ON THE EFFICACY OF REPRESENTATION

Atsushi Shimojima

Submitted to the faculty of the Graduate School
in partial fulfillment of the requirements
for the degree
Doctor of Philosophy
Indiana University

May, 1996

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

Doctoral Committee	K. Jon Barwise (Principal Advisor)
	Anil Gupta
	Michael Friedman
April 22, 1996	Alice ter Meulen

For my parents, Kunihiko and Yumiko Shimojima.

Acknowledgements

First and foremost, I would like to thank my principal advisor, Professor Jon Barwise, for being the excellent role model of brilliant, creative, and honest researcher. Among other things, he taught me how to discern the interesting from the uninteresting, how to be wild in one's imagination, and why to be honest in writing. It is, however, a different issue how much of this I have made my own virtues. Professor Anil Gupta greatly helped improving my theory, with his usual penetrating but judicious comments on the earlier drafts. Professor Michael Friedman played the role of commentator from the philosophical side, while Professor Alice ter Meulen exposed me to issues in natural language semantics that are relevant to my project. Special thanks go to my friend Gerard Allwein, who were always willing to help me in the mathematical side of this project. I am also grateful to the old and new members of Indiana University Visual Inference Laboratory, Kata Bimbo, Michael Chui, Ruth Eberle, Kathi Fisler, Eric Hammer, Isabel Luengo, and Mike Wollowski, who patiently listened to and commented on the often rough ideas that I presented in the regular colloquia. Finally, I would like to express my gratitude to my parents, Kunihiko and Yumiko Shimojima, who trusted me and kept supporting my family throughout my graduate career.

Abstract

In this dissertation, “representations” are external objects that we use to present information about some other objects on the basis of some fixed semantic rules. All the following objects count as representations: a set of Japanese declarative sentences describing Mount Fuji, a time table of the Boston subway system, a geometry diagram used to demonstrate the Pythagorean theorem, a state map of the United States, a relief map of a Rocky terrain, a ball-and-stick model of a molecular, and a scale model of the Rockfellar center.

The dissertation studies how different modes of presenting information exhibit different degrees and kinds of efficacy as an aid for human reasoning. The main hypothesis, which we dub the “Constraint Hypothesis,” is that the varience in inferential potentials of different modes of representation is largely attributable to particular ways in which structural constraints on representations match and mismatch with constraints on their targets. If proven true, the hypothesis provides a definite perspective for the study of information representation in general, which is rapidly taking the shape of an independent branch of science.

As the basis for our analysis, we adopt the mathematical framework of the “qualitative information theory” being developed by Dretske, Barwise, Perry, and Seligman. We analyze several key properties that account for the efficacy of particular modes of representation, including the capacity of providing “free rides” in inference, the

property of being over-specific in representing information, and the property of being self-consistent in information content. We also propose an analysis of the conceptual boundary between "linguistic" modes and the "graphical" modes of representation, and explain some of the differences in their cognitive potentials. All of these analyses crucially refer to the structural constraints on representations and their match and mismatch with the constraints on the targets, and hence lend piecemeal yet sturdy supports to the Constraint Hypothesis that we are advocating.

Contents

Chapter 1

Introduction

In this dissertation, “representation” means a relationship holding between two objects when we (the cognitive agents) use one to present information about the other on the basis of particular semantic rules. We develop a model-theoretic characterization of different *modes* of representation, intended to account for their varying potentials in the process of human reasoning. In its subject, this work is a descendant of Nelson Goodman’s “general theory of symbols” conceived as a “systematic inquiry into the varieties and functions of symbols” (Goodman 1968, p. xi). In its method, this work is a non-standard extension of model-theoretic semantics, broadly conceived as the application of the tools of modern mathematics to the analysis of the relationships between representations and the things that they represent. In its aspiration, our work is a part of the research program of qualitative information theory, which was first envisioned by Dretske (1981) and has been developed mainly by Barwise, Perry, and Seligman. We will start this introductory chapter by elaborating these points, to give more precise ideas about the subject, method, and background of our work. We will then state the main hypothesis of this dissertation, and describe what steps we will take to substantiate it.

1.1 Subject

Two terminological points first. Preserving the double usage of the English term, we will use "representation" not only to denote a relation between two objects, but also to denote *the thing that represents*, namely, the object that has the relation of representation to the other object. All the following things count as representations in this sense: a set of Japanese sentences describing Mount Fuji, a set of first-order formulas describing the premises of a syllogism, a time table of the Boston subway system, a geometry diagram used to demonstrate the Pythagorean theorem, a state map of the United States, a picture of Mount Fuji by Hokusai, a relief map of a terrain in the Rockies, a ball-and-stick model of a molecular, and a scale model of the Rockfeller Center.

The term "representation," however, will *not* be used for so-called "mental representations" postulated in certain theories of cognitive psychology (e.g. Braine 1978, Kosslyn and Pomerantz 1977, Johnson-Laird 1983, Johnson-Laird and Byrne 1991, Rips 1994, Shepard and Cooper 1982). Thus, anything that you have in mind, whether it is an image, a sentence, or a model, is not a representation within this dissertation. The term only denotes what might be called "external representations," the objects placed in certain spatio-temporal locations and accessed primarily through perception. The mental representations are not within the direct scope of this dissertation, although our analysis of different modes of external representations may have an application to the understanding of the theoretical status of mental representations.

We use the term "mode of representation" in a sharp contrast to the term "content of a representation." Different representations can have the same content while presenting it in different modes. A set of Japanese sentences describing Mount Fuji might, perhaps, present the same information as a picture by Hokusai, but they are presenting the information in different modes. A set of English sentences, a Venn diagrams, and an Euler diagram may say the same thing about the relationships among certain sets, but they are presenting that content in three different modes. A numerical table, a scatter plot, and a chartmap in statistics may present the same data set, but they are presenting it in different modes. Bertin (1973) discusses one hundred different modes in which graphic representations present the same data set about the distribution of work force in ninety counties in France. On the other hand, different representations can have different contents while presenting them in the same mode. Two sets of Japanese sentences may be presenting different contents (in describing Mount Fuji and Atsushi's tiny apartment, for example), but they are presenting their contents in the same mode (at a certain level of analysis at least). Similarly for two scatter plots that present different data sets, two Venn diagrams that present different sets of premises of syllogisms, and two massing models of a building whose design is being planned.

Consider two representations in different modes, say, the following Venn diagram and Euler diagram.

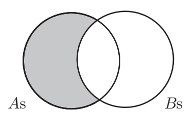


Figure 1.1

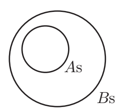


Figure 1.2

Presumably, these two representations have the same information content, namely, that all As are Bs. But they present this information in different modes. What then constitutes this difference in mode? First of all, the information corresponds to different syntactic features of the diagrams. In the Venn diagram, the information is indicated by the syntactic fact that the complement of the circle labeled "Bs" with respect to the circle labeled "As" is shaded. In the Euler diagram, the same information is indicated by the syntactic fact that the circle labeled "As" is enclosed in the circle labeled "Bs." Thus, *different semantic rules* concerning the indication relation are associated with these diagrams. Secondly, while one circle is completely inside the other circle in the Euler diagram, the overlap between the two circles in the Venn diagram is *partial* in that each circle has a portion that does not overlap with the other. This feature is not particular to this Venn diagram, but a general rule for all Venn diagrams. Namely, as a syntactic convention, all circles appearing in a Venn diagram must overlap but must do so only partially. Notice that this convention is not in effect for the Euler diagram. Thus, *different syntactic rules* are associated with these two diagrams.

Roughly, then, difference in mode may consist in differences in both semantic rules and syntactic conventions. The above diagrams differ in both counts. Given a mode, then, we can collect representations that obey the syntactic and semantic rules that determines the mode. Since all these representations are governed by the same syntactic and semantic rules, we can say that they constitute a *system*. In fact, it is becoming common to use the phrase "system of representation" or "representation system" as a loose synonym to "mode" in the sense just characterized. We will follow this practice, and use this convenient double locution. Later, however, we will assign a technical meaning to the phrase "representation system" and use it as an *explication* of the notion of mode.

Given two representations in different modes, we often have an intuition about their relative efficacy. We may feel that one is *clearer*, *more illuminating*, or *more efficient* than the other as a representation. We may simply feel that one is *better* than the other, without any articulate idea about the respect in which one is "better" than the other. For example, Hokusai's picture of Mount Fuji is certainly more *vivid* than the corresponding description of the mountain with Japanese sentences. We translate the premises of a syllogism from English to a Venn diagram to obtain a *clearer* expression of the premises. We translate a table of statistical data into a bar chart to obtain a more *suggestive* presentation of the data. A scale model of a building often gives us a *better* idea, whatever that means, about the design of the building than a pile of blueprints.

The subject of this dissertation is the origins of these intuitions about the efficacy of different modes of representation. More specifically, we wish to study different modes in which representations present information, analyze the potentials of these different modes to mediate, hinder, assist, and control the user's reasoning, and thus account for the different degrees and kinds of efficacy exhibited by the different modes. It is in this respect that our project is a descendant of Goodman's "general theory of symbols," which aims at a "comprehensive grasp of the modes and means of reference and of their varied and pervasive use in the operations of the understanding" (Goodman 1968, p. xi).

There are, however, two different ways in which the efficacy of two representations may differ. One is due to the difference in the modes in which they present their contents. All the examples in the last paragraph are intended to illustrate this kind of difference. The other kind of difference is due to the contents that the two representations represent. For example, the sentence "Jones is insane and Smith's body was mutilated" is more useful than the sentence "Jones is insane" in figuring out who murdered Smith simply because the former has *more* information than the other. It is not because they present their contents in different modes. Also, a picture of Smith's body may be more or less useful than a picture of the entire scene of the murder. Whichever it may turn to be, the difference comes from the contents of the pictures, not from the modes the pictures present their contents. One may be curious about what makes some information more relevant or useful to the given problem than some other set of information. But it is an issue that can be addressed independently of the modes of representation in which the information is presented. Our focus is on the difference of efficacy due to the difference in the *modes* of representation, not the difference due to the *contents* of representations. In case the efficacy of two representations differ in both respects, we are interested in analyzing the difference insofar as it is due to the variance in mode.

1.2 Method

Barwise and Perry (1983) succinctly characterize the basic method of model-theoretic semantics in the following way (p. 27):

> Semantics is the study of linguistic meaning, of the relationships that hold between expressions of language and things in the world. This study can be conducted in a precise way using the tools of modern mathematics. The approach is generally called "model-theoretic semantics," since model theory is the part of logic concerned with the relation between the linguistic expressions of mathematics and the mathematical structures they describe.

Here, we can roughly equate "the tools of modern mathematics" with set-theoretic tools such as sets, pairs, relations, unions, products, functions, partial functions, functions from functions to functions, and so on. Our own method consists in using these set-theoretic constructions as building blocks to model the relationship between representations and the things represented, and in this basic method, our work is a descendant of model-theoretic semantics. We, however, extend this method to all modes of representation in general, not just natural language. Apart from this difference in subject, there are important aspects of model-theoretic semantics that our framework does not inherit. Here we will confine ourselves to give a general grasp of the difference, leaving the detail for a later chapter.

It is illuminating to compare our stance to model-theoretic semantics with the stance of "situation semantics," proposed by Barwise and Perry (1983). While admitting the great value of the basic method of model-theoretic semantics, Barwise and Perry extensively criticize the particular framework of model-theoretic semantics that was standard when they wrote. (It is still standard now, although to a lesser extent perhaps.) According to Barwise and Perry, standard model-theoretic semantics is flawed in its attempt of adapting the techniques of mathematical model theory directly to the analysis of natural language. They write (p. 28):

> But the heritage of model theory, however illustrious, is a mixed blessing. For the founders of modern logic—Frege, Russell and Whitehead, Gödel, and Tarksi—were preoccupied with the language of mathematics. Because of this preoccupation, many assumptions and attitudes about this language were built into the very heart of model theory, and so came to be assumptions about the nature of language in general. These assumptions have made it increasingly difficult to adapt the ideas of standard model theory to the semantics of natural languages.

In particular, Barwise and Perry point out that a theory within standard model-theoretic semantics is usually judged by how well it accounts for the entailment relation, why it is that a sentence like "Socrates is mortal" follows from "All men are mortal" and "Socrates is a man." Barwise and Perry see this preoccupation with the entailment relation between sentences as an unwelcome heritage from mathematical model theory, and claim that "there is much more evidence than just entailments for which a semantic theory must account" (p. 28). They cite six such additional phenomena to be explained by an adequate semantic theory, putting the strongest emphasis on the so-called "efficiency of language," namely, the fact that linguistic expressions can be "recycled, can be used over and over again in different ways, places, and times and by different people, to say different things" (p. 32). According to Barwise and Perry, it is not enough to treat the efficiency of language and other five phenomena as "minor headaches" to be explained by amending the existing model-theoretic semantics (p. 28). Barwise and Perry's situation semantics is the result of reshaping model-theoretic semantics in a way to capture these additional phenomena. While preserving the idea of using set-theoretic tools to develop semantics of natural language, it filters out unwelcome assumptions and attitudes from mathematical model theory, and adds necessary ingredients to capture the additional phenomena fundamental in natural language.

Our own framework has a similar stance toward standard model-theoretic semantics. Like situation semantics, it preserves the idea of using set-theoretic tools to model the relationships between representations and the represented. Like Barwise and Perry, we try to reshape the framework of mathematical model theory to fit our particular need of analysis. The difference is that situation semantics targets at natural language, while our framework targets at all modes of representations in general, including, therefore, the language of mathematics itself. Like Barwise and Perry, we demand more of a semantic theory than just modeling the entailment relation between representations. The difference is that while situation semantics focuses on the efficiency of language, i.e., different sorts of contextual dependency of the interpretation of linguistic expression, we focus on the comparison of modes of representation in their efficacy. Methodologically, situation semantics is an older sibling of our framework, if model-theoretic semantics is a parent.

There exists, however, an alternative framework that applies the technique of model theory to analyze the workings of different modes of representation. The framework, which we will call the "logic" framework for convenience, has been applied by Shin (1990), Hammer (1995), and Luengo(1995) to the diagrammatic modes of representation, such as the system of Venn diagrams, Peirce's α-system, Harel's system of higraphs, and diagrams for plane geometry. In my opinion, the framework has played and will play a major role in analyzing the workings of non-linguistic modes of representation. The framework, however, has its own limitations. It adapts the technique of mathematical model theory more straightforwardly than we do, and this makes our framework stand to the logic framework roughly as situation semantics stands to standard model-theoretic semantics. We will discuss the relationship between our framework to the logic framework after we develop our framework in more detail.

1.3 Background

As the reader may have realized, we use the term "information" rather freely, as though it were an entity on its own. For example, our definition of the term "representation" already contains the term "information," declaring that representation is a relation holding between two objects when we use one to present *information* about the other on the basis of particular semantic rules.

Of course, one can avoid the realistic talk about information while preserving the idea that information is something that is presented. That, perhaps, better fits the common view of information, which Dretske (1981) characterized as follows (p. vii):

> It is common to think of information...as something that depends on the interpretive efforts—and, hence, prior existence—of intelligent life. According to this view, something only *becomes* information when it is assigned a significance, interpreted as a sign, by some cognitive agent. Beauty is in the eye of the beholder, and information is in the head of the receiver. To speak of information as *out there*, independent of its actual or potential use by some interpreter, and antedating the historical appearance of all intelligent life, is bad metaphysics. Information is an artifact, a way of describing the significance *for some agent* of intrinsically meaningless events. We *invest* stimuli with meaning, and apart from such investment, they are informationally barren.

We could, perhaps, adopt this anti-realistic view of information, and render the expression "present information" in our definition of "representation" as a *way of speaking*, as a locution that does not commit itself to the existence of the entity that it *prima facie* talks about. Here, the relation of presenting information is subject to further analysis, which ultimately gets rid of the talk of information, in favor of, say, the description of the significance of a representation as the physical object to the agent as a biological being.

Dretske rejects the anti-realistic view of information by attributing it to the confusion between the notion of information and that of meaning. He contrasts the view with the following realistic view of information (p. vii):

> Once this distinction [between information and meaning] is clearly understood, one is free to think about information (though not meaning) as an objective commodity, something whose generation, transmission, and reception do not require or in any way presuppose interpretive processes.

