Part I

Introduction

Lecture 1

Information Flow: A Review

The three lectures in the first part of the book present an informal overview of a theory whose technical details will be developed and applied later in the book. In this lecture we draw the reader's attention to the problems that motivated the development of the theory. In Lecture 2 we outline our proposed solution. A detailed example is worked out in Lecture 3.

In the course of the first two lectures we draw attention to four principles of information flow. These are the cornerstones of our theory. We do not attempt to present a philosophical argument for them. Rather, we illustrate and provide circumstantial evidence for the principles and then proceed to erect a theory on them. When the theory is erected, we shall be in a better position to judge them. Until that job is done, we present the principles as a means of understanding the mathematical model to be presented – not as an analysis of all and sundry present day intuitions about information and information flow.

1.1 The Worldly Commerce of Information

In recent years, information has become all the rage. The utopian vision of an information society has moved from the pages of science fiction novels to political manifestos. As the millennium approaches fast on the information highway, the ever-increasing speed and scope of communication networks are predicted to bring sweeping changes in the structure of the global economy. Individuals and companies are discovering that many transactions that used to require the movement of people and goods, often at great expense, may now be accomplished by the click of a mouse. Information can travel at the speed of light; people and goods cannot.[1] The result is no less than a reshaping of

[1] If information about the future is possible, as it seems to be, then information travels *faster* than the speed of light.

our shrinking planet as cultural and commercial boundaries are transformed, for better or worse, by the increasing volume of information flow.

No doubt such future-mongering should be taken with more than a pinch of salt, but there can be little doubt that the prospect of life in "cyberspace" has caught the imagination of our age. Even in the most sober of society's circles, there is a mixture of heady excitement and alarm, a sense of revolution of an almost metaphysical sort.

Once one reflects on the idea of information flowing, it can be seen to flow everywhere – not just in computers and along telephone wires but in every human gesture and fluctuation of the natural world. Information flow is necessary for life. It guides every action, molds every thought, and sustains the many complex interactions that make up any natural system or social organization. Clouds carry information about forthcoming storms; a scent on the breeze carries information to the predator about the location of prey; the rings of a tree carry information about its age; a line outside the gas station carries information about measures in the national budget; images on a television screen in Taiwan can carry information about simultaneous events in Britain; the light from a star carries information about the chemical composition of gases on the other side of the universe; and the resigned shrug of a loved one may carry information about a mental state that could not be conveyed in words.

With this perspective, the current revolution appears to be primarily technological, with people discovering new and more efficient ways to transform and transmit information. Information is and always was all around us, saturating the universe; now there are new ways of mining the raw material, generating new products, and shipping them to increasingly hungry markets.

This book, however, is not concerned with technology. Our primary interest is not so much in the ways information is processed but in the very possibility of one thing carrying information about another. The metaphor of information flow is a slippery one, suggesting the movement of a substance when what occurs does not necessarily involve either motion or a substance. The value of the metaphor lies largely in the question it raises: How *do* remote objects, situations, and events carry information about one another without any substance moving between them?

The question is not a new one. A variety of answers have been proposed by philosophers, mathematicians, and computer scientists. Our starting point was the work of Dretske, which will be discussed below. However, before going into details it is worth asking what such an answer is meant to achieve.

Consider the following story:

Judith, a keen but inexperienced mountaineer, embarked on an ascent of Mt. Ateb. She took with her a compass, a flashlight, a topographic map, and a bar of Lindt bittersweet

chocolate. The map was made ten years previously, but she judged that the mountain would not have changed too much. Reaching the peak shortly after 2 P.M. she paused to eat two squares of chocolate and reflect on the majesty of her surroundings.

At 2:10 P.M. she set about the descent. Encouraged by the ease of the day's climb, she decided to take a different route down. It was clearly indicated on the map and clearly marked on the upper slopes, but as she descended the helpful little piles of stones left by previous hikers petered out. Before long she found herself struggling to make sense of compass bearings taken from ambiguously positioned rocky outcrops and the haphazard tree line below. By 4 P.M. Judith was hopelessly lost.

Scrambling down a scree slope, motivated only by the thought that down was a better bet than up, the loose stones betrayed her, and she tumbled a hundred feet before breaking her fall against a hardy uplands thorn. Clinging to the bush and wincing at the pain in her left leg, she took stock. It would soon be dark. Above her lay the treacherous scree, below her were perils as yet unknown. She ate the rest of the chocolate.

