tout ce qui n'amène pas la différenciation voulue sur le point voulu, les groupements associatifs et les types syntagmatiques sont tous deux en jeu.*" (Chapter VI, section 2)

Further, except for the singular work of David Lewis (1969), neither linguists nor philosophers have quite conceived of utterances and interactions within a system of choices. Unfortunately, Lewis thought of game theory as mere "scaffolding" that was ultimately dispensable.[48] More importantly, he conceived of speaker meaning as entirely conventional because his primary aim was to analyze the concept of convention and he believed Grice's concept of nonnatural or speaker meaning fell within his concept of convention. The particular types of game models he considered, based on the work of Schelling (1960), were also fairly basic and lacked the complexity required for modeling communication that was ambiguous and contextual and costly. (The first work to develop appropriate game models was Parikh 1987b.)[49]

Thus, Saussure considered choice but not interaction between speakers and addressees, and philosophers with the exception of Lewis considered action and interaction but not choice. Both these elements of choice in the context of strategic interaction need to be brought together in a way that results in an equilibrium, a composite idea that neither component idea by itself suggests. To the best of my knowledge, the only mathematical framework that does this *adequately* is game theory, particularly situated games of partial information, the kinds of games invented in Parikh (1987b) and developed throughout this book. The kind of system that emerges from taking choice and strategic interaction as fundamental

---

48. To quote from Lewis's (1969) introduction:

My theory of convention had its source in the theory of games of pure coordination—a neglected branch of the general theory of games of von Neumann and Morgenstern, very different in method and content from their successful and better known theory of games of pure conflict. Coordination games have been studied by Thomas C. Schelling and it is he who supplied me with the makings of an answer to Quine and White.

Yet, in the end, the theory of games is scaffolding. I can restate my analysis of convention without it. The result is a theory along the lines of Hume's, in his discussion of the origin of justice and property.

49. The key insight required was the modeling of communication by representing the relevant interactions and choices via games in so-called *extensive* form as initially developed by Kuhn (1953) and as elaborated by Kreps and Wilson (1982) for games of incomplete information. The extensive form enabled making explicit certain crucial information involving ambiguity, something that is obscured in the so-called *strategic* form representations Lewis used. These ideas are introduced in chapter 3 and in Appendix A. Lewis also never considered the costs of communication.

properties of linguistic and communication systems is rather different from what has emerged in the past.

To take just one important example, the approach taken here is very different from the approach that Montague (1974b) took in "The Proper Treatment of Quantification in Ordinary English." There, Montague gave a method of translating an interesting fragment of English into the language of Intensional Logic (IL). Since the interpretation of IL relative to a model is a straightforward affair, English could then be interpreted indirectly by piggybacking on IL. Montague's accomplishment is an impressive one, standing as the culmination of the "logicist" approach that has its origins in the work of Frege and Russell. This type of approach, which remains the dominant paradigm in (linguistic) semantics, maps surface linguistic forms into an abstract logical representation, Logical Form. Logical Form is a syntactic level of representation that shares many properties with an artificial logical language such as First Order Logic. This abstract language can then be interpreted relative to a model.

However, unlike an artificial language—such as standard First Order Logic or a computer programming language, both of which can be interpreted directly—Logical Form necessarily contains elements that can only be interpreted relative to a context. Thus, Logical Form requires a pragmatic component that will fill in details from the context and yield the content of the utterance. Thus, these theories give the linguist (and philosopher) *two* degrees of freedom: first, there is the unseen level of Logical Form, which can be quite remote from observable sentences;[50] second, the fact that one then maps Logical Forms to contents means that Logical Form can be remote from intended contents as well. While this situation is not, in itself, a fatal one, these two degrees of freedom with an unobservable and relatively remote Logical Form in the center do mean that theories of this sort will be relatively unconstrained by empirical factors. Any evidence contradicting such a theory could, in principle, be fixed by ad hoc adjustments to the intermediate layer of Logical Form. It seems therefore that conceptually simpler theories responsive to the empirical facts of language use might be preferred.[51]

50. Notice that surface syntax is already quite remote from what we can observe directly since it involves abstract objects such as constituents and grammatical categories. Logical Form adds another level of unobservable structure.