Apart from Dretske's argument concerning the confusion of information and meaning, we have a reason for retaining realism about information, at least for a while. As scientists, we are curious where the realistic view about information will take us in understanding various semantic phenomena about truth, reference, and representation. We simply want to explore the potentials of the notion of information in explaining the semantic phenomena. The enterprise may or may not be fruitful, and it is perhaps not anything that Descartes would recommend. But at the same time, we do not think that a priori discussion of the ontological status of information would settle the issue either. And at least while we are pursuing this enterprise, it does not make sense to reduce the talk of information to anything different, say, to the talk of the "significance" relation between an physical event and an agent. This dissertation is a part of this enterprise, and we aim to add an positive evidence to the prospect of this enterprise. For this reason, we follow the lead of Dretske, and adopt at least "methodological" realism about information. The reader will see realism reflected in our framework, where we postulate independent entities called "pieces of information."

I am not, of course, the first person engaged in this research program. Dretske (1981) already envisioned it under the name of "semantic theory of information." Dretske contrasts it to mathematical information theory, or communication theory, initiated by Hartley (1928), Shanon (1948), and Wiener (1948). Communication theory can be seen as an endeavor to provide a measure of how much information is to be associated with a given state of affairs and, in turn, a measure of how much of this information is transmitted to, and thus available at, other points. As such, communication theory "deals with *amounts of information*," but not "the information that comes in those amounts" (p. 3). In contrast, the *semantic* theory of information is concerned with the message, the information itself, that flows in a particular act of communication. The theory is concerned with information as "a commodity that is capable of yielding knowledge," and as "a commodity that can be transmitted, received, exchanged, stored, lost, recovered, bought, and sold" (p. 47). The semantic theory of information tries to tell "what this thing is," and thereby solve the semantic issues concerning truth, reference, and representation.

This research program on the general theory of information underlies in Barwise and Perry's situation semantics, which aims to "account for how language fits into the general flow of information" (1983, p. 45). Barwise takes up the program of qualitative information theory more explicitly in his papers (1991, 1993) and the paper (1995) with Gabbay and Hartonas. These papers aim at a general mathematical framework with which we understand various phenomena involving a flow of information in one way or another. The program is still growing in this direction, mainly due to the effort of Barwise and Seligman (1993, 1996).

An application of the qualitative information theory to the semantic analysis of different modes of representation has been envisioned by Barwise and Etchemendy (1990b) and carried out in Barwise and Etchemendy (1990a). They aim at "an information-based theory" of valid inference that "does not presuppose that information is not presented linguistically, or, for that matter, in any particular medium" (1990b, p. 21). It is a theory of deduction "rich enough to assess the proofs that use multiple forms of representations" and to assess "inference that is not inextricably tied to linguistic forms of representation" (1990b, p. 9). Thus, unlike our project, its focus is on the *validity* of inferences carried out with representations, rather than on the *efficacy* of representations in inferences.

As a part of this research program, our work owes many critical ideas, technical and non-technical, to these predecessors. We will spell out our debts in a later chapter, where we fully spread out our conceptual framework.

1.4 The Constraint Hypothesis

Our goal is to study different *modes* of representation, and thereby account for the varying degrees and kinds of efficacy exhibited by them in the process of human reasoning. This task is not straightforward. As a rough synonym to "effectiveness," "efficacy" generally means "power or capacity to produce a desired effect." Thus, the efficacy of a tool is relative to the use to which it is put, and as the use of the tool is diverse, there are many different kinds of efficacy the tool may exhibit.

And, yes, the use of a representation in human problem solving *is* diverse. We may use a representation as an *inventory* of the information we gather. We may use a representation as a site to *analyze* the gathered information. If we work together to solve a problem, we may use a representation as a *communication tool*.[1] Depending on our styles of problem solving, we may place more emphasis on the use of a representation for a quick *retrieval* of information, as opposed to the comprehensiveness of the stored information. Also, we may want to use a representation not as a static record of information, but as a site for *updating information*.[2] We may also want to use a representation to *check the consistency* of gathered information.

All this means that our task is not an easy one. There are many different uses of representations in human reasoning, and there are accordingly many different notions of efficacy that can be applied to different modes of representation. Does this mean that it is hopeless to analyze the relative efficacy exhibited by different modes of representation? No. It only means that such a study should be sensitive to the existence of many varieties of efficacy, and that our overall intuition about the relative efficacy of a mode of representation (its clarity, efficiency, suggestiveness, and others) can be a conglomerate of such different varieties of efficacy. The diversity of the subject simply calls for a careful treatment of the subject.

In fact, the main thesis of this dissertation asserts the possibility of a *principled* study of the subject. Using "inferential potential" as a blanket term for different degrees and kinds of efficacy of a mode of representation as an aid for reasoning, we can formulate the thesis in the following conjunction:

> **Constraint Hypothesis** Representations are objects in the world, and as such they obey certain structural constraints that govern their possible formation. The variance in inferential potential of different modes

[1] These uses of representations roughly correspond to Bertin's three-way classification of the uses of graphics: (1) recording information, (2) communicating information, and (3) processing information. He also discusses the crucial properties of graphics relative to each of these uses. See Bertin (1973), pp. 160–164.

[2] Barwise and Etchemendy (1990) points out that this "dynamic" use is typical for diagrams (such as Venn diagrams and position diagrams used to solve GRE-style analytical problems). We will later discuss the topic of dynamic use of representations in more detail.

> of representation is largely attributable to different ways in which these structural constraints on representations match with the constraints on targets of representation.

This dissertation is devoted to the substantiation of this thesis. If this dissertation is successful, it will give a definite perspective from which we grasp varying and elusive phenomena concerning the efficacy of different modes of representation. The reader may be wondering what kind of regularities we mean by "constraint on representations," or what kind of match and mismatch we have in mind between constraints on representations and constraints on their targets. The entire dissertation will serve as an answer to such questions. Let us not, therefore, dwell on the content of the hypothesis anymore here. We will instead sketch how we will go about arguing for this thesis.

1.5 Organization of the dissertation

There are two basic steps that we take for the substantiation of the Constraint Hypothesis: (1) identify a property or phenomenon, commonly associated with certain representation systems, that accounts for a particular variety of efficacy (or inefficacy) exhibited by the systems, and (2) show that the property is attributable to a particular way in which structural constraints on the representations of the systems match with constraints on the targets of representation. To facilitate the step (2), we will develop a conceptual framework designed to highlight structural constraints on representations, those on targets, and their match and mismatch. Following steps (1) and (2), we will account for several important varieties of efficacy exhibited by representation systems, and use the success of each account as an piecemeal argument for the Constraint Hypothesis. The result will be an accumulated evidence for the hypothesis. There might also be an a priori argument for the hypothesis, but no attempt will be made at it in this dissertation.

We divide the labor into five chapters. In chapter 2, we identify the first property of certain representation systems, namely, the capacity of providing what we call "free rides" in reasoning. It accounts for the mode of representation when we use a representation as a site of dynamic update of information. We will give an informal analysis of the property which is, nonetheless, specific enough to show that it is due to a particular way in which the constraints on the representations of a representation system match with the constraints on targets of the system.

In chapter 3, we identify another property of representation systems, called "content specificity." Here, content specificity means the inability of presenting a certain chunk of information in isolation, without presenting certain other. The property thus accounts for expressive inflexibility of a representation system, an important variety of inefficacy. Again, our analysis of the property is informal, but specific enough to show that the property is due to a particular mismatch between the constraints on the representations of a representation system and those on the targets.

Chapter 4 formally develops the conceptual framework that we have used in chapters 2 and 3 and will use in chapters 5 and 6. It is specifically designed to capture (1) constraints governing the representations of a system, (2) constraints governing the targets of the system, and (3) match and mismatch between the two sets of constraints. As we indicated above, the formal framework is a species of model-theoretic semantics in a broad sense, but it owes much of its conceptual machinery to the framework of situation semantics and the qualitative information theory developed by Dretske, Barwise, Perry, Seligman, and others. We use the framework to refine our informal analyses of free ride and content specificity in chapters 2 and 3. We also discuss another important property of representation systems, called self-consistency, which endows a system with certain "model-theoretic" capacities. Our analysis of this new property will lend another piecemeal support to the Constraint Hypothesis. We will also spend a part of the chapter to compare our framework with the logic framework, a framework that develops rather standard model-theoretic semantics for non-linguistic representation systems.

In chapter 6, we will ask ourselves, what, if any, is the boundary between so-called "linguistic" modes of representation and "diagrammatic" modes of representation. After reviewing several previous proposals about the boundary, we will use our conceptual framework to identify the property common in diagrammatic representation systems and absent from linguistic representation systems. More specifically, we will claim that every diagrammatic system of representation supports a certain matching of a "nomic" constraint on representation with a constraint on targets, while a linguistic system specifically avoids such a matching. We will show that this distinction accounts for the difference in the kinds of efficacy and inefficacy usually attributed to linguistic and diagrammatic systems of representation. We hope to convince the reader that the proposed contrast is the final answer to the long-standing issue about the distinction between "linguistic" and "diagrammatic."

Chapter 2

Free Rides

This chapter has two purposes. One is to isolate and analyze the phenomenon, dubbed "free ride," which accounts for an important variety of efficacy exhibited by a wide range of representation systems. According to our analysis, the phenomenon is due to a particular way in which a structural constraint governing representations matches with a constraint governing the targets of representation. By showing the significance of this specific variety of constraint matching, we also hope to lend a good initial support to the Constraint Hypothesis, and motivate the conceptual framework (to be developed in chapter 4) that makes explicit what constraints govern representations and how they accord with constraints on the targets.

In section 2.1, we distinguish two different categories of efficacy that may be applied to representation systems. One is concerned with efficacy in using a representation as a *static* record of information. The other is concerned with efficacy in using a representation as a site for *updating* the obtained information. Hohauser (1982), Larkin and Simon (1987), Funt (1980), and Barwise and Etchemendy (1990b) point out one important variety of efficacy in the second category, which is due to a system's capability of providing a free ride from information to information. In section

2.2, we look at simple cases of free ride, and propose our analysis of the phenomenon. Section 2.3 tests the proposed analysis on more complex cases cited by the above authors. Our analysis of the free ride phenomenon remains informal in this chapter. The complete analysis will be given in chapter 4, where the conceptual framework for this dissertation is laid out formally.

2.1 The phenomenon of free ride

As we mentioned in Chapter 1, there are many different ways in which we use a representation in problem solving. We can divide these uses into two categories. One may be called the "static" use, and comprises the cases in which we use a representation as a static record of information, making no changes to its informational content during the process. Imagine that you are using an AAA map of Indianapolis to figure out how to get to the Speedway from the downtown. Except occasional circling and crossing with a pencil, you do not make any significant changes to the map. The focus is on searching and retrieving the information that is presented in the map once and for all.

The other category may be called the "dynamic" use, and comprises the cases in which we actively operate on the representation, altering its information contents successively. Imagine that you are making a list of restaurants where you and your family may want to dine for the next trip to San Francisco. You add names of restaurants, cross out some, write under some names an approximate cost for a dinner entry, and even put some tentative evaluations such as "excellent," "good," "so-so," and "not preferred." Here the focus is on updating the information in your list, on the basis of the newly obtained information or the results of your own thinking. The list is a valuable tool for the problem solving precisely because the information content of your list is constantly changed. Note, however, that a single process of problem solving may contain both the static use and the dynamic use of representations at different stages.

In the terminology of Larkin and Simon (1987), a representation “consists of both data structures and programs operating on them to make new inferences” and the “computational efficiency” of representations depends on how efficient it is (1) to search the data structure and recognize relevant information and (2) to modifying and augment the data structure to draw new inferences (p. 71). Borrowing this idea, we divide the notion of efficacy associated with representation systems into two broad categories: efficacy for the static use of a representation as a record of information, and efficacy for the dynamic use of a representation as a site for updating information.

In his book on architectural modeling, Hohauser (1982) cites one paradigmatic example that exhibits the second type of efficacy (p. 177). It is a story about a ninety-three year old man, named Harry Lieberman, who constructed a “memory map” of his home village with the help of an artist. The following is my own reconstruction of the story, rather than a historically accurate report, designed to highlight the point that we are making.

Harry is asked to describe the geographical features of the village in which he grew up, as accurately as possible, by memory. After some failed trials of recollecting the geographical features of his home village with no tools, he decides to draw an approximate map of his home town. On the basis of fragments of his memory, he draws lines and curves on a sheet of paper to represent the streets, pathways, rivers, and such. He then uses wood blocks to represent the buildings that he remembers to have existed, and places them on his map, to represent the approximate locations of those buildings. (He keeps revising and supplementing the map, and eventually obtains a map that represents his home town to the best of his memory.)

At the beginning of this procedure, Harry remembers the locations of a river, two roads, and several houses, and constructs the following tentative map:

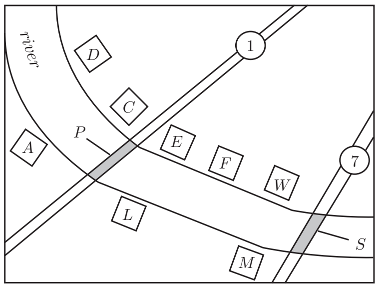


Figure 2.1

He recollects one more piece of information about his home village, that is:

(θ_1) The house K was halfway between the houses L and M.

To present this new fragment of memory in his map, Harry puts a wood block standing for the house K between the wood blocks standing for the houses L and M:

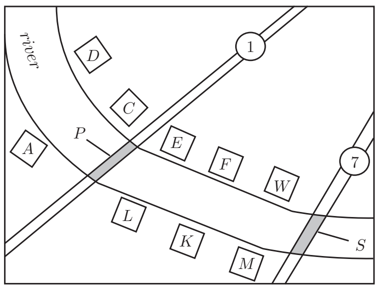


Figure 2.2

As the result of this simple operation, Harry's map now presents many pieces of new information, *other than* θ_1, that were absent from the initial map. Among them are:

(θ_2) The house K was across the house F over the river.

(θ_3) The house K was closer to road 1 than the house M was.

(θ_4) The house M was closer to the bridge S than the house K was.

(θ_5) The house K and the house A had the road 1 in between.

To get the sense of utility of the system Harry's memory map, imagine how many deduction steps would be needed if he tried to obtain the same results with pure thought on the basis of the principles of geometry. By operating on his map in the way described above, Harry has skipped all these complications of computation, and obtain the information θ_2, θ_3, θ_4, and θ_5 almost "for free." The operation is extremely efficient, for the purpose of updating the information content of his map toward the solution of the problem.

Barwise and Etchemendy (1990b) are concerned with the same kind of efficacy when they make the following remark on the utility of diagrams in reasoning (p. 22):

> Diagrams are physical situations. They must be, since we can see them. As such, they obey their own set of constraints. In our example from *Hyperproof*, when we represent tetrahedron a as large, a host of other facts are thereby supported by the diagram. By choosing a representational scheme appropriately, so that the constraints on the diagrams have a good match with the constraints on the described situation, the diagram can generate a lot of information that the user never need infer. Rather, the user can simply read off facts from the diagram as needed.

Although this passage already contains the notion of "constraint" that is of central importance to our project, let us now focus on the phenomenon that they are pointing to. Here is the example that they mention as the "example from *Hyperproof*." We are given the following picture that depicts the sizes, shapes, and locations of different blocks on a chess board[1]:

[1]Barwise and Etchemendy (1994) have developed an interactive computer program, called "Hyperproof," for teaching basic principles of analytical reasoning. This is a typical picture that shows up in a proof that a student constructs with the program.

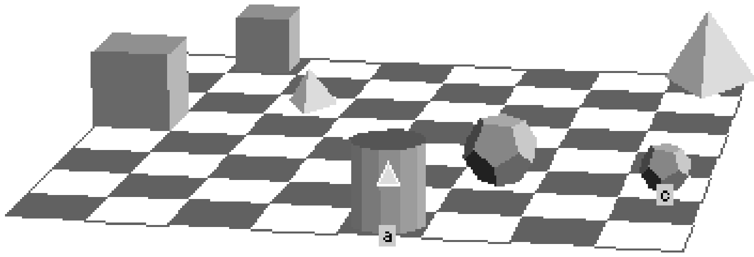


Figure 2.3

We interpret this diagram as we do for ordinary pictures. Namely, the large cube icon on the left most column of the grid picture represents an (actual) large cube on the corresponding position on the chess board, the small dodecahedron icon labeled "*c*" represents a small dodecahedron, named "*c*," on the corresponding position on the chess board. Note that the cylinder icon with a triangle symbol represents the tetrahedron *a* whose size is unknown. So, this picture has an element of indeterminacy, unlike usual photographs. Now, we change the cylinder icon in this diagram to a large tetrahedron icon, to add the following information:

(θ_6) *a* is large.