Suddenly, she remembered the flashlight. It was still working. She began to flash out into the twilight. By a miracle, her signal was seen by another day hiker, who was already near the foot of the mountain. Miranda quickly recognized the dots and dashes of the SOS and hurried on to her car where she phoned Mountain Rescue. Only twenty minutes later the searchlight from a helicopter scanned the precipitous east face of Mt. Ateb, illuminating the frightened Judith, still clinging to the thorn bush but now waving joyously at the aircraft.

Two previously unacquainted people, Judith and the helicopter pilot, met on the side of a mountain for the first time. How did this happen? What is the connection between the helicopter flight and Judith's fall, such that the one is guided by the location of the other?

Naturally, common sense provides answers – in broad outline, at least. We explain that the helicopter was flying over that part of the mountain because the pilot believed that there was someone in danger there. The location had been given to the pilot at the Mountain Rescue Center shortly after the telephone operator had turned Miranda's description into a range of coordinates. Miranda also conveyed the description by telephone, but her information was gained from the flashes of light coming from Judith's flashlight, which was responding to the desperate movements of Judith's right thumb as she clung to the mountainside.

This establishes a physical connection between the two events, but a lot is left unsaid. Most events are connected in one way or another. How is this particular connection capable of conveying the vital piece of information about Judith's location? Consider the nature of the connection. It is made up of a precarious thread of thoughts, actions, light, sound, and electricity. What is it about each of these parts and the way they are connected that allows the information to pass?

A full explanation would have to account for all the transitions. Some of them may be explained using existing scientific knowledge. The way in which the switch on Judith's flashlight controlled the flashes, the passage of light from

the flashlight to Miranda's eyes, and the transformation of spoken words into electrical signals in telephone wires and radio waves may all be explained using models derived ultimately from our understanding of physics. The motion of Judith's thumb muscles and the firing of Miranda's retinal cells, as well as the many other critical processes making up the human actions in this story, require physiological explanation. A knowledge of the conventions of mapmakers and English speakers are needed to explain the link between the combinations of words, map coordinates, and the actual location on the mountain. And, finally, the psychology of the various people in the story must be understood to bridge the considerable gap between perception and action: between Judith's fall and her moving the switch, between the light falling on Miranda's retina and her mouthing sounds into a cellular phone, between the sounds coming from the telephone, the scribbled calculations on a message pad, and the pilot's hurried exit from the building.

A full explanation would have to include all these steps and more. It is no wonder that we speak of information, knowledge, and communication; life is too short to do without them. Yet it is not just the complexity of the explanation that makes the prospect of doing without information-based vocabulary so daunting. Stepping below the casual uniformity of talk about information, we see a great disunity of theoretical principles and modes of explanation. Psychology, physiology, physics, linguistics, and telephone engineering are very different disciplines. They use different mathematical models (if any), and it is far from clear how the separate models may be linked to account for the whole story. Moreover, at each stage, we must ask why the information that someone is in danger on the east face of Mt. Ateb is carried by the particular event being modeled. This question is not easily stated in terms appropriate to the various models. To explain why the pattern and frequency of firings in Miranda's retinas carry this information, for instance, we need more than a model of the inside of her eye; the question cannot even be stated in purely physiological terms.

What are the prospects for a rigorous understanding of the principles of information flow? It is relatively uncontroversial that the flow of information is ultimately determined by events in the natural world and that the best way of understanding those events is by means of the sciences. But explanations based on the transfer of information are not *obviously* reducible to scientific explanations, and even if they are, the hodgepodge of models and theoretical principles required would quickly obscure the regularities on which the information-based explanations depend. The possibility exists that a rigorous model of information flow can be given in its own terms; that such phenomena as the chaining together of information channels, the relationship between error and gaps in the chain, and the difference between a reliable information source and accidental

correlations can be explained in a precise way, not by reducing the problem to physics or any other science, but by appealing to laws of information flow. This book aims to provide such a model.

1.2 Regularity in Distributed Systems

In determining the information carried by Judith's flashlight signal, a spectrograph is of little use. Whatever information is carried depends not on intrinsic properties of the light but on its relationship to Judith and other objects and events.

It is all too easy to forget this obvious point if one focuses on information conveyed by means of spoken or written language. The information in a newspaper article appears to depend little on the particular copy of the newspaper one reads or on the complex mechanism by which the article was researched, written, printed, and distributed. The words, we say, speak for themselves. The naïviety of this remark is quickly dispelled by a glance at a newspaper written in an unfamiliar language, such as Chinese. There is nothing in the intricate form of Chinese characters inscribed on a page that conveys anything to the foreigner illiterate in Chinese.