51. It appears that the neuroscientific evidence also indicates that choices are actively considered in the brains of interlocutors during communication (see Glimcher, 2004, and more recent books on neuroeconomics); interactions

The approach developed here shares some properties with the analysis found in Montague's (1974a) ''English as a Formal Language.'' There, he tried to interpret a fragment of English directly relative to a model. However, the objects he chose to translate were sentences and their parts, not utterances. My approach is similar though I connect *utterances* directly to model-theoretic structures, that is, to situation theory, and give a method for computing the interpretation of complex expressions and their parts in a non-Fregean way via the *fixed point principle*. As will be shown, because my objects of analysis are utterances and not sentences, and because utterances involve both choice and strategic interaction, game theory plays a fundamental role in this computation.

I am thus advancing an empirical hypothesis about choice, strategic interaction, and equilibrium, and the nature of linguistic meaning and communication as well as developing a set of mathematical tools for linguistic and philosophical exploration. A consequence of the analysis will be that there is no principled distinction between semantics and pragmatics. There are currently many views about the relationship between semantics and pragmatics, but they all appear to share the view that certain things are inherently semantic—for example, conventional meanings—and other things are inherently pragmatic—for example, the role of context. This starting point originated perhaps in Grice's (1975) pioneering work on implicature, but it has now become almost an a priori commitment for researchers.

My suggestion is that it is worthwhile to start afresh from first principles and see what kind of framework emerges before we draw a line between the two subfields, if any such line needs to be drawn. In the end, it is the techniques and analyses of natural language meaning that are interesting, not some putatively inherent distinction between subfields. Game theory and equilibrium semantics allow us to look at ''the problem of meaning'' in a new way.

To return to the main theme of equilibrium, the details of this idea in the framework will become clear as we develop it. It incorporates the other three principal features we have discussed: reference, use, and indeterminacy. Below I describe a global view of the idea.

Language and its use may be viewed as a system of constraints in the specific sense of ''constraint'' introduced earlier. I will not bother to con-

---

among agents are of course obviously observable. So the elements underlying equilibrium—choice and strategic interaction—appear to have empirical validation.

vert a relationship that obtains into a formal constraint in this sense, but when this word is used, my intention is that the relationship can be so expressed. In other words, equilibrium semantics will itself be expressible as a system of constraints within situation theory.

My concern in this book will be with the following four sets of constraints:

1. Syntactic constraint (**S**)
2. Conventional constraint (**C**)
3. Informational constraint (**I**)
4. Flow constraint (**F**)

Equilibrium semantics may be described compactly via its system of constraints **SCIF**. When speech is considered, a fifth set of phonological constraints **P** would have to be added. This last constraint will be addressed very briefly in chapters 5, 6, and 7.

S contains some account of the syntax of the language being considered. I will not adopt any particular approach to grammar, but will use this constraint informally as it is not my chief concern. Syntax interacts with and is influenced by meaning, that is, the other three constraints in **SCIF**, but we will simply take it as given, except in chapter 7. **S** plays a critical role in the derivation of content. As an aside, we observe that it is primarily with respect to this constraint that different symbol systems differ: language is the symbol system with the most elaborate and determinate syntax.

C is a set of conventional constraints that maps every word into one or more properties or relations. This map is called the *conventional map* and can be largely extracted from a dictionary,[52] except for a relatively small class of syncategorematic words such as determiners and conjunctions. The conventionality of the meanings implies that they are independent of context.[53] Again, to a lesser degree, different symbol systems differ with respect to this constraint as well: some symbol systems such as language are more or less fully conventional whereas others such as images may be partly naturalistic. The exact role and place of conventional meanings in communication will be examined from chapter 3 onward.