As the result, we obtain the following picture[2]:

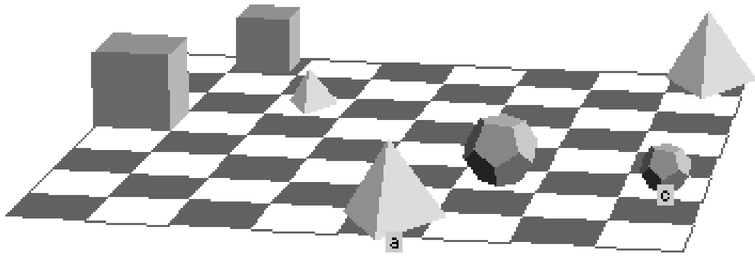


Figure 2.4

[2]This operation would take a substantial effort if it were done by hand, not with a computer. But it is not relevant to the point that we are making.

In addition to the information θ_6, our picture now presents the following information (among others) that was absent from the first picture:

(θ_7) a is larger than c.

(θ_8) a is of the same size as the cube on the leftmost column.

(θ_9) Every object in front of c is a large tetrahedron.

(θ_{10}) There are exactly two large tetrahedrons.

(θ_{11}) There are no medium tetrahedrons.

In this example, a relatively small operation of adding a single piece of information θ_6 makes the resulting picture "generate a lot of information," as Barwise and Etchemendy put it. This generation of information in turn saves a substantial number of inferences on our part—we "never need infer" the information θ_7, θ_8, θ_9, θ_{10}, and θ_{11}, but "simply read [them] off" from the resulting diagram.

Funt (1980) seems to have observed the same general phenomenon when he discusses "experimental feedbacks from diagrams." Funt presents a program, called WHISPER, that can observe and operate on diagrams to solve certain problems about the motions of collapsing objects[3]. He characterizes WHISPER in the following way (p. 51):

> Usually feedback is thought of in terms of a robot immersed in a real world environment. In WHISPER's case, however, the feedback is from a situation analogous to that in the real world—the diagram and diagram transformations—rather than from observation of actual falling objects....Using this feedback WHISPER is able to find when and where

[3]Funt describes, in some detail, how WHISPER controls a machine to "observe" and draw diagrams, but we are not concerned with these issues here.

> discontinuous changes in an object's motion occur without being forced to use sophisticated, "number-crunching" algorithms for touch tests for arbitrary shapes.

One of the "experiments" that Funt refers to is the following. WHISPER is given a diagram of a structure built of four blocks, A, B, C, and D, that is about to collapse:

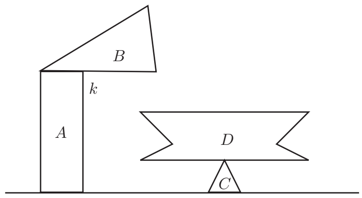


Figure 2.5

The blocks are assumed to have uniform density and thickness, and the problem is to predict the sequence of events occurring during the presumed collapse. As the qualitative solution of the problem, WHISPER outputs a set of diagrams that depict a sequence of events during the collapse. Observing the first diagram with its "retina," WHISPER determines that B is in the dominant instability, and computes a clockwise rotation of A of 28° on the pivot point k. WHISPER figures that the rotation terminates when:

(θ_{12}) The bottom surface of B inclines clockwise at 28°.

WHISPER changes the initial diagram so that it may depict the termination condition of the rotation that reflects this information:

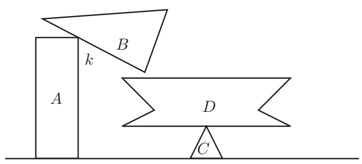


Figure 2.6

The resulting diagram presents the following information:

(θ_{13}) There is a gap between the lowest surface of B and the top surface of D.

On the basis of this information, WHISPER corrects its computation of A's rotation, and outputs the following "correct" diagram that depicts the termination condition of A's rotation:

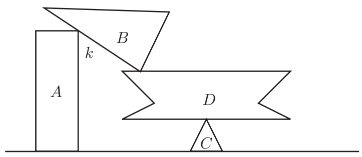


Figure 2.7

Repeating this trial and error, WHISPER produces the following snapshots of the presumed collapse before it ends its problem-solving procedure:

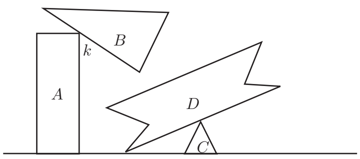


Figure 2.8

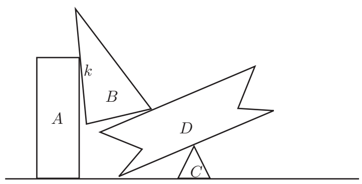


Figure 2.9

Note that θ_{12} and θ_{13} are distinct pieces of information. The information θ_{13} is the information that becomes presented in the diagram *as the result of* changing the original diagram to reflect the information θ_{12}. In Funt's terminology, θ_{13} is the "experimental feedback" of changing the diagram on the basis of the information θ_{12}. Relying on this feedback, WHISPER modifies its calculation of the initial rotation of B, and it proceeds to the next "experiment" on diagrams. According to Funt, derivation of the information θ_{13} out of the information θ_{12} in the given circumstance takes "sophisticated, 'number-crunching' algorithms."

Thus, WHISPER's experiment is another case in which an action that presents a piece of information in a representation generates an independent piece of information in the same representation, which in turn saves a significant amount of inferential task on the part of the user. Funt's example is different from the examples given by Hohauser and by Barwise and Etchemendy in that some of the snapshots produced by WHISPER depict different stages of the collapse of the blocks, rather than one and the same state of these blocks. However, the free ride that WHISPER obtains in drawing an *individual* diagram is about a single situation of the blocks. For example, the information θ_{12} and θ_{13} are both about the same situation of the blocks, although the situation is an instantaneous one.

We also find a reference to the same general phenomenon in Larkin and Simon (1987). They note two facts: (a) drawing a pair of parallel lines cut by a transversal on paper generates the information that there are eight angles, four exterior and four interior, defined by the corresponding parallel lines and transversal, and (b) drawing a rectangle and its two diagonals on paper generates the information that there is a point of intersection of the corresponding diagonals. Larkin and Simon write (p. 74):

> The process of drawing the diagram makes these new inferences which are then displayed explicitly in the diagram itself....Of course, the same information can also be inferred from the sentential representation, but these latter inference processes may require substantial computation, and the cost of this computation must be included in any assessment of the relative efficiency of the two representations.

In the same vein, Lindsay (1988) takes up the case of geometry diagrams, and discusses their ability to "permit the drawing of inferences without explicit use of rules of deduction" (p. 112). He calls a representation system with this capacity "non-deductive," and says that a nondeductive system "requires no separate computational inference-making stage: the operation of the construction process entails the 'making' of the inferences" (p. 115).

It would be useful to characterize, if tentatively, the common phenomenon that the above authors commonly point to. It is a case in which (a) the user applies a certain operation to a representation to present a given piece of information, (b) as the result of this action, the representation somehow comes to present pieces of information different from the given information, and (c) this "generation" of information saves the user the task of inferring this information from the original information.

How is it possible that an operation lets a representation "generate" information? What makes it possible for us to "experiment" on representations, with the expectation of a "feedback"? What is the mechanism in which an action on a representation saves the cost of inference through "free rides"? As the above examples testify, this phenomenon is real. Also, it apparently accounts for efficacy of many different kinds of representations in their dynamic use. Nevertheless, no elaborate analysis has been offered, except a few illuminating, yet sketchy insights offered by the above authors who pointed to the phenomenon. We will devote the rest of this chapter (and a part of chapter 4) to the analysis of this phenomenon.

2.2 An informal analysis

We will eventually analyze the rather complicated examples of the phenomenon discussed in the last section. To get a good first grasp of the phenomenon, however, let us start by looking at simple example. Consider the following two scenarios:

Example 2.1 We use Venn diagrams to check the validity of the following syllogisms:

(θ_{14}) All Cs are Bs.

(θ_{15}) No Bs are As.

(θ_{16}) (Therefore) no Cs are As.

We start with drawing three circles, labeled "As," "Bs," and "Cs" respectively. On the basis of the assumptions θ_{14} and θ_{15} of the syllogism, we shade the complement of the B-circle with respect to the C-circle (Figure 2.10) and then shade the intersection of the B-circle and the A-circle (Figure 2.11). Observing that the intersection of the C-circle and the A-circle is shaded as a result, we read off the conclusion θ_{16} of the syllogism, and decide that the syllogism is valid.

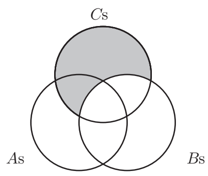


Figure 2.10

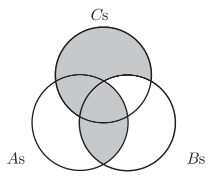


Figure 2.11

Example 2.2 We use Euler circles to solve the same problem. On the basis of the assumptions θ_{14} and θ_{15}, we draw a circle labeled "Cs" inside a circle labeled "Bs" (Figure 2.12), and then draw a circle labeled "As" completely outside the B-circle (Figure 2.13). We observe that the C-circle and the A-circle do not overlap, and read off the conclusion θ_{16} of the syllogism.

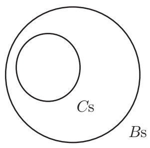


Figure 2.12

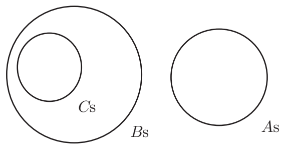


Figure 2.13

Each of these examples is an instance of the phenomenon that we identified in the last section—updating a diagram on the basis of the information θ_{14} and θ_{15} lets the diagram generate the information θ_{16}, which in turn lets us decide that the syllogism is valid. The question is: what is responsible for the generation of the information θ_{16}?

Simply stated, it is a structural constraint governing Venn diagrams (in example 2.1) or Euler diagrams (in example 2.2) that is responsible for the generation of this information. This answer has been already anticipated in Barwise and Etchemendy's passage cited in section 2.1. They point out (for the case of diagrams) that the generation of free information in a representation is due to the fact that we "choose a representational scheme appropriately, so that the *constraints on the diagrams* have a good match with the constraints on the described situation" (1990b, p. 22, italic added). Also, Lindsay (1988) points to "constraint propagation" or "constriant satisfaction" that "force the display of conclusions" as the underlying mechanism of what he calls "non-deductive inference procedures" (p. 126). The following analysis, therefore, can be taken as a substantiation of these authors' somewhat sketchy characterization of the process of free ride.

In example 2.1, we present the premises θ_{14} and θ_{15} of the syllogism in a Venn diagram. Let us look closely at what exactly occurs *on the part of diagrams* during the operation in question. In order for a Venn diagram to present this information, it must have certain properties, that is, certain states of affairs that indicate this information must hold in the Venn diagram. What are they? According to the semantic rules associated with Venn diagrams[4], a Venn diagram presents the information θ_{14} and θ_{15} if it supports the following states of affairs:

[4]Shin (1991a) has developed a logical system, with a model-theoretic semantics, that formalizes one of the standard ways in which we use Venn diagrams in solving syllogisms. Our discussion of example 2.1 is based on the semantic rules for Venn diagrams formalized in her paper.

(σ_1) The complement of a circle labeled "*B*s" with respect to a circle labeled "*C*s" is shaded.

(σ_2) The intersection of the circle labeled "*B*s" and a circle labeled "*A*s" is shaded.

In fact, the operations that we take in the example are ones that realize these states of affairs in a Venn diagram. Now let us ask ourselves: is there anything that follows, *on the part of the diagram*, from these two facts? Well, as a matter of geometrical constraint, whenever the states of affairs σ_1 and σ_2 hold in a Venn diagram, the following state of affairs *must* hold in the Venn diagram:

(σ_3) The intersection of a circle labeled "*C*s" and a circle labeled "*A*s" is shaded.

Notice that this fact is a "side-effect" of the operations that we have taken: the operations are taken in order to realize the states of affairs σ_1 and σ_2, not σ_3. Nevertheless, on the semantic convention associated with Venn diagrams, this fact has an independent semantic value, namely, the information θ_{16}. We get the information θ_{16} "for free," and decide that the syllogism is valid.

To express the mechanism of this "free ride" schematically:

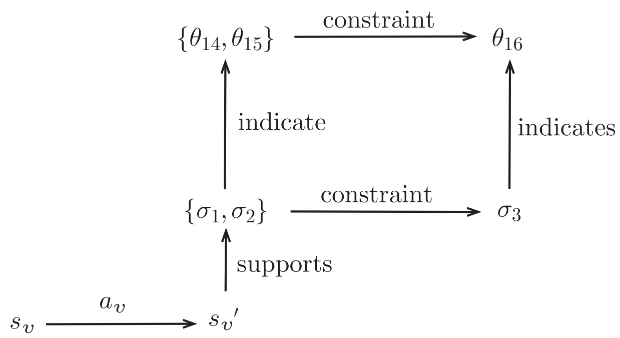


Figure 2.14

Here a_v is the operation (or sequence of operations) that we apply to present the set of assumptions $\{\theta_{14}, \theta_{15}\}$. The operation a_v changes a (blank) diagram s_v to another diagram $s_v{}'$, which supports the states of affairs σ_1 and σ_2. Since σ_1 and σ_2 respectively indicate the information θ_{14} and θ_{15} on our semantic convention, we say that $s_v{}'$ presents the assumptions $\{\theta_{14}, \theta_{15}\}$. Now, as a matter of geometrical constraint, the set of states of affairs $\{\sigma_1, \sigma_2\}$ entails the state of affairs σ_3, which in turn indicates the information θ_{16}. Thus, to check the validity of the syllogism, we do not have to infer the conclusion θ_{16} from the assumption $\{\theta_{14}, \theta_{15}\}$ on our part. The constraint governing the Venn diagram takes over the necessary inference.

Note that this "inference" is *valid* indeed. The conclusion θ_{16} that we obtain for free is in fact a consequence of the assumptions $\{\theta_{14}, \theta_{15}\}$ that we start with. In other words, there is a constraint, governing the relationships among the sets of As, Bs, and Cs, that makes θ_{16} a consequence of $\{\theta_{14}, \theta_{15}\}$. Note the match between this constraint on the target of representation (depicted in the upper part of Figure 2.14) and the constraint on the representation $s_v{}'$ (depicted in the lower part). This match of constraints guarantees our free ride to be valid, and partly explains the efficacy of the system of Venn diagrams for checking the validity of syllogisms.

Example 2.2 admits a similar analysis. We start with presenting the set of premises $\{\theta_{14}, \theta_{15}\}$ of the syllogism in an Euler diagram. We do this by realizing the following states of affairs in an Euler diagram:

(σ_4) A circle labeled "Cs" is inside a circle labeled "Bs."

(σ_5) A circle labeled "As" is completely outside the circle labeled "Bs."

Notice that these states of affairs indicate the desired information θ_{14} and θ_{15} according to the semantic rules we adopt for the Euler diagrams. Now, as a matter of geometrical constraint governing Euler circles, these states of affairs force the following state of affairs to hold in the same diagram:

(σ_6) A circle labeled "*C*s" appears completely outside a circle labeled "*A*s."

Although the fact σ_6 is a side effect of the operation we take, it has an independent semantic value, and lets us read off the conclusion θ_{16} of the syllogism. Again, we get a piece of information for free. We can express the mechanism of this free ride in the following schematic way, letting a_e be the operation (or the sequence of operations) that we take in this example, and s_e and $s_e{}'$ be the Euler diagrams before and after the operation a_e:

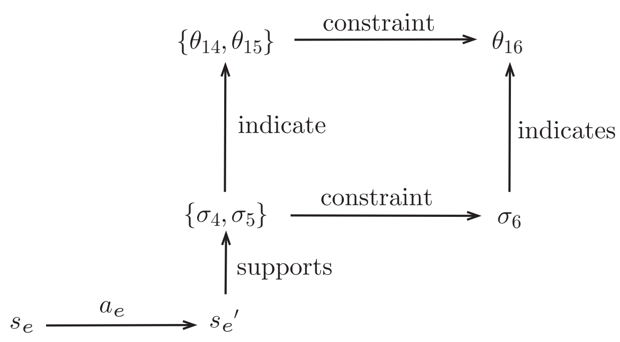


Figure 2.15

In each of examples 2.1 and 2.2, we get a free ride from the premises to the conclusion of the syllogism. The constraints responsible for the free rides are different in the two scenarios—in example 2.1, the relevant constraint is one that governs the shadings of different areas of overlapping circles (the constraint from $\{\sigma_1, \sigma_2\}$ to σ_3), while in example 2.2, the relevant constraint is one that governs the enclosure-disclosure relations among circles of different sizes (the constraint from $\{\sigma_4, \sigma_5\}$ to σ_6). However, these different constraints spare us the same deduction steps—someone using the standard first-order calculus would have to go through two applications of modus ponens and a universal generalization.