Logical and linguistic investigations into the topic of information give the impression that one should be concerned with properties of sentences. Even when it is acknowledged that information is not a syntactic property of sentences and that some system of interpretation is required to determine the information content of a sentence, the role of this system is typically kept firmly in the background. In Tarskian model theory, for example, and in the approaches to natural language semantics based on it, an interpretation of a language consists of an abstract relation between words and sets of entities. No attempt is made to model what it is about human language-using communities that makes this relation hold.[2]

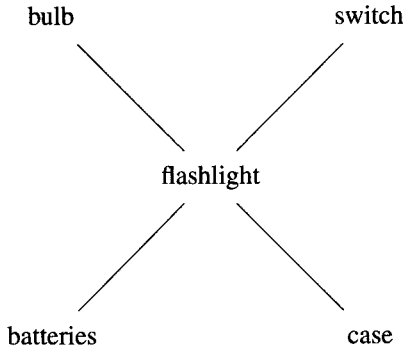
By contrast, when one looks at nonlinguistic forms of communication, and the many other phenomena that we have listed as examples of information flow, the spatial and temporal relationships between parts of the system cannot be ignored. The very term "information flow" presupposes, albeit metaphorically, a spatial and temporal separation between the source and the receiver of the information.

[2] The criticism is not that the topic has been ignored, as philosophers of language have had much to say about it, but that their proposals have not been incorporated into logico-linguistic theories. The relation between a name and its bearer, for example, is taken to be a primitive relation of semantic theory; the contingent facts of language use that establish the relation are ignored on the grounds that the only semantically relevant feature of a name is its bearer.

We draw attention to the importance of examining information flow in the context of a system in our first Principle.

The First Principle of Information Flow: Information flow results from regularities in a distributed system.

Two features of this principle require immediate comment: that the systems in which information flows are *distributed* and that the flow results from *regularities* in the system. By describing a system in which information flows as "distributed," we mean that there is some way in which it is divided into parts, so that the flow of information is from one part (or parts) to another. For example, we may consider Judith's flashlight to be a system in which information flows: the lit bulb carries the information that the switch is on and the battery is charged and so forth. The flashlight may be divided into the following parts:



We do not intend to suggest that this division into parts is unique or comprehensive. Each of the components of the flashlight has parts that are not represented (but could be) and there are a host of different ways of decomposing the system that would not list any of the parts depicted above. (Indeed, the relativity of this decomposition is crucial to the story we will tell.)

The parts of an information system are often spatial or temporal parts but they do not need to be. The conception of information flow we develop is very broad, encompassing abstract systems such as mathematical proofs and taxonomic hierarchies as well as concrete ones like the above. In abstract systems the relation of whole to part is also abstract, and the metaphor of "flow" has to be interpreted even more loosely. We place no restriction on what kind of thing may count as a part, only that the choice of parts determines the way in which we understand what it is for information to flow from one part to another.

The first principle of information flow also states that information flow results from *regularities in the system.* It is the presence of regularities that links

the parts of a system together in a way that permits the flow of information. The fact that the components of Judith's flashlight are more or less adjacent in space and have overlapping temporal extents is not sufficient to explain the information carried by the lit bulb about the position of the switch or the state of the batteries. It is of the greatest importance that the components of the flashlight are connected together in such a way that the whole flashlight behaves in a more or less predictable manner. In this case, the regularities that ensure the uniform behavior of the system are mostly electrical and mechanical in nature. The contacts in the switch, the design of the case, and many other details of the construction of the flashlight go to ensure that the flashing of the bulb is systematically related to the position of the switch.

The behavior of the system need not be entirely predictable for information to flow. Properties of components of the flashlight, such as the discoloring of the plastic case due to exposure to sunlight, are not at all predictable from properties of the other components; yet this indeterminacy does not interfere with the regular behavior of the system in the informationally relevant respects. More complex systems may even be highly nondeterministic while still allowing information flow. Yet, as a general rule, the more random the system the less information will flow.

The range of examples in the story of Judith on Mt. Ateb show that information flows may be due to a wide range of factors. Some of them are "nomic" regularities, of the kind studied in the sciences; others, such as those relating a map to the mapped terrain, are conventional; and others are of a purely abstract or logical character. Sometimes the regularity involved is very difficult to pin down. Parents can often tell when their children are getting ill just from the look of their eyes. The relationship between the appearance of the eyes and the condition permits this inference reliably, but even an ophthalmologist would hesitate to say on what exactly the inference depends.