---

52. With appropriate modifications to take account of the kinds of criticisms Kripke (1972/1980) and Putnam (1975) have made. See section 7.1.2.

53. This independence is actually partial in a sense that is made precise in section 7.1.2.

I maps the properties and relations obtained from the conventional map into certain special situation-theoretic objects introduced in the next chapter. Which particular objects they get mapped into is in part determined by **S** and in part by our informational space or ontology relative to a context or utterance situation u. This map is called the *informational map*. The general form of this constraint is by and large the same for all symbol systems, but as **S** influences its behavior, its details vary from system to system.

Finally, **F** is in some sense the main constraint, the one that embodies much of the framework of equilibrium semantics. Essentially, a system of situated games provides a model of the utterance situation u, so that together with the sentence and its phrase structure, we can infer its meaning. Again, this constraint works in the same way for different symbol systems at a general level, but differs in particulars.

At this stage, this very brief description of **SCIF** will necessarily be rather abstract. The rest of the book will spell it out.[54]

The general idea of equilibrium in equilibrium semantics is that all four sets of constraints are in equilibrium—within each constraint and across constraints, both in the context of the system of meaning and grammar and in the context of utterances. In this book, we explore the central part of this ideal conception.

1.5   The Scope of Game Theory in Language

As explained in Parikh (2007), there are broadly two levels at which game theory can address the problems posed by language, a situational level and a structural level. At the situational level, the interest is in solving the problem of content in situated communication. At the structural level, the interest is in solving the problem of how the various structures— conventional meaning, semantical rules, linguistic variation—emerge to enable communication. Of course, both these levels coexist and codetermine each other, as discussed in section 7.1.2, but this book will try to offer a more or less complete account at the situational level since this is the central problem of semantics in philosophy and linguistics. I hope to address the structural level and also relate the two levels elsewhere. In analogy with the field of economics where the main division is between microeconomics and macroeconomics, these two levels could well be called *microsemantics* and *macrosemantics* since the first deals with com-

---

54. See Parikh and Clark (2006, 2007) for a short introduction to this framework.

munication between and among individuals and the second deals with attributes of language that emerge in entire populations.

As I have argued in this chapter, a new framework that reorganizes and reconceptualizes semantics and pragmatics is badly required. The mainstream consensus has reached an impasse. The account of equilibrium semantics does *three* things: it provides a new theory, a new framework, and a new paradigm for language and meaning. It would therefore be a fatal mistake to view the game-theoretic and related apparatus developed in *Language and Equilibrium* as simply grafting a piece of mathematical machinery onto a mainstream, largely Grice-inspired framework in order to formalize relatively informal areas of the field such as implicature or lexical disambiguation. It is a great deal more than this: I am seeking a radical reframing of the problem of meaning and, indeed, am promising a whole new way to think about it.

In contrast, much other recent game-theoretic work primarily in the field of linguistics, especially by Arthur Merin, Robert van Rooij, Gerhard Jäger, Anton Benz, and others,[55] while notable in its own right, has generally taken the mainstream view of semantics and pragmatics as *given* and so works primarily within what has come to be called *formal pragmatics* as there is relatively little that can be done within the sphere of literal meaning if it is assumed to be given by conventional rules. As a result, much though perhaps not all such work has focused on the formalization of *received* pragmatic factors (like the Gricean maxims and various types of implicature) involved in communication rather than attempting something more fundamental to transform our very view of meaning and, in particular, the relation between semantics and pragmatics. In all of this work, the idea of equilibrium is simply a computational technique to solve games rather than a pervasive empirical fact about language. For such work, game theory is, in a sense, *everything*. Without it, there would be nothing to say. Reading such work, a philosopher or linguist could be reasonably secure that not much will need to change in his or her broad picture of language. But, of course, this is a serious problem because many (e.g., those favoring a more radical

---

55. See Merin (1999) and Benz et al. (2006) for instance. The title of the latter collection of papers—*Game Theory and Pragmatics*—indicates that their contributions are viewed as belonging only to pragmatics and, indeed, this is what they have been in the main. The corresponding semantics is left primarily to researchers in formal semantics and the combination of the two to a Gricean view of the whole.

pragmatics) feel that this foundational mainstream view of language is breaking down.