We now describe these process of free ride in general terms. Let a be an operation (or a sequence of operations) applied to a representation s and resulting in a representation s'. The operation a triggers a *free ride* from the set Θ_1 of assumptions to the information θ under the following conditions:

1. The operation a realizes in the representation s' a set Σ_1 of states of affairs which, on the semantic rules associated with s', indicate the set Θ_1 of assumptions.

2. A structural constraint from Σ_1 to σ holds on the representation s'. (Hence s' supports the extra state of affairs σ.)

3. On the semantic rules associated with s, the state of affairs σ indicates the information θ.

4. A constraint from Θ_1 to θ holds on the target of representation.

On this analysis, the utility of a free ride consists in the fact that it lets us update a representation s to s' in the way that saves us the valid inference from the assumptions Θ_1 to the information θ (although the reasoner still has to read off θ from the representation s'). This process crucially depends on the constraint from Σ_1 to σ holding on the representation s'. To wit, suppose there were no such constraint. Then, realizing the states of affairs Σ_1 in a representation s' would only result in Σ_1 in s', and nothing else. No additional facts would hold in s', and hence no additional information could be read off from the representation.

The utility of this process also depends on the existence of the constraint from Θ_1 to θ on the target of representation. Without it, the conclusion θ that we obtain for free in this process may not hold of our target, even if the assumptions Θ_1 do hold of it. We do get a ride, but it is a ride to an invalid conclusion. Hence, it is not enough that we have a constraint from Σ_1 and σ for a correct free ride. The constraint from Θ_1 to θ, to which it corresponds via the indication relation, must actually hold on the target of representation. It is in this particular sense of "matching" in which we claim that the free ride phenomenon is due to a matching between a constraints on representations and a constraint on targets.

2.3 Testing the analysis

At the core of our analysis is the observation that the phenomenon of free ride is attributable to a particular matching between a constraint on representations and a constraint on targets. We now test our analysis on the various examples that we have seen in section 2.1.

The system of Harry's image map

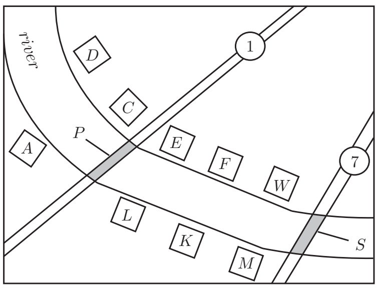


Figure 2.2

Let us start with the example, cited by Hohauser, where Harry updates a memory map to another by adding the information θ_1:

(θ_1) The house K was halfway between the houses L and M.

The resulting map (Figure 2.2, reproduced above) generates various information. Let us consider the generation of the information θ_2:

(θ_2) The house K was across the house F over the river.

According to our analysis given above, there must be some structural constraint on the map that is responsible for the generation of this information. In fact, there is such. It is the constraint from:

(σ_7) The block F is halfway between the blocks E and W.

(σ_8) The block L is across the block E over the band labeled "river."

(σ_9) The block M is across the block W over the band labeled "river."

(σ_{10}) The part of the band labeled "river" is straight between the blocks L and M.

(σ_{11}) The block K is halfway between the blocks L and M.

to:

(σ_{12}) The block K is across the block F over the band labeled "river."

The reader should convince himself or herself that this is a reasonable constraint to assume to hold of the map. Note that the states of affairs $\sigma_7, \ldots, \sigma_{10}$ already hold in the initial map. Harry's operation preserves these states of affairs while realizing σ_{11} in the resulting map. According to the semantic rules associated with this memory map, the states of affairs $\sigma_7, \ldots, \sigma_{10}$ and σ_{11} indicate the following information respectively:

(θ_{17}) The house F is halfway between the houses E and W.

(θ_{18}) The house L is across the house E over the river.

(θ_{19}) The house M is across the house W over the river.

(θ_{20}) The part of the river is straight between the houses L and M.

(θ_1) The house K is halfway between the houses L and M.

In our scenario, this much is a part of what Harry has already recollected about his home village. Now, due to the constraint specified above, the updated map also supports the state of affairs σ_{12}, which in turn indicates the information θ_2. This information is not a part of what Harry has recollected, although it is a consequence of what he has. Thus, we obtain a free ride from the set of assumptions $\{\theta_{17}, \theta_{18}, \theta_{19}, \theta_{20}, \theta_1\}$ to the information θ_2, exactly in the way that our general analysis of the phenomenon specifies. Here the structural constraint from $\{\sigma_7, \sigma_8, \sigma_9, \sigma_{10}, \sigma_{11}\}$ to $\sigma 12$ on Harry's map matches with the constraint from $\{\theta_{17}, \theta_{18}, \theta_{19}, \theta_{20}, \theta_1\}$ to θ_2 via the indication relation. Our analysis applies also to the free rides to the other pieces of information θ_3, θ_4, and θ_5.

The system of Hyperproof diagram

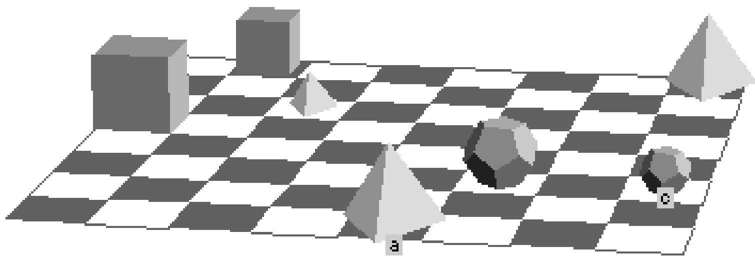


Figure 2.4

In this example, we take an operation that adds the information θ_6 to a Hyperproof diagram.

(θ_6) a is large.

The operation lets the resulting diagram (Figure 2.4, reproduced above) present the information θ_7, θ_8, θ_9, θ_{10}, and θ_{11}. We consider how the information θ_7 is generated:

(θ_7) a is larger than c.

We can identify the relevant constraint on the Hyperproof diagram as the one from the set of states of affairs:

(σ_{13}) The icon labeled "c" is small.

(σ_{14}) The icon labeled "a" is shaped as a tetrahedron, a cube, or a dodecahedron and it is large.

to the state of affairs:

(σ_{15}) The icon labeled "a" is shaped as a tetrahedron, a cube, or a dodecahedron and it is larger than the icon labeled "c."

As in the case of Harry's memory maps, the state of affairs σ_{13} already holds in the Hyperproof diagram before the update operation. Our operation preserves this state of affairs while adding another state of affairs σ_{14}. According to the semantic convention associated with Hyperproof diagrams, the state of affairs σ_{13} indicates the information:

(θ_{21}) c is small.

while σ_{14} indicate:

(θ_6) a is large.

Note that to indicate the information θ_6, it is not enough that an icon labeled "a" is large. As we mentioned earlier, the cylinder icons in Hyperproof diagrams are interpreted as inderterminate about the sizes of the blocks that they stand for. Thus, if an icon is shaped as a cylinder, then even if it is large, it does not indicate that the corresponding block is large, according to the Hyperproof semantics. This is why we say that the information σ_{15} is indicated by a somewhat complex state of affairs σ_{14}, not by the simpler state of affairs that the icon labeled "a" is large.

Now, due to the structural constraint on Hyperproof diagrams specified above, the updated diagram also supports the state of affairs σ_{15}, which in turn indicates the information θ_7. The information θ_7 is a consequence of the set of assumptions $\{\theta_{21}, \theta_6\}$ that we started with. We thus get a free ride, exactly as our analysis prescribes. Similar accounts apply to the free rides to the other information θ_8, θ_9, θ_{10}, and θ_{11}.

The system of Whisper diagram

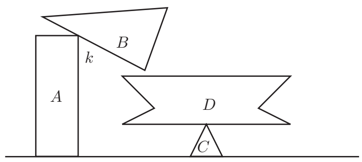


Figure 2.6

In updating the initial diagram to the second diagram (Figure 2.6, reproduced above), Whisper obtains a free ride to the following information:

(θ_{13}) There is a gap between the lowest edge of B and the surface of D.

What structural constraint on *Whisper*'s diagrams is responsible for this free ride? It is the constraint from the states of affairs:

(σ_{16}) The point "k" is 1/4 inch leftward, and 3/16 inch upward from the top edge of the hexagonal icon "D."

(σ_{17}) The bottom edge of the triangular icon "B" is 5/8 inch long.

(σ_{18}) The top edge of the hexagonal icon "D" is 7/8 inch long.

(σ_{19}) The point "k" divides the bottom edge of the triangular icon "B" into 7:9.

(σ_{20}) The bottom edge of the triangular icon "B" inclines at 28° clockwise.

to the state of affairs:

(σ_{21}) The right end of the the bottom edge of the triangular icon "B" does not touch the top edge of the hexagonal icon "D."

(Again, the reader should check if this is a constraint likely to hold on the WHISPER diagram.) The states of affairs σ_{16}, σ_{17}, σ_{18}, and σ_{19} indicate the following information:

(θ_{22}) The point "k" is 4 inch leftward, and 3 inch upward from the top edge of the hexagonal block D.

(θ_{23}) The bottom edge of the triangular block B is 10 inch long.

(θ_{24}) The top edge of the hexagonal block "D" is 14 inch long.

(θ_{25}) The point "k" divides the bottom edge of the triangular block "B" into 7:9.

while the state of affairs σ_{20} indicates the information:

(θ_{12}) The bottom edge of the triangular block "B" inclines at 28° clockwise.

Presumably, the information θ_{13} is a consequence of the assumptions $\{\theta_{22}, \theta_{23}, \theta_{24}, \theta_{25}, \theta_{12}\}$. Thus, this is case of free ride, as we characterize it.

The system of geometry diagram

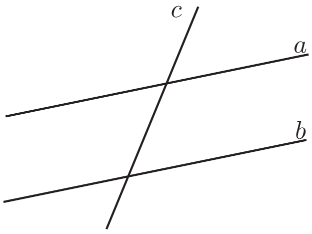


Figure 2.16

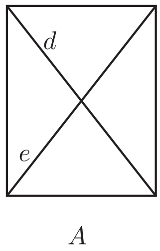


Figure 2.17

Larkin and Simon's examples are concerned with diagrams drawn in geometry proofs. We can think of such diagrams as depicting points, lines, and figures on the abstract space whose characteristics are defined by Hilbert's axiomatic theory of geometry. In particular, Figure 2.16 presents the following information about three lines on Hilbert's space:

(θ_{26}) The lines a and b are parallel.

(θ_{27}) The line c cuts both a and b.

What states of affairs in the diagram indicate these pieces of information? They are:

(σ_{22}) The line icons labeled "a" and "b" are approximately parallel.

(σ_{23}) The line icon labeled "c" cuts both line icons labeled "a" and "b."

(By "line icons" we mean the black, thin, approximately straight marks on the paper that you are looking at. They must be distinguished from the lines on the Hilbert's space that they stand for.)

Now, as a matter of structural constraint governing the diagram, the state of affairs σ_{22} and σ_{23} holding in the diagram yield the following state of affairs in the same diagram:

(σ_{24}) The line icons labeled "a," "b," and "c" make eight angles, four interior and four exterior.

This state of affairs in turn indicates the following information:

(θ_{28}) The lines a, b, and c make eight angles, four interior and four exterior.

Thus, due to the structural constraint from $\{\sigma_{22}, \sigma_{23}\}$ to σ_{24}, we obtain a free ride from the assumptions $\{\theta_{26}, \theta_{27}\}$ to the information θ_{28}. This free ride is again a correct one—there is a constraint on Hilbert's space that makes θ_{28} a consequence of $\{\theta_{26}, \theta_{27}\}$ (as we can ascertain from the fact that the former is provable from the latter in Hilbert's axiom system).

We can give a similar analysis to the other example provided by Larkin and Simon, where drawing a rectangle (labeled "A" in Figure 2.17 above) and its two diagonals (labeled "d" and "e") on paper generates the following information:

(θ_{29}) There is a point of intersection of the segments d and e.

The structural constraint on diagrams responsible for this free ride is the one from the states of affairs:

(σ_{25}) The segment icons labeled "d" and "e" are the diagonals of the rectangular icon labeled "A."

to the state of affairs:

(σ_{26}) There is a point of intersection of the segment icons labeled "d" and "e."

Note that θ_{29} is a piece of information about Hilbert's space while σ_{26} is a state of affairs holding in the diagram that we have drawn. We leave the reader to elaborate the analysis.

2.4 Conclusion

We started this chapter by isolating a phenomenon, dubbed "free ride," that often occurs when we operate upon representations dynamically, updating their information contents in the process of problem solving. We have seen that the free ride phenomenon accounts for an important variety of efficacy in such processes, in that it spares the reasoner a number of inferences that would be otherwise needed. By examining a number of instances of the phenomenon, we have shown that the capability of providing a free ride accounts for the efficacy of a wide range of representation systems, including the system of Venn diagrams, of Euler diagrams, of certain maps, of Hyperproof diagrams, of Funt's block diagrams, and of geometry diagrams. In the core part of the chapter, we proposed an analysis of the phenomenon, and showed that the phenomenon heavily depends on the existence of a particular structural constraint governing the representations, and its match with a constraint governing the target of relation.

Thus, this chapter lends a partial support to the general hypothesis, the Constraint Hypothesis, that we are advocating throughout this dissertation. At the same time, our result in this chapter should provide a motivation for chapter 4, where we develop a conceptual framework that is specifically designed to capture the structural constraints governing representations and the ways they match with the constraints on the targets of representations. We will use the framework to offer a more precise analysis of the capability of a representation system to provide a free ride.

Before closing our informal discussion of free rides, let us sharpen our understanding of the phenomenon by looking at rather different instances of the phenomenon than we have seen so far. Think of the cases of so-called "hardware simulation" where people use dummies for auto-accident tests, scale models of buildings for ventilation tests, and miniature river systems for hydrological tests. In such a simulation, the testers create a situation (the simulating situation) that stands for the real world situation (the simulated situation) whose behaviors under certain circumstances we wish to know. The testers (1) adjust the simulating situation so that it may represent the circumstances of the simulated situation that we are interested in, and then (2) observe what happens in the simulating situation. The steps (1) and (2) correspond to the steps of free ride in which we ($1'$) operate on a representation so that it represents our assumptions, and ($2'$) observing the results produced under the structural constraints governing the representation. Thus, cases of accurate simulations correspond to cases of correct free rides, and cases of inaccurate simulations to cases of incorrect free rides. From this perspective, then, the cases of free rides that we cited in section 2.1 are special instances of simulations that use diagrams as the simulating situations. Funt used the expression "experimenting on diagrams." More accurately, we were simulating with diagrams.

Chapter 3

Content specificity

Some representations are more specific in their informational content than other representations. For example, the diagram of a triangle drawn on paper is more specific in its informational content than the sentence, "There is a triangle." The former specifies (at least) the relative sizes of the three sides and the three angles of the triangle in question while the latter does not. Now, there are important cases in which *all* representations in a certain representation system *have to* be content specific in certain respects—cases in which they *cannot* present certain chunks of information in isolation, without adding certain other information. For example, no diagrams commonly used in geometry can represent a figure as a triangle without specifying the relative lengths of the three sides (at least on a layman's interpretation). This chapter is concerned with this phenomenon of "system-wide" content specificity. Why is it not possible for a representation in a certain system to present a certain set of information in isolation? Under what conditions does a system of representation exhibit this limitation?

In section 3.1, we cite a number of examples of the phenomenon, and try to identify the phenomenon more precisely. The examples are also intended to show the ubiquity of the phenomenon. We will illustrate how the property of content specificity accounts for a particular kind of *inefficacy* of the representation system. In section 3.2, we consider a few simple examples in detail, and propose an informal analysis of the phenomenon, just as we did in section 2.2 of the last chapter. In section 3.3, we return to some of the examples cited in section 3.1, and test our analysis on them. At the core of our analysis is the claim that the content specificity of a representation system is due to structural constraints governing the representations in the system, and to a particular way they fail to match with the constraints on the targets. Thus, our analysis in this chapter combines with the analysis in the last chapter to lend a strong support for the Constraint Hypothesis and give a motivation for the conceptual framework to be developed in the next chapter.

3.1 The phenomenon of content specificity

Let us look again at the case of Harry's memory maps, cited in Hohauser (1982). As in the last chapter, we are mainly concerned with what *would* have happened in the use of such maps. The following scenario is by no means a historically accurate report of what Harry Lieberman actually did.