Despite the wide range of regularities that permit information flow, it is important to distinguish genuine regularities from merely accidental, or "statistical" regularities. Accidents are not sufficient for information flow. To give an example, suppose that on the occasion of Judith's fall, Miranda caught sight of the light of the moon reflected from a small waterfall on the mountainside. By chance, we suppose, the waterfall reflected the moonlight in a sequence of flashes very similar to the Morse code SOS. In such circumstances, Miranda might have formed the belief that there was someone in trouble on the mountain in the approximate location of Judith's fall. Her belief would be true but she would not have the information. Information does not flow from A to B just because someone at B happens to have been misled into believing something correct about what is going on at A.

Now suppose that over the course of several months, a large number of climbers are saved from the treacherous slopes of Mt. Ateb by Mountain Rescue, alerted by reports of SOS signals flashed from the mountain. In fact, the flashes were all caused by reflected moonlight and a spate of visiting climbers, rather overzealous in their desire to report potential accidents. It is important that we assume that there really is no more to it that this, that is, no mysterious link between troubled climbers and water spirits. The result would be that a statistical regularity is established between flashes from the mountainside and climbers in distress. It should be clear that this spurious regularity is no more able to establish information flow than the one-time coincidence considered previously.[3]

1.3 Information and Knowledge

There is a close connection between information and knowledge. Puzzles similar to those discussed in the previous section are used by philosophers to test different theories of knowledge. Indeed, the origin of the work presented here was an attempt to elaborate and improve on a theory of knowledge presented by Fred Dretske (1981) in his book *Knowledge and the Flow of Information*. The informational role of regularities in distributed systems will be better appreciated if seen in the light of Dretske's theory.

Since Gettier's famous paper of 1963, philosophers have been looking for the missing link between true belief and knowledge. The "traditional" account is that knowledge is *justified* true belief. Miranda knows that someone is in trouble on the mountain because her belief is justified by her knowledge of Morse code and other relevant considerations. But consider the case in which the flashes were produced by reflected moonlight. Miranda's belief and its justification would remain the same, but she would not *know* that someone is in trouble. This is Gettier's argument and there have been many responses to it.[4]

Recently, the topic has largely been put aside, not because an agreed upon solution has been found but because many have been proposed and no clear victor has emerged. Dretske's solution was one of the first. He proposed that information is the missing link. Very roughly, Dretske claims that a person knows that p if she believes that p and her believing that p (or the events in her head responsible for this belief) carries the information that p. To the extent

[3] An interesting question, which tests the best of intuitions, is whether information is carried by Judith's flashlight signals against a background of fortuitously correct accident reports of the kind considered in our example.

[4] For a survey, see Shope (1983).

that our beliefs carry information about the world, they play an invaluable role in guiding our actions and in communicating with others.

We take Dretske's account of the relationship between information and knowledge to be an important insight.[5] As a bridge between two subjects, we can use it both as a means of applying our theory to epistemology and also as a way of incorporating epistemological considerations into the theory of information.

For example, the first principle of information flow is illuminated by considering some related epistemological puzzles. A perennial problem in the philosophy of knowledge is that of accounting for a person's knowledge of remote facts. Miranda's knowledge of the world must stem from her experience of it, and yet a great deal of her knowledge concerns things that are not and never were part of our immediate physical environment. She may never have climbed Mt. Ateb herself, and she is certainly not able to see what happened on the scree slope in the dark. How is she able to know about parts of the world that are beyond her experience?

A rough answer is as follows: Things outside Miranda's experience are connected in lawlike ways to things within her experience. If the world beyond her senses bore no relationship to her experience, then she would not be able to know about it. It is the regularity of the relationship that makes knowledge possible. This answer places the philosophical problem squarely within the scope of our current investigations. Miranda and the remote objects of her knowledge form a distributed system governed by regularities. The fact that the system is distributed gives rise to the problem; the regularity of the system provides the only hope of a solution.

Dretske's approach to the theory of knowledge is not dissimilar to those who claim that it is the *reliability* of the belief-producing process that constitutes the difference between knowledge and mere true belief (Goldman, 1979, 1986; Nozick, 1981; Swain, 1981). Indeed, there is a close connection between information flow and reliability. For a signal to carry information about a remote state of affairs, it must have been produced by a reliable process. An unreliable process will not permit information to flow.

Reliability, however, is clearly a matter of degree. Some processes are more reliable than others, and what counts as sufficiently reliable may vary with circumstances. Consider Judith's flashlight. The information that the bulb is lit carries the information that the switch is on, because they are linked by a

[5] We do not accept Dretske's account of information based on probability theory, and motivation for the design of our theory can be traced back to inadequacies in his proposals. Dretske's proposals will be discussed in more detail in Section 1.5 of the present chapter.