For my account, on the other hand, game theory (and situation theory) are merely the best currently available tools to forge this new way of thinking about language and meaning that I have called equilibrium semantics. My primary interest is not in the apparatus but in the reframing and transformation it makes possible. Indeed, I have added significantly to the tools where they fell short of the task.

As such, I urge the reader to be open to this omnipresent but largely unnoticed idea and image of equilibrium in language.

1.6 A Note to the Reader

Meaning is central to life. It is what makes us human. As such, this book is intended for a wide readership in the cognitive sciences: philosophers influenced by either Anglo-American or Continental or non-Western thought, linguists, artificial intelligence researchers, computer scientists, neuroscientists, psychologists, economists and other social scientists interested in language and communication, and even formally inclined theorists in the arts, especially literary and visual.

Readers will bring their own projects and philosophical commitments to the book and I want to alert them to one overriding principle as a guide to their reading: while the material is presented largely as a tightly knit and almost seamless framework, it has in fact many separable and interacting parts and levels that can be discerned. I have selected these constituents because I believe they are the components best suited to the ends I had in mind, but people with different goals or with different tastes may make different choices for these individual elements. A reader may accept my view of a foundational issue without being persuaded by some particular analysis or vice versa. I naturally hope that readers will find the entire structure credible and appealing, but some may pick and choose from the offerings and build their own wholes. I want readers to keep such variant architectonic possibilities in mind as they read the book.

It is also useful to briefly go over what is background and what is new. There are three principal elements that form the context for this book.

Foremost is the rich backdrop of the last century of semantics itself. I have tried to identify some weaknesses in mainstream approaches to meaning and to relate my proposals to this setting in chapter 1. Some more discussion is presented later, especially in chapter 5 where I consider the issue of indeterminacy and the Gricean challenge of defining meaning

and communication, and in chapter 6 where I address especially Kripke's arguments for a Russellian view of definite descriptions. I hope to attempt a more detailed comparative analysis with other semantic accounts in the future.

The second has to do with the perspective of situation theory, the required parts of which I have recounted in chapters 1 and 2 for the uninitiated reader. I hope this background will also help those who are familiar with its history to break free of some of the prejudices that have surrounded its decline. As I have argued, especially in chapter 2, it proved inadequate because some key pieces were missing, and I have tried to supply these items in what follows. Among these are a new perspective on infons (and therefore on partiality and fine-grainedness, two of situation theory's important strengths) and a new operation of unification.

The last is the context of game theory which has now been around for almost a century.[56] I have presented its basic ideas from scratch in chapter 3 but from my own perspective of *situated* choice. It is difficult to separate the background from this viewpoint but the experienced reader will have no trouble differentiating between the two. For newcomers, I have provided references to standard texts in the field.

Classical game theory is also concerned largely with single games (or with repetitions of a single game). A key innovation is the idea of interactions between and among games. This has led to many new elements, not least the product operation on games, conceiving the initial probability distributions as strategic variables, and interdependent games with a double fixed point solution—all introduced in chapter 4—but, again, it is not easy to produce a complete list of additions and alterations. Appendix A also contains an elaboration of my perspective on situated games and their solution that exploits an analogy made possible by the universality result of chapter 4.

I have followed a certain convention throughout the book: unless displayed, linguistic expressions are almost always mentioned in double quotes or small capitalization, and meanings or contents are italicized. The use of different alphabets or styles for the many symbols also follows a pattern that is harder to describe but should become familiar as the reader progresses through the book.

---

56. Zermelo (1913) was probably the first person to publish a theorem in the theory of games.