Recall that Harry has constructed the following memory map on the basis of his somewhat fragmentary memory about his home village:

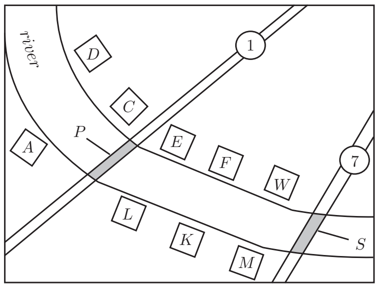


Figure 3.1

Harry now recalls that:

(θ_1) The house B was somewhere between the house A and the house K

Before forgetting it again, he wants to record this information in his map by placing the B block between the A block and the K block. However, this requires him to put the B block either between the A block and the road line 1, or between the road line 1 and the L block, or between the L block and the K block. He cannot decide which alternative to take because his memory does not tell him which of the following was true:

(θ_2) The house B was between the hose A and the road 1

(θ_3) The house B was between the road 1 and the house L

(θ_4) The house B was between the house L and the house K

Harry's trouble comes from the content specificity of the system of his memory maps. Let Θ_m be the set of pieces of information presented in the map depicted in figure 3.1. (As before, we assume that there are fixed rules of interpreting his memory maps.) The problem is that Harry's memory maps cannot present the set of information $\Theta_m \cup \{\theta_1\}$ in isolation, without presenting either θ_2 or θ_3 or θ_4 at the same time. In other words, any representation in Harry's memory map system must be specific about the choice among θ_2, θ_3, and θ_4 if it is to present the information $\Theta_m \cup \{\theta_1\}$. Stenning and Oberlander (1995) aptly characterize the content specificity as "the demand by a system of representation that information in some class be specified in any interpretable representation" (p. 98). We borrow a part of this idea, and say that the system of Harry's memory maps is *over-specific in presenting* $\Theta_m \cup \{\theta_1\}$.

Now, the efficacy of a representation system partly depend on its ability to present, in an isolating manner, *whatever* chunk of information that has been obtained in the course of reasoning. In Levesque's words (1988), the efficacy of a system depends on "what it allows you to leave unsaid" as well as "what it allows you to say" (p.370). In this respect, Harry's system of memory maps is ineffective due to the content specificity that we have just seen. We *can* present the information $\Theta_m \cup \{\theta_1\}$ in a map, but cannot present it in isolation, abstracting away from the alternatives θ_2, θ_3, and θ_4. In Stenning and Oberlander's words, the system "limits abstraction" (p. 97). The content specificity of the system leads directly to inefficacy in the this particular respect.

Furthermore, the over-specificity of Harry's system can lead the representations to hold so-called "accidental features." Although Harry actually knows better, let us suppose he went ahead and placed the B block between the A block and the K block, choosing a particular location, say, between the A block and the road line 1. The following map would result:

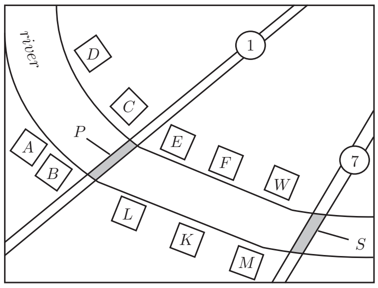


Figure 3.2

This map support not only the state of affairs:

(σ_1) The block B between the A block and the road line 1

but also:

(σ_2) The block B is across the C block over the “river” band.

(σ_3) The block B is closer to the bridge icon P than the L block is.

(σ_4) The block A is closer to the B block than the L block is.

Note that the states of affairs σ_2, σ_3, and σ_4 are all consequences of Harry's arbitrary choice of putting the B block between the A block and the road line 1—they are *accidental features* of the map. Nevertheless, they have the following independent semantic values respectively:

(θ_5) The house B is across the house C over the river.

(θ_6) The house B is closer to the bridge P than the house L is.

(θ_7) The house A is closer to the house B than the house L is.

This is why the above accidental features are potentially dangerous. In the course of his problem solving, Harry may forget that σ_1, σ_2, σ_3, and σ_4 are accidental features, and may "appeal to" one of them to base his future reasoning on the information indicated by it.

The issue of content specificity has a long history in philosophy, although it is mainly discussed in connection to internal representations, rather than external representations. Consider Berkeley's famous discussion of the nature of mental representations in the introduction to his *Treatise* (1710). In demonstrating why our mental representations cannot be "abstract" in Locke's sense, Berkeley claims that it is impossible to imagine a hand or an eye without imagining some particular shape or color of it or to frame the idea of a man without making him either black or white or tawny, straight or crooked, tall or low or middle-sized (paragraph 10). Likewise, we can never attain to have the general idea of a triangle that is 'neither oblique nor rectangle, equilateral nor scalenon, but all and none of these at once"' (paragraph 13). Pursuing the same line of argument, Hume (1739) contends, "it is impossible to form an idea of an object that is possessed of quantity and quality, and yet is possessed of no precise degree of either" (Book 1, section VII). From these observations, Berkeley and Hume conclude that no mental representations are abstract.

Pylyshyn (1973) discusses the content specificity of mental representations with an assumption opposite to Berkeley's and Hume's conclusion—he has an empirical evidence that mental representations are often *not* specific. Thus, for example, "it would be quite permissible...to have a mental representation of two objects with a relationship between then such as 'beside,"' while "such a representation need not contain a more specific spatial-relation term such as 'to the left of' or 'to the right of"' (p. 11). Starting with this evidence, Pylyshyn argues that mental representations cannot be anything like images, because "it would seem to be an unreasonable use of the word 'image,' however, to speak of an image of two objects side by side without one of the relations between them being either 'to the left of' or 'to the right of."' In the same spirit, Dennett (1960) argues (p. 136):

> Consider the Tiger and his Stripes. I can dream, imagine or see a striped tiger, but must the tiger I experience have a particular number of stripes? If seeing or imagining is having a mental image, then the image of the tiger *must*—obeying the rules of images in general—reveal a definite number of stripes showing, and one should be able to pin this down with such questions as 'more than ten?', 'less then twenty?'.

Since Berkeley and Hume assume that all mental representations are image-like, they conclude that mental representations cannot be abstract (opposing to Locke). Since Dennett and Pylyshyn know, as an empirical fact, that mental representations are often *un*specific in their contents, they conclude that they cannot be image-like (opposing to pictorialists such as Kosslyn and Pomerantz (1977) and Shepard and Cooper (1982)). What is common among all these philosophers is the observation that certain external representations, commonly called "pictures" or "images," exhibit particular content specificity—the observation that those representations cannot present a certain combination of information in isolation, without thereby presenting additional pieces of information. The arguments of those philosophers would collapse, if images and pictures did not have such properties.

As we have seen in the case of Harry's memory maps, the over-specificity of a system of representations lead to its inflexibility of expressing arbitrary chunk of information. The diagrams used in Barwise and Etchemendy's Hyperproof (1994) exhibit this limitation (although the system has natural ways of circumventing the weakness).[1] The logical system underlying Hyperproof is a heterogeneous system that allows interactions of diagrammatic representation and sentential representations in a single proof. Suppose that we have obtained the following diagrams and sentences in a proof:

[1]Barwise and Etchemendy (1990b) themselves discuss the expressive limitation of Hyperproof diagrams (p. 22). I am making the same point as theirs with a slightly different example.

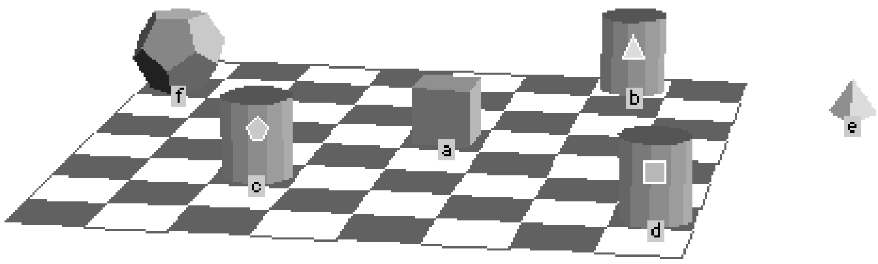


Figure 3.3

- $Small(d)$
- $SameCol(e, a)$
- $SameSize(b, f) \vee SameSize(c, f)$

Hyperproof has an rule of inference, called "Apply," that allows us to incorporate the information content of a sentence into a diagram. How do we "Apply" the first sentence into the given diagram? Well, we can simply change the size of the cylinder icon labeled "*d*" to a small icon. It is easy and intuitive. What about the second sentence? We know that we should put the tetrahedron icon labeled "*e*" somewhere in the same column as the cube icon labeled "*a*." Unfortunately, we cannot decide on exactly where. Similarly for the third sentence. We know that we should change the cylinder icon labeled "*b*" or the cylinder icon labeled "*c*" into the same size as the cube icon labeled "*f*," but we do not know which to change. This means that we cannot incorporate the information content of the second or the third sentence into the above Hyperproof diagram, without making an arbitrary choice about exactly where the block e is located or about which of the blocks b and c is of the same size as the block f. To be more precise, let Θ_h be the information content of the first Hyperproof diagrams, and θ_{h1}, θ_{h2}, and θ_{h3} be the information contents of the three sentences in that order. Then, our observation is that no Hyperproof diagram can present the chunk of information $\Theta_h \cup \{\theta_{h2}\}$ or the chunk $\Theta_h \cup \{\theta_{h3}\}$ in isolation, wheareas we *can* present the chunk $\Theta_h \cup \{\theta_{h1}\}$. The Hyperproof diagrams are inflexible in these respects.

The example of Harry's memory maps also showed that when a representation system exhibits over-specificity, there is a chance that representations in the system support accidental features in the process of reasoning. Accidental features abound especially in the case of geometry diagrams, and it is standard in high school classes in geometry that a proof appeals to an accidental feature of the diagram being used, and becomes invalid. Berkeley (1710) is alluding to such a fallacy when he notes, "Having demonstrated that the three angles of an isosceles rectangular triangle are equal to two right ones, I cannot therefore conclude this affection agrees to all other triangles which have neither a right angle nor two equal sides" (Introduction, paragraph 16). In the same vein, Barwise and Etchemendy (1990b) discuss the possibility of appealing to accidental feature of the following diagram used in a proof of the Pythagorean Theorem:

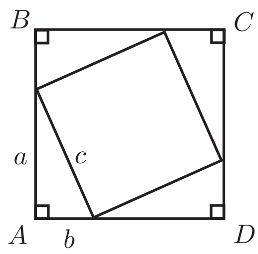


Figure 3.4

Barwise and Etchemendy note, "if any piece of our reasoning had appealed to the observation that, in our diagram, a is greater than b, the proof would have been fallacious, or at any rate would not have been as general as the theorem demands"(p. 14).

Note that the accidental features mentioned in these examples have their root in the over-specificity of the system of geometry diagrams that is commonly used. The diagram in Berkeley's example has the features of being rectangular and equilateral *because* no geometry diagrams represent a figure as a triangle without representing it as having particular angles and particular sides. Of course, we do not have to represent it as a right equilateral triangle, but in any case, we have to make particular choices about the sizes of angles and sides if we are to draw a triangle at all. Likewise, the diagram in Barwise and Etchemendy's example did not have to represent the triangle so that a may be greater than b. But all geometry diagrams have to make a choice anyway about the relative lengths of the three size of the right triangle. This is due to the over-specificity of the system of geometry diagrams about the lengths of the sides of the triangle to be represented.

Of course, the above remarks are only true of the cases in which geometry diagrams are interpreted *on the semantic rules that a layman would adopt.* After a few courses of highschool geometry, people seem to develop a sophisticated way of interpreting geometry diagrams, and learn to ignore the particular features that may mislead the proofs. Relative to such interpretation rules, the system of geometry diagrams is not over-specific as we describe above. Luengo (1995) formulates a logical system that formalizes the rules for such valid use of geometry diagrams. Although this is an interesting topic, our aim here is not to investigate such a system of geometry diagrams, but to point to the over-specificity based on the unsophisticated, yet prevalent interpretation of geometry diagrams.

We now have seen a number of examples in which the representations in a given system cannot present particular combinations of information in isolation. What is, then, the mechanism behind this phenomenon of over-specificity? Why is it not possible for a representation in a system to present certain sets of information in isolation?

Note that we are not concerned with what it is for a representation to be more or less specific in its content. For this question, it is perhaps sufficient to say that it means that the representation presents more or less information, with the notion of "presenting" understood in one's favorite view of the relation of presenting. Or if one is more logically minded, one might want to develop a model theory for the representation system in question, and say that a representation is more or less specific in informational content if there are greater or lesser numbers of models that make it true. In fact, Stenning and Oberlander (1995) takes this line of characterizing content specificity of representations. This, however, does not answer the question that we are asking. Our question is not concerned with the content specificity of *particular* representations. It is concerned with the exact conditions under which *all* representations in a system have to be content specific in certain respects.

One might answer this question immediately, by saying that it all depends on the syntactic and semantic rules stipulated in the representation system in question. But what does it mean that it all *depends* on the syntactic and semantic rules? Of course, without such stipulations, there are nothing that regulate the formation and interpretation of representations. So, trivially, the over-specificity of a system depends on such rules. But if the claim is that *every* case of over-specificity can be explained *purely* as a matter of syntactic and semantic stipulations, it is false. Of course, there is a species of over-specificity that is purely the result of the syntactic and semantic stipulations in the relevant system. For example, Stenning and Oberlander (1995) consider several representation systems of this kind. One of them is a first-order language in which only certain conjunctions of atomic sentences count as well-formed. In this system, the representations are all over-specific purely as the result of this syntactic stipulation. In fact, there is only one model for each well-formed sentence in this system.

However, this is rather a special case of over-specificity, and cannot be taken as the representative of all cases. Take the system of geometry diagrams in which no diagrams can represent a figure just as a triangle without representing it as a particular triangle. Is *this* a part of the syntactic and semantic rules associated with the geometry diagrams? Is it even a logical consequence of such rules? Or take the system of Hyperproof diagrams. Did anybody stipulate that Hyperproof diagrams cannot represent a block as being between two other blocks without representing it as a specific position on the chess board? Obviously, there are many instances in which the over-specificity of a representation system is neither a part nor a logical consequence of the explicit syntactic and semantic stipulations in the system. More pointedly formulated, then, our question is: given the syntactic and semantic stipulations of a representation system, what is the extra factor that forces particular over-specificity onto the representations in the system?[2]

3.2 An informal analysis

In this section, we will propose an informal analysis of the phenomenon of content specificity. We will follow the practice of chapter 2, and tackle the question by first looking at simple instances of the target phenomenon.

Example 3.1 We use Euler diagrams to see what conclusions can be deduced from the following premises:

[2]I do not think that Stenning and Oberlander are committed to the view that I have just criticized. Rather, they are pursuing a different question than ours when they discuss representation systems with stipulated over-specificity. However, since they do not explicitly discuss different ways in which over-specificity is imposed on representations, they fail to support their claim that content specificity is the feature distinguishing graphical and linguistic representations (p. 98). We will return to this point in chapter 5.

(θ_8) All Cs are Bs.

(θ_9) No Bs are As.

(θ_{10}) All Bs are Ds.

Noticing that the first two premises θ_8 and θ_9 are the same as what we had in example 2.1 in the last chapter, we start with drawing the same diagram as in that example (Figure 3.5).

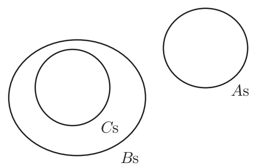


Figure 3.5

We now want to present the premise θ_4 in the diagram. According to the semantic convention we adopt for Euler diagrams, this requires us to enclose the circle labeled "Bs" in a circle labeled "Ds." However, such a D-circle must either have no overlap with the A-circle (Figure 3.6) or have some overlap with it (Figure 3.7).[3] Given our semantic convenion, the two alternatives indicate the following information respectively:

(θ_{11}) Some Ds are As.

(θ_{12}) No Ds are As.

Since neither follows from the given set of premises $\{\theta_8, \theta_9, \theta_{10}\}$, we cannot take an action required to present the premise θ_{10}.

[3]Actually, this alternative has two sub-alternatives: that a circle labeled "Ds" partially overlaps with a circle labeled "As" and that A circle labeled "Ds" encloses a circle labeled "As." Figure 3.7 is for the first sub-alternative. However, our example will serve as an example of over-specificity without taking these sub-alternatives into account.

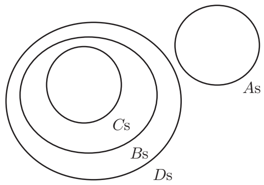


Figure 3.6

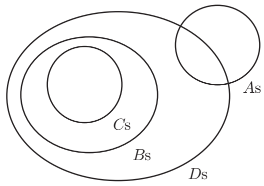


Figure 3.7

Example 3.2 This time, we are given the following set of premises:

(θ_8) All Cs are Bs.

(θ_9) No Bs are As.

(θ_{13}) All As are Ds.

The first two premises are the same as those in the last example, so we start with drawing the diagram in Figure 3.5. We now want to present the premise θ_{13} in the diagram. On our semantic convention, this requires us to enclose the A-circle in a circle labeled "Ds." However, such a D-circle must either have no overlap with the B-circle and the C-circle (Figure 3.8), or overlap only with the B-circle (Figure 3.9), or overlap with both (Figure 3.10).[4] The three alternatives indicate independent pieces of information:

(θ_{14}) No Ds are Bs and no Ds are Cs.

(θ_{15}) Some Ds are Bs but no Ds are Cs.

(θ_{16}) Some Ds are Bs and some Ds are Cs.

Sine none of these follow from the given set of premises $\{\theta_8, \theta_9, \theta_{13}\}$, we have to give up adding the premise θ_5 to the diagram.

[4]Each of the last two alternatives has two sub-alternatives, concerning enclosure and partial overlapping between circles. Again, it is not essential to list them separately for our purpose.

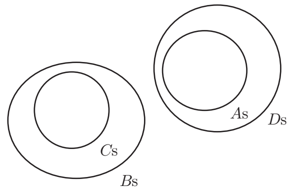
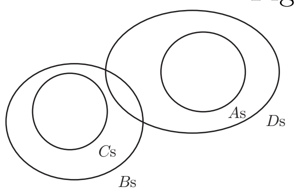


Figure 3.8



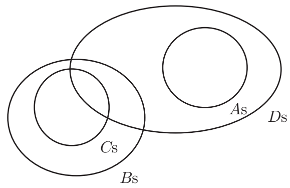


Figure 3.9 Figure 3.10

Both examples are instances of the phenomenon of over-specificity we are concerned with. The first example points to the over-specificity of the system of Euler diagrams about the alternatives θ_{11} and θ_{12} in presenting the information $\{\theta_8, \theta_9, \theta_{10}\}$. The second example points to the over-specificity of the same system about the alternatives θ_{14}, θ_{15}, and θ_{16} in presenting the information $\{\theta_8, \theta_9, \theta_{13}\}$. Our question is what imposes this particular over-specificity on the Euler diagrams, given the syntactic and semantic rules we adopt for them.

Again, our strategy is to look closely what happens on the part of the diagrams during the operation in the examples. In example 3.1, we first present the premises θ_8 and θ_9 in a diagram. Given our semantic convention for Euler diagrams, this means that we realize the following states of affairs in a diagram:

(σ_5) A circle labeled "*C*s" is inside a circle labeled "*B*s."

(σ_6) A circle labeled "*A*s" is completely outside the *B*-circle.

Now, to present the assumption θ_{10}, we want to realize the following state of affairs in the diagram (while preserving σ_5 and σ_6):

(σ_7) A circle labeled Ds encloses the B-circle in it.

The trouble is that doing so makes the diagram present either the information θ_6 or the information θ_7. But what does this mean in terms of the states of affairs holding in the diagram? It means that whenever the diagram supports state of affairs σ_6 while preserving σ_4 and σ_5, it *must* support either of the following states of affairs:

(σ_8) A circle labeled "Ds" has no overlap with a circle labeled "As."

(σ_9) A circle labeled "Ds" overlaps with a circle labeled "As."

In other words, there is a structural constraint governing the Euler diagram from the set of states of affairs $\{\sigma_4, \sigma_5, \sigma_6\}$ to the *disjunction* of the states of affairs σ_7 and σ_8. It is this structural constraint on Euler diagrams that gives rise to the particular over-specificity of the Euler system we have observed. The following chart summarizes how it is so:

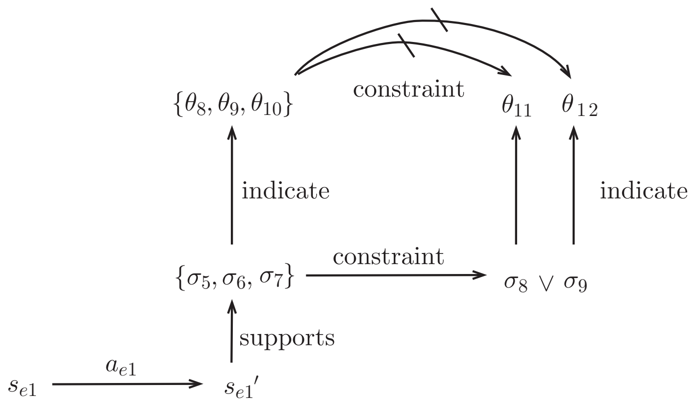


Figure 3.11

Here we apply an operation a_{e1} to an Euler diagram s_{e1} to realize the set of states of affairs $\{\sigma_5, \sigma_6, \sigma_7\}$. These states of affairs indicate the set of assumptions $\{\theta_8, \theta_9, \theta_{10}\}$, so we can say that the resulting diagram $s_{e1}{}'$ presents those assumptions. Now, because of the structural constraint from $\{\sigma_5, \sigma_6, \sigma_7\}$ to $\sigma_8 \vee \sigma_9$ holding on the diagram $s_{e1}{}'$, the diagram must support either of the states of affairs σ_8 and σ_9, while these respectively indicate θ_{11} and θ_{12}. Thus, the operation a_{e1} ends up adding in the diagram $a_{e1}{}'$ either the information θ_{11} or the information θ_{12}, while neither follows from the assumptions $\{\theta_8, \theta_9, \theta_{10}\}$. Notice that, on the semantic convention for Euler diagrams, $\{\sigma_5, \sigma_6, \sigma_7\}$ is the only set of states of affairs that indicate the assumptions $\{\theta_8, \theta_9, \theta_{10}\}$. It follows that *no* Euler diagrams can present these assumptions without thereby presenting either θ_{11} or θ_{12}.

Example 3.2 is a similar story. After drawing an Euler diagram (Figure 3.5) to present the premises θ_8 and θ_9, we wish to add the premise θ_{13} to the diagram while letting the diagram still present θ_8 and θ_9. For this purpose, we need to realize the following state of affairs in the diagram:

(σ_{10}) A circle labeled "Ds" enclose the A-circle in it

while preserving the states of affairs σ_4 and σ_5. *As a matter of structural constraint* governing Euler diagrams, however, whenever an Euler diagram supports the set of states of affairs $\{\sigma_4, \sigma_5, \sigma_{10}\}$, it must support either of the following states of affairs:

(σ_{11}) The circle labeled "Ds" has no overlap with circles labeled "Bs" and "Cs."

(σ_{12}) The circle labeled "Ds" overlaps with a circle labeled "Bs" but not with a circle labeled "Cs."

(σ_{13}) The circle labeled "Ds" overlaps with circles labeled "Bs" and "Cs."

On the semantic convention we adopt, these alternative states of affairs indicate the information θ_{14}, θ_{15}, and θ_{16} respectively. Unfortunately, none of these follows from the given set of premises $\{\theta_8, \theta_9, \theta_{13}\}$. Thus, we must either give up presenting the premise θ_{13} or tolerate our diagram to present one of the above invalid conclusions.

Again, a structural constraint governing Euler diagrams gives rise to particular over-specificity of the system, namely, the specificity about the alternatives θ_{14}, θ_{15}, and θ_{16} in presenting the information $\{\theta_8, \theta_9, \theta_{13}\}$.

The following chart schematizes this situation:

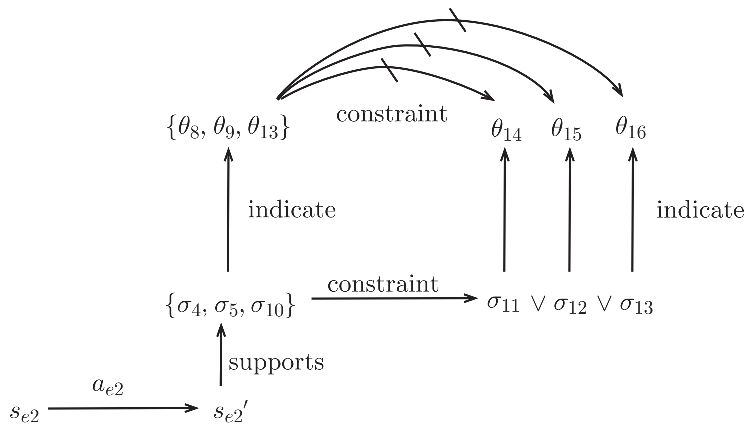


Figure 3.12

Because of the structural constraint from $\{\sigma_4, \sigma_5, \sigma_{10}\}$ to $\sigma_{11} \vee \sigma_{12} \vee \sigma_{13}$, the action a_{e2} that puts the set of premises $\{\theta_8, \theta_9, \theta_{13}\}$ in an Euler diagram s_{e2}' ends up adding either θ_{14} or θ_{15} or θ_{16}. Since $\{\sigma_4, \sigma_5, \sigma_{10}\}$ is the only set of states of affairs that indicate the set of assumptions $\{\theta_8, \theta_9, \theta_{13}\}$, there is no Euler diagram that presents these assumptions without thereby presenting either θ_{14} or θ_{15} or θ_{16}. Note that none of these follow from the assumptions.

We will now describe, in general terms, how a structural constraint on representation gives rise to particular over-specificity of a system. A system of representation is *over-specific* in representing the set of assumptions Θ_1 iff, for all set of states of affairs Σ_1 that indicate Θ_1, the following conditions hold:

1. There holds on the representations of the system a structural constraint from

Σ_1 to the *disjunction* of states of affairs $\sigma_1^* \vee \sigma_2^* \vee \ldots \vee \sigma_n^*$ (this can be an infinite disjunction),

2. Each disjunct σ_i^* indicates a piece of information θ_i^* such that the constraint from Θ_1 to θ_i^* does not hold on the targets of the system.

Under this condition, no representation in the system can present the assumptions Θ_1 in isolation. To wit, suppose you want to present the set of assumptions Θ_1 by realizing a set of states of affairs Σ_1 in a representation. By clause 1 above, there holds on the representation a structural constraint from Σ_1 to the disjunction $\sigma_1^* \vee \sigma_2^* \vee \ldots \vee \sigma_n^*$. Thus, the representation must support one of the disjuncts, say σ_i^*. But σ_i^* indicates the information θ_i^* that does not follow from the assumptions Θ_1 (clause 2). Thus, the operation ends up presenting the extra information θ_i^*. It follows that, under the above conditions, we cannot use any representation in the system to present the assumptions Θ_1 in isolation from one of the alternative pieces of information $\theta_1^*, \theta_2^*, \ldots, \theta_n^*$.

If we try to override this limitation, and present the assumptions Θ_1 in a representation of the system, then the representation will support so-called accidental features. For, due to the constraint specified in clause 2, such a representation supports one of the states of affairs $\sigma{*}_1, \sigma{*}_2, \ldots \sigma{*}_n$ along with all the consequences $\sigma{*}_{n+1}, \sigma{*}_{n+2}, \ldots \sigma{*}_{n+m}$ that our arbitrary operation has on the part of our representation. Given the semantic convention for the system, this means that we have to tolerate our diagram presenting one of the pieces of information $\theta_1^*, \theta_2^*, \ldots \theta_n^*$ plus any of the information indicated by $\sigma{*}_{n+1}, \sigma{*}_{n+2}, \ldots \sigma{*}_{n+m}$. There is no guarantee that these pieces of information follow from our original assumptions Θ_1.

Note that this over-specificity of the system is largely due to the structural constraint from Σ_1 to $\sigma_1^* \vee \sigma_2^* \vee \ldots \vee \sigma_n^*$ governing the representations of the system plus the particular way in which this constraint fails to match with the constraints on the targets. If it were not for this constraint, we could realize the set of states of affairs Σ_1 without realizing none of the states of affairs σ_1^*, σ_2^*, ..., σ_n^*, and thus put the information Θ_1 in our representation without committing ourselves to any of the unwarranted pieces of information θ_1^*, θ_2^*, ..., θ_n^*. Also, if there were a constraint, on the targets, from Θ_1 to at least one of θ_1^*, θ_2^*, ..., θ_n^*, then there would be a chance that we present the assumptions Θ_1 with that piece of information, and thus avoid choosing an unwarranted piece of information. Because of the particular mismatch between the constraints on the two domains, we do not have this chance.

There are, however, two different cases of over-specificity, depending on whether the underlying constraints on representations are *nomic* or *stipulated*. An example of *stipulated* structural constraints for a system of Euler diagrams would be that no closed curves touch each other without overlapping. This stipulation would be necessary to avoid the ambiguity of interpretation of such curves, concerning whether the sets denoted by them intersect or not. One of the stipulated constraints for a system of Venn diagrams is that all closed curves in a Venn diagram partially overlap. This stipulation is necessary to make sure that every possible intersection and complement of the denoted sets is represented by some region of the diagram. Another example is the stipulation for Stenning and Oberlander's system of first order language cited in section 3.1 that says that every well-formed representation must consist of a special sequence of atomic sentences conjoined by the sign "$\wedge$."

The structural constraints we have seen in examples 3.1 and 3.2 are not stipulated in this sense. The constraint relevant in example 3.1 is a special case of the general geometrical constraint that whenever there are two circles on a plane, they overlap or do not overlap. The constraint relevant in example 3.2 is a special case of the general constraint that whenever you have three circles on a plane with one enclosing another, the other circle overlaps with either none of the others, only one, or both. It would be indeed quite redundant if we try to stipulate these constraints for a system of Euler diagrams. We call a structural constraint "purely nomic" if it holds on a set of representations without needing any stipulation on our part. We call a constraint "purely stipulative" if it holds on a set of representations purely in virtue of the syntactic stipulations we make for the representations. We call a constraint "partially nomic" if it holds on a set of representations due to both nomic and stipulated constraints.

On our analysis, then, the view that I criticized in section 3.1 is the claim that all case of over-specificity of representation system are cases in which the constraint from Σ_1 to $\sigma_1^* \vee \sigma_2^* \ldots \vee \sigma_n^*$ is purely stipulative. What we have claimed, by contrast, is that in many cases, the constraint from Σ_1 to $\sigma_1^* \vee \sigma_2^* \ldots \vee \sigma_n^*$ is partially or purely nomic. In fact, we have just shown that both examples 3.1 and 3.2 are cases in which the over-specificity of a system is due to purely nomic structural constraint governing the representations in the system. As we will see , some cases of over-specificity cited in section 2.1 are due to purely nomic structural constraints on representations.

3.3 Testing the analysis

In this section, we will take up some examples cited in section 3.1, and show that our informal analysis of the phenomenon of over-specificity fit those examples. For those sympathetic readers who already believe that our analysis is correct, the following will show *how* our general analysis applies to particular cases. We will consider the examples of Harry's memory map cited in Hohauser (1982), of Hyperproof diagrams, and of geometry diagrams for the Pythagorean theorem cited in Barwise and Etchemendy (1990b).

The system of memory map

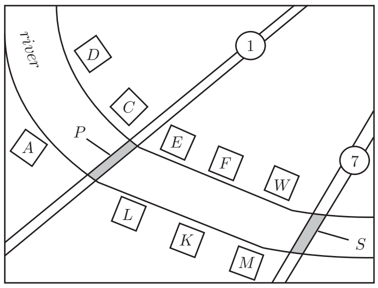


Figure 3.1

This is the case in which the system of Harry's memory map exhibit the over-specificity about the alternatives θ_2, θ_3, and θ_4 in presenting θ_1 in the above map. If our analysis is correct, we should be able to specify a structural constraint governing Harry's maps that is responsible for this over-specificity. Consider the constraint from the set of states of affairs:

(σ_{14}) The *B* block is between the *A* block and the *K* block.

(σ_{15}) The road line 1 is between the *A* block and the *L* block.

(σ_{16}) The *L* block is between the road line 1 and the *K* block.

to the disjunction of states of affairs:

(σ_1) The *B* block is between the *A* block and the road line 1.

(σ_{17}) The *B* block is between the road line 1 and the *L* block.

(σ_{18}) The *B* block is between the *L* block and the *K* block.

The reader should check that this is a *partly nomic* constraint reasonable to be assumed to govern Harry's memory maps. Among the above states of affairs, σ_{15} and σ_{16} already hold on the above memory.

Recall that we let Θ_m be the entire set of information presented in this map. Let Σ_m be the set of states of affairs, holding in the above diagram, that indicate this set of information Θ_m. Then, since σ_{15} and σ_{16} are members of Σ_m and the above constraint holds on Harry's maps, it follows that the constraint from $\Sigma_m \cup \{\sigma_{14}\}$ to $\sigma_1 \vee \sigma_{17} \vee \sigma_{18}$ also holds on Harry's maps. It is this constraint that is responsible for the over-specificity at issue.

Harry wants to add the piece of information θ_1 by realizing the state of affairs σ_{14} in this map. However, due to this structural constraint on the map, if he realizes σ_{14} while preserving the states of affairs Σ_m, he must realize one of the alternative states of affairs σ_1, σ_{17}, and σ_{18}. These states of affairs respectively indicate the pieces of information θ_2, θ_3, and θ_4. Unfortunately, neither follows from the information $\Theta_m \cup \{\theta_1\}$ that Harry is willing to keep in his memory map. Moreover, according to the semantic convention associated with Harry's memory map, there is no other way of putting this information in a memory map. More precisely, there is no other set of states of affairs than $\Sigma_m \cup \{\sigma_{14}\}$ that indicate the information $\Theta_m \cup \{\theta_1\}$. Thus, the system of Harry's memory map is over-specific in presenting the information $\Theta_m \cup \{\theta_1\}$, exactly in the way our analysis characterizes.

The system of Hyperproof diagram

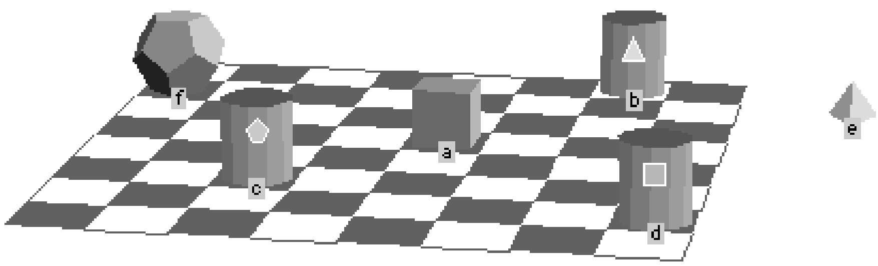


Figure 3.3

- $SameCol(e, a)$
- $SameSize(b, f) \vee SameSize(c, f)$

We saw that the system of Hyperproof diagram is over-specific in presenting the content of each of these sentences in the above diagram. Let us look at the case of the first sentence. Here is the information presented in the sentence:

(θ_{h2}) The block e is in the same column as the block a.

Since we let Θ_h be the entire information content of the diagram, we are considering the over-specificity of the system of Hyperproof diagrams in presenting the information $\Theta_h \cup \{\theta_{h2}\}$. Consider the constraint, holding on Hyperproof diagrams, from the set of states of affairs:

(σ_{19}) The tetrahedron icon labeled "*e*" is in the same column as the cube icon labeled "*a*"

(σ_{20}) The cube icon labeled "*a*" is in row 5 and column 5.

to the disjunction of states of affairs:

(σ_{21}) The tetrahedron icon labeled "*e*" is in row 1 and column 5.

(σ_{22}) The tetrahedron icon labeled "*e*" is in row 2 and column 5.

(σ_{23}) The tetrahedron icon labeled "*e*" is in row 3 and column 5.

(σ_{24}) The tetrahedron icon labeled "*e*" is in row 4 and column 5.

(σ_{25}) The tetrahedron icon labeled "*e*" is in row 6 and column 5.

(σ_{26}) The tetrahedron icon labeled "*e*" is in row 7 and column 5.

(σ_{27}) The tetrahedron icon labeled "e" is in row 8 and column 5.

The reader should convince him/herself that this is a partially nomic constraint that is reasonable to assume on Hyperproof diagrams. Here, σ_{20} is one of the states of affairs that already hold on the diagram above.

Recall that we let Θ_h be the entire set of information presented in the above diagram. Let Σ_h be the set of states of affairs, holding in the diagram, that indicate this chunk of information. Since σ_{20} is a member of Σ_h, we can assume that the constraint from $\Sigma_h \cup \sigma_{19}$ to $\sigma_{21} \vee \sigma_{22} \vee \sigma_{23} \vee \sigma_{24} \vee \sigma_{25} \vee \sigma_{26} \vee \sigma_{27}$ holds on Hyperproof diagrams. Because of this constraint, when we add the information θ_{h2} by realizing the state of affairs σ_{19} in the above diagram, we must realize one of the states of affairs σ_{21}, σ_{22}, σ_{23}, σ_{24}, σ_{25}, σ_{26}, and σ_{27}. Each of these indicates a piece of information about the exact row in which the tetrahedron e is located, but, unfortunately, none follow from the information $\Theta_h \cup \{\theta_{h2}\}$ that we wish to present in the diagram. Furthermore, under the semantic rules for Hyperproof diagrams, there is no set of states of affairs, other than $\Sigma_h \cup \sigma_{19}$, that indicate the information $\Theta_h \cup \{\theta_{h2}\}$. This is, therefore, a case of over-specificity as we characterize in our analysis. A similar account explains why we cannot incorporate the content of the sentence $SameSize(b, f) \vee SameSize(c, f)$ into the given diagram.

Reflecting this limitation of its diagrammatic part, the logical system underlying Hyperproof does not allow us to "Apply" the sentences $SameCol(e, a)$ or $SameSize(b, f) \vee SameSize(c, f)$ to the diagram. This is to prevent the diagram from supporting accidental features.[5]

[5]Thus, the diagrammatic part of the Hyperproof representation system is ineffective for the purpose of presenting *any old* combination of information we obtain (or assume) in a given stage of problem solving. However, the logical system of Hyperproof has a special rule, called "Case Exhaustive," that lets us consider several diagrams that embody the alternative states of affairs, and draw a valid conclusion that is common in all these diagrams. Furthermore, the sentential part of the system supplements this limitation of the diagrammatic part. This is suggestive about the utility of heterogeneous system of representation that comprises two or more kinds of representations.

The system of geometry diagram

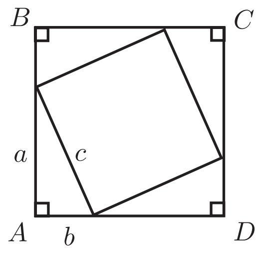


Figure 3.4

The structural constraint responsible for the over-specificity of geometry diagrams in our example is the one from:

(σ_{28}) There is a triangle icon whose sides are labeled "*a*," "*b*," and "*c*."

to the disjunction of states of affairs:

(σ_{29}) The "*a*" side of the triangle icon is longer than the "*b*" side.

(σ_{30}) The "*a*" side of the triangle icon is as long as the "*b*" side.

(σ_{31}) The "*a*" side of the triangle icon is shorter than the "*b*" side.

This constraints is purely nomic again. We will leave the reader to elaborate how this structural constraint gives rise to the over-specificity we saw about this example. Unlike the cases of Hyperproof diagram, we do not stop presenting our assumption in the face of the danger of accidental features σ_{29}, σ_{30}, and σ_{31}. It seems that after one or two courses of elementary geometry, we develop a certain tolerance against such accidental features, and learn to conduct valid reasoning in their presence. Luengo (1995) formulates a logical system that formalizes the rules for such valid use of geometry diagrams. Although this is an interesting topic, our aim here is not to investigate such tolerance. We just wish to emphasize that the accidental features typical in geometry diagrams have their root in the over-specificity of the geometry diagrams, which in turn has its root in the structural constraint governing the geometry diagrams.

3.4 Conclusion

We started out by identifying the property of over-specificity of representation systems. In examining a number of examples, we have seen that the over-specificity of a system accounts for the inefficacy for the purpose of using a representation to present the arbitrary chunk of information that has been obtained or assumed in the course of reasoning. We then went onto the analysis of the phenomenon, asking why it is that no representation in a system can present certain chunks of information without presenting certain other. According to our analysis, the over-specificity of a representations is due to a structural constraint governing the representation and to a particular mismatch of this constraint to the constraints on the targets. We distinguished two kinds of over-specificity according as whether the responsible constraint is purely or partially nomic or purely stipulated. This way, we believe, we have isolated and analyzed an important species of inefficacy associated with representation systems.

Before we close our discussion, let us compare the phenomenon of content specificity with the phenomenon of free ride discussed in the last chapter. What is common between these two phenomena? Both involves a constraint on representations that prevents us from presenting certain information without presenting certain other. In a case of free ride, we present a set of assumptions Θ_1 by means of a certain set of states of affairs Σ_1, and then, due to a structural constraint, we end up representation an additional piece of information θ. The latter information is enforced, so to speak, just as an extra piece of information θ_i^* is enforced in the case of over-specificity. Of course, the free ride case differs from the over-specificity case in that (a) θ is a consequence of Θ_1 and (2) θ is not one of the alternative pieces of information that may be enforced, while θ_1^* can be one of such alternatives. However, it is suggestive that these different phenomena, which account for totally different kinds of efficacy and inefficacy of a representation system, has this common element, namely, a constraint on representations that enforces an extra piece of information. In fact, coupled with a few additional conditions, this element distinguishes so-called graphical representation systems from so-called linguistic representation systems. We will explore this topic in chapter 5.

Chapter 4

The Conceptual Framework

In chapters 2 and 3, we identified the phenomena of free ride and content specificity and showed that they account for important varieties of efficacy and inefficacy of representation systems. We proposed informal analyses of the phenomena, and demonstrated that structural constraints on representations play a crucial role in both phenomena. These results lend a good initial support for the Constraints Hypothesis we are advocating, which claims that we can account for the inferential potentials of different representation systems largely in terms of structural constraints on representations and their match and mismatch with the constraints on the targets. This in turn gives us a motivation for a conceptual framework that is designed to make explicit (a) what structural constraints govern the representations in a given system, (b) what constraints govern its targets, (c) how these two sets of constraints match and fail to match (via some semantic relation), and (d) how these match and mismatch underlies the perceived efficacy and inefficacy of the system. We will now develop such a framework.

Section 4.1 will explicate the main concepts of our framework. They include *situation*, *state of affairs* holding in a situation, *representation* as a species of situation, *indication* relation from states of affairs to states of affairs, *signaling* relation from representations to situations, *constraint* (on source and target domains), and finally *representation system*. Most of these concepts have been used in the previous chapters in their intuitive senses. We will now give them more definite meanings based on the explicitly defined conceptual framework. In section 4.2, we will compare our conceptual framework with an alternative framework for modeling the functions of representation systems, which we dub the "logic framework." We will use the comparison not to criticize the alternative framework, but to highlight the advantages and disadvantages of our own framework in contrast to the alternative. In the final section, we will revisit the phenomena of free ride and content specificity, and use our conceptual framework to refine the informal analyses of the phenomena given in the previous chapters. We will introduce the notion of constraint projection as a property of representation systems, and reconstruct the phenomena of free ride and content specificity out of this property. We will also characterize the property of self-consistency of representation systems, newly introduced in this chapter.

The formal analyses given in this chapter show that capacity for free ride, content specificity, and self-consistency are all special instances of matching and mismatching between constraints on representations and constraints on targets. They therefore gives a strong support for the Constraint Hypothesis. In the next chapter, we will use the framework introduced in this chapter, especially the notion of constraint projection, to analyze the conceptual difference between so-called graphical systems and so-called linguistic systems of representation.

4.1 Main concepts

We divide our main concepts into three groups. The first group contains the notions of situation and state of affairs, which are the basic building blocks in the ontology of our framework. The second group include the notions of representation, source domain, target domain, indicating relation, and signaling relation, which are mainly concerned with what it is for a situation to present a piece of information about some other situation. The third group are concerned with the central notions of our framework, namely, the concepts of (nomic or stipulative) constraint on representations and constraint on targets.[1]

Situations and states of affairs

The basic building blocks in our ontology are situations and states of affairs. Some of the situations are actual while the others are not actual. For example, the situation s_B that today's *Herard Times* describes as Bloomington's music scene is an actual situation, but the situation s_G that the movie *Batman* describes as Gotham City's crime scene is not an actual situation.

Intuitively, situations are "sites" in which various states of affairs hold or do not hold. For example, that Jazz Fables is famous is a state of affairs that holds in the situation s_B, but that Atsushi is famous is a state of affairs that does not hold in s_B. Also, that Batman is famous is a state of affairs that holds in s_G. We can thus consider states of affairs as the properties of situations that classify the situations in

[1]Many of the concepts in the first group have their origins in the literature on situation theory, especially, Barwise and Perry (1980, 1983), Barwise (1989), Devlin (1991), and Westerståhl (1990). Many concepts in the second and the third group are borrowed from more recent literature on the qualitative theory of information, especially, Barwise (1993), Barwise, Gabbay, and Hartonas (1995), and Barwise and Seligman (1996).

which they hold. When a state of affairs σ holds in a situation s, we also say that s *supports* σ , written $s \models \sigma$.

Each state of affairs σ has a unique dual $\bar{\sigma}$ such that $\sigma = \bar{\bar{\sigma}}$. Although σ and $\bar{\sigma}$ never hold together in a single situation, it is possible that neither σ nor $\bar{\sigma}$ holds in a given situation. Situations need not settle the issue of whether σ or $\bar{\sigma}$, so to speak. We say that a situation s is *determinate* relative to a state of affairs σ if and only if either σ or $\bar{\sigma}$ holds in s. For example, although the state of affairs that Atsushi is famous does not hold in the situation s_B, s_B is determinate relative to this state of affairs, since its dual (the state of affairs that Atsushi is not famous) holds in Bloomington's music scene. On the other hand, s_B is not determinate relative to the state of affairs that Atsushi is bald. For, although I am not bald, it is not Bloomington's music scene that settles the issue of whether I am bald or not. Likewise, the situation s_B is determinate relative to neither state of affairs θ_1 nor θ_2 in the following list, but it is determinate relative to both states of affairs θ_3 and θ_4:

(θ_1) May 3 is the deadline for submission of all materials for PhDs to be issued on May 4

(θ_2) For every two points A and B there exists a line that contains each of the points A and B

(θ_3) There will no live performance at Bluebird this Thursday

(θ_4) Bear's Place is the only place where the Jazz Fables play

We will also have conjunctions and disjunctions of states of affairs. Thus, a situation s supports $\sigma_1 \wedge \sigma_2$ iff s supports both σ_1 and σ_2; s supports $\sigma_1 \vee \sigma_2$ iff s supports either σ_1 or σ_2, and so on.

We use English lowercases $s, s', t, t', \ldots$ for situations, actual or only possible, and use greek lowercases $\sigma, \sigma', \theta, \theta', \ldots$ for states of affairs. English capitals $S, S', T, T', \ldots$

are for sets of situations, and Greek capitals $\Sigma, \Sigma', \Theta, \Theta', \ldots$ are for sets of states of affairs.

Representation

So far, we have been using the term "representation" rather intuitively, to mean objects such as sentences, diagrams, and models that we use to present information about particular objects on the basis of certain semantic rules. We will now give clearer meaning to the term. We will start out using the notion "representation system" still intuitively, but will spell out its meaning gradually, as we proceed to lay out our conceptual framework.

In our conception, representations are situations that we (cognitive agents) create in order to present information about particular objects to ourselves or to others. Thus, the English sentence that you are reading now is a situation that I created (or let the printer create) to present certain information about our concept of situation. The map of Indianapolis that I drew yesterday to show my wife the location of the Indianapolis zoo is a situation that I created to present the information about the zoo.

Given a representation system $\mathcal{R}$, there is a range of information that the representations in the system can possibly present. We model this range of information by means of a set of states of affairs Θ'. To each state of affairs θ in Θ', we assign one or more states of affairs σ, typically one different from θ itself, with the agreement that whenever σ holds in a representation s, we can read off the information θ from s. If a state of affairs σ lets us read off a piece of information θ as the result of this agreement, we say "σ indicates θ" and write $\sigma \Rightarrow \theta$. We call the binary relation $\Rightarrow$ the "indication relation" for the representation system $\mathcal{R}$. Thus, representations present information partly by virtue of the semantic rules and the states of affairs that hold in them.

Let $dom(\Rightarrow)$ be the domain of the indication relation $\Rightarrow$, and S be all the situations (actual or non-actual) that are determinate relative to each element of $dom(\Rightarrow)$. Let Σ be the set of state of affairs relative to which each element of S is determinate. Then, S consists of situations that may or may not support states of affairs in Σ. Let $\models_{\mathcal{S}}$ be this binary relation of supporting defined on $S \times \Sigma$. We call the triple $\mathbf{S} = \langle S, \Sigma, \models_{\mathbf{S}} \rangle$ the "source domain" of the given representation system $\mathcal{R}$. Likewise, let $ran(\Rightarrow)$ be the range of the indication relation $\Rightarrow$, and T be all the situations (actual or non-actual) that are determinate relative to each element of $ran(\Rightarrow)$. Let Θ be the set of state of affairs relative to which each element of T is determinate. Again, T comprises situations that may or may not support states of affairs in Θ. Letting $\models_{\mathbf{T}}$ be the supporting relation defined on $T \times \Theta$, we call the triple $\mathbf{T} = \langle T, \Theta, \models_{\mathbf{T}} \rangle$ the "target domain" of the representation system $\mathcal{R}$.[2]

Note that only the states of affairs in $dom(\Rightarrow)$ are meaningful within the system $\mathcal{R}$, and only the states of affairs in $ran(\Rightarrow)$ are information within the coverage of $\mathcal{R}$. In this dissertation, we assume that every state of affairs σ in $dom(\Rightarrow)$ has a unique state of affairs that it indicates. Thus, the indicating relation $\Rightarrow$ is a (often partial) function from Σ into Θ. For sets of states of affairs $\Sigma_i \subseteq \Sigma$ and $\Theta_i \subseteq \Theta$, we say Σ_i indicate Θ_i and write "$\Sigma_i \Rightarrow \Theta_i$" iff Θ_i is the image of Σ_i under the indication relation $\Rightarrow$.

Suppose a state of affairs σ holds in a situation s and $\sigma \Rightarrow \theta$. We may say that s presents the information θ. But which situation in T does s carry the information θ about? The information θ is true of some elements of T but false of the others,

[2]Thus, the source domain $\mathbf{S}$ and the target domain $\mathbf{T}$ of a representation system constitute "classifications" as defined in Barwise and Seligman (1996). This makes our conceptual framework easily embeddable in the general theory of information flow that Barwise and Seligman are developing.

and depending on which situation s is about, s may be either true or false as a representation. Thus, for s to be a really meaningful representation, *something* must determine which situation in T the situation s carries information θ about—which situation s is targeting at. We model this targeting relation by means of a binary relation from the set of situation S to the set of situation T, writing "$s \leadsto t$." Thus, given a representation system $\mathcal{R}$, the domain of the targeting relation $\leadsto$ consists of all the situations in S that have definite targets in T, and the range of $\leadsto$ consists of all the situations in T that are targeted by some situations in S.

We are in the position of characterizing what it is for a situation to present information about another situation in a representation system.

Definition (Presenting information) Let $\mathbf{S} = \langle S, \Sigma, \models_{\mathbf{S}} \rangle$ and $\mathbf{T} = \langle T, \Theta, \models_{\mathbf{T}} \rangle$ be the source domain and the target domain of a representation system $\mathcal{R}$. Let $s \in S$, $t \in T$, and $\theta \in \Theta$. We say that the situation s *presents* the information θ *about* the situation t in $\mathcal{R}$ iff there is a state of affairs $\sigma \in \Sigma$ such that:

1. $s \models_{\mathbf{S}} \sigma$,
2. $\sigma \Rightarrow \theta$,
3. $s \leadsto t$.

The representation s is *true in* the system $\mathcal{R}$ iff each information that s presents about its target t in $\mathcal{R}$ is supported by t.

Constraints

Central to the Constraint Hypothesis and to the analyses developed in chapters 2 and 3 are the notions of "constraint holding on the representations" and "constraint holding on the targets." We will now illuminate these notions on the basis of the conceptual framework we are developing.

Let us start with the first notion. What exactly is it for a constraint to hold on the representations in a representation system? To answer this question, we need first determine what constitutes a representation *within* a representation system. Consider the source domain $\mathbf{S} = \langle S, \Sigma, \models_{\mathbf{S}} \rangle$ of a representation system $\mathcal{R}$. Remember that the set S contains *any possible* situation that is determinate relative to the set of states of affairs $dom(\Rightarrow)$. For two reasons, S is too large to be considered as the set of representations within the system $\mathcal{R}$. First, S may contain situations whose syntactic features are "ill-formed" with respect to the syntactic formation rules stipulated in $\mathcal{R}$. Second, S may contain "abnormal" situations that we do not want to count as representations of the system. Let us expand these points one by one.

The first point is rather obvious. A representation system often comes with a set of more or less explicitly accepted syntactic conventions, or "formation rules." These conventions impose so-called "well-formedness conditions" on the representations of the system. However, these well-formedness conditions do not extend over the set S, which comprises any possible situation that is determinate relative to the set of states of affairs $dom(\Rightarrow)$. Obviously, we do not count those ill-formed situations in S to be representations within the system. Doing so would amount to ignoring the formation rules stipulated in the system altogether, which may have important effects on the overall efficacy and inefficacy of the system.

The second point is the following. The world is governed by many different sorts of regularity that we call "natural laws." We can subdivide these laws into different systems, such as logical laws, arithmetical laws, Euclidean laws, mechanical laws, and so on. As situations residing in the world, most members of the set S are subject to these laws. However, since we define S as all possible situations that are determinate relative to the set of states of affairs Σ, S may well contain "abnormal situations" that make exceptions to these "normal" natural laws.[3] Now, what we are interested in in this dissertation is the efficacy and inefficacy of a given representations system in the *normal* circumstance. More specifically, we are interested in how effectively we can fulfill our problem-solving tasks when we use *normal* representations according to the syntactic and semantic rules prescribed in the system. We do not want to praise the system as effective or blame it as ineffective on the basis of the cases in which we use *abnormal* representations that violate the natural regularities that we can reasonably expect. For this reason, we do not count any abnormal situations in S as representations of the system.

We thus decide to count as proper representations of the representation system only those situations in S that are both normal and well-formed. We collect those situations in S that obey the normal set of natural laws, and name the resulting set "S_n." We also collect those situations in S that obey the formation rules in $\mathcal{R}$, and call the set "S_s." The set of representations of the system $\mathcal{R}$ is then the intersection $S_n \cap S_s$.

With these notions at hand, we can now elucidate what it is for a constraint to hold on the representations in the system $\mathcal{R}$. We call any pair $\langle \Gamma, \Delta \rangle$ of subsets of Σ a "source sequent" and write it as "$\Gamma \vdash \Delta$." We say, "a situation s in S *respects* a source sequent $\Gamma \vdash \Delta$" to mean that if s supports *all* members of Γ, then s supports *at least one* member of Δ. What we call "constraints on representations" are those source sequents that are respected by all representations in $\mathcal{R}$. More precisely:

Definition (Constraint on representations) A source sequent $\Gamma \vdash \Delta$ is a *constraint on the representations of* $\mathcal{R}$ iff: all situations in $S_n \cap S_s$ respect $\Gamma \vdash \Delta$.

The two sets S_n and S_s let us clarify the notions of "purely nomic," "purely

[3] We will shortly discuss the issue of what makes a situation "normal" or "abnormal."

stipulated," and "partially nomic" constraint on representations that we used in the previous chapters:

Definition (Nomic and stipulative constraint on representations) Let $\Gamma \vdash \Delta$ be a constraint on the representations of a representation system $\mathcal{R}$.

- The constraint $\Gamma \vdash \Delta$ is *purely nomic* iff: all situations in S_n respect $\Gamma \vdash \Delta$.
- The constraint $\Gamma \vdash \Delta$ is *purely stipulative* iff: all situations in S_s respect $\Gamma \vdash \Delta$ but some in S_n do not.
- The constraint $\Gamma \vdash \Delta$ is *partly nomic* iff: all situations in $S_n \cap S_s$ respect $\Gamma \vdash \Delta$ but some in S_n and some in S_s do not.
- The constraint $\Gamma \vdash \Delta$ is *nomic* iff it is either purely nomic or partly nomic.
- The constraint $\Gamma \vdash \Delta$ is *trivial* iff $\Gamma \cap \Delta \neq \emptyset$.

What is it, then, for a constraint to hold on the *targets* of a representation system? What constitute a target of the representation system in the first place?

Consider the target domain $\mathbf{T} = \langle T, \Theta, \models_{\mathbf{T}} \rangle$ of the representation system. We can think of different kinds of regularities that are imposed upon most of the elements of T. Some regularities are due to natural laws such as mechanical laws. Some regularities are due to stipulative laws such as the U.S. laws and hotel regulations. In fact, it is these regularities that make it possible for us to make a valid inference about the target domain at all. Now, since the set of situations T is defined as the set of all possible situations that are determinate relative to every state of affairs in $ran(\Rightarrow)$, S may well comprise exceptions both for these natural and stipulative regularities. However, we are interested in the efficacy of the representation system when it is used to help reason about normal targets, rather than when it is used to help reason about exceptional abnormal targets. For this reason, we count only normal situations in S as proper targets of the representation system.

Thus, we collect the set of situations in T that obey the normal regularities and take the resulting set, called T_{ns}, to be the set of targets of the representation system. We call any pair $\langle \Gamma, \Delta \rangle$ of subsets of Θ a *target sequent* and write it as "$\Gamma \vdash \Delta$." We define a constraint holding on the targets of the representation system in the following way:

Definition (Constraint on targets) A target sequent $\Gamma \vdash \Delta$ is a *constraint on the targets of the system* $\mathcal{R}$ iff: all situations in T_{ns} respect $\Gamma \vdash \Delta$.

Note that we followed a slightly different procedure in defining the *targets* of the representation system than in defining the *representations* of the system. In defining the representations in the system, we distinguished the regularities due to natural laws and the regularities due to syntactic stipulations. We accordingly have two subsets S_n and S_s of situations that respect these two different kinds of regularities. The purpose of doing so was to differentiate two kinds of constraints governing the representations of the system, namely, the nomic and the stipulative constraints. In the case of targets, there is no need for such a distinction (not at least for our present purpose). We treat both nomic regularities and stipulative regularities on the targets on a par. Accordingly, T_{ns} is the set of constraints that respect both kinds of regularities.

We said that we are interested in the efficacy of a representation system under the circumstance in which we operate on "normal" representations that represent "normal" situations. Now, the criterion for normality may be different among different people in different occasions. To accommodate these different views on normality, we should actually not fix the set S_n of "normal" source situations and T_{ns} of "normal" target situations for a representation system. We should rather make them independent variables whose values affect the efficacy of the representations system modeled in our framework. Mainly for the sake of simplicity, however, we do not take this path, and make the simplifying assumption that there is unique criteria of normality for both source and target situations. We hope that the reader tolerate this simplification, and adjust the values of S_n and T_{ns} on the basis of his or her own criteria of normality.

We end up associating seven elements to the intuitive notion of representation system. They are a source domain $\mathbf{S} = \langle S, \Sigma, \models_{\mathbf{S}} \rangle$, a target domain $\mathbf{T} = \langle T, \Theta, \models_{\mathbf{T}} \rangle$, an indicating relation $\Rightarrow$, an signaling relation $\leadsto$, a set of nomically normal representations S_n, a set of well-formed representations S_s, and a set of normal targets T_{ns}. Henceforth, we assign the title of "representation system" to any tuple

$$\langle \mathbf{S}, \mathbf{T}, \Rightarrow, \leadsto, S_n, S_s, T_{ns} \rangle$$

that satisfies the conditions specified above. More specifically:

Definition (Representation system) A representation system is a septuple $\mathcal{R} = \langle \mathbf{S}, \mathbf{T}, \Rightarrow, \leadsto, S_n, S_s, T_{ns} \rangle$ such that:

- $\mathbf{S}$ is a source domain $\langle S, \Sigma, \models_{\mathbf{S}} \rangle$;
- $\mathbf{T}$ is a target domain $\langle T, \Theta, \models_{\mathbf{T}} \rangle$;
- $\Rightarrow$ is a (possibly partial) function from Σ into Θ;
- $\leadsto$ is a (possibly partial) function from S to T;
- S_n and S_s are subsets of S;
- T_{ns} is a subset of T.

4.2 Comparison to the logic framework

We have laid out the basic tennets of the conceptual framework. There has been developed, however, an alternative conceptual framework that is designed to model the workings of different representation systems in reasoning. We devote this section for the comparison of our framework with this alternative framework, which we call the "logic" framework for convenience.

Outline of the logic framework

In a recent paper (1995), Barwise and Hammer address the general question, "What is a logical system?" On their view, a logical system is "a mathematical model of some pretheoretic notion of consequence and an existing (or possible) inferential practice that honors it" (p. 19). Here, an "inference practice" primarily means an inference as "concerning relationships among structured representations in some sort of conventional representation system." Thus, Barwise and Hammer grasp a logical system as a mathematical model that illuminates how we use representations in particular representation systems in everyday practice of reasoning.

It is not our concern whether their view is exactly to the taste of all logicians who develop particular logical systems. Different people develop logical systems for different purposes, and some may be able to demonstrate that his/her own system is a definite counter-example to Barwise and Hammer's thesis. However, at least the following part of their view seems to be true: the conceptual framework used in developing logical systems can be productively applied to analyze the working of representation systems in everyday reasoning. Let us see how this is so, on the basis of Barwise and Hammer's conception of logical systems.

According to Barwise and Hammer (pp. 42–43), a typical logical system $\mathcal{S}$ contains the following four components: (a) grammar, (b) model structures, (c) a notion of logical consequence, and (d) a (possibly empty) set of rules of inference. If we are to use the system $\mathcal{S}$ as a model of the working of a representation system $\mathcal{R}$ in reasoning, we can take (a) as the specifications of the types of external representations that count as well-formed representations in the system $\mathcal{R}$. For this purpose, we can take (b) to stand for the targets of the system $\mathcal{R}$, (c) to stand for the nomic or stipulated constraints governing the targets, and (d) to stand for the rules that we follow in manipulating well-formed representations to update their information contents. In this conception, the completeness theorem for the logical system $\mathcal{S}$ rigorously establishes the fact that the manipulations rules of the system $\mathcal{R}$ let us update the information contents of our representation in *every* possible way that is valid with respect to the targets. The soundness theorem for the system $\mathcal{S}$ establishes that the manipulations rules of the system $\mathcal{R}$ let us update the information contents of our representation *only* in a way that is valid with respect to the targets. The strength of the logic framework consists in the fact that it can tell all these things about the representation system by means of relatively simple mathematical tools (that are familiar to every student of logic).

In fact, several logical systems have been developed that are designed to model the workings of particular representation systems in this line. Shin (1991b) develops two logical systems, Venn-I and Venn-2, that model the use of two slightly different kinds of Venn diagrams in everyday practice. She has proved the soundness and completeness of both systems for the cases where we start with a finite number of Venn diagrams. Hammer (1995) develops a logical system that contains certain first-order sentences in addition to well-formed Venn diagrams in Shin's Venn-I. He has proved the soundness and completeness of his system for the general cases where we start with an infinite number of Venn diagrams and first-order sentences. Hammer also applies the logic framework to model the workings of several other representation systems, including the system of higraphs (Harel 1988) and Peirce's α-system of existential graphs. Luengo (1995) develops a logical system, DS2, that models the use of certain diagrams in the proofs of Hilbert's plane geometry. Although the well-formed diagrams in DS2 can represent only points, lines, and angles, Luengo has proved the soundness of the system, and thus captured, to an extent, the valid use of diagrams in geometry proofs. On this conception, the logical system behind Barwise and Etchemendy's Hyperproof can be seen as a model of our everyday practice of inference that uses both sentences and diagrams.

Comparison

How does our own framework compare to this logic framework? Obviously, our framework has more structures than the logic framework does (it models a representation systems as a septuple!). What justify these additional structures? There are at least two aspects of representation systems made explicit in our framework that are implicit in the logic framework.

(1) The logic framework has the part called "grammar" which specify the syntactic structures of well-formed representations of the representation system being modeled. As such, the syntactic formation rules in this grammar part only specify the purely *stipulative* constraints on the syntactic features of the well-formed representations. They are silent about the non-stipulative, *nomic* constraints on the representations in the system. As is clear from our Constraint Hypothesis, however, it is crucial for our purpose to make explicit the nomic constraints on the representations as well as stipulative constraints. Also, we need to make it clear what it is for a constraint (stipulated or nomic) to hold on a representation. For these reasons, we have to expand the grammar part of the logic framework into a source domain $\mathbf{S} = \langle S, \Sigma, \models_{\mathbf{S}} \rangle$,

accompanied by two special subsets S_n and S_s of S. By specifying a source sequents $\Sigma_1 \vdash \Sigma_2$ respected by the members of a special subset $S_n \cap S_s$ of S, we make it clear what it is for the sequent $\Sigma_1 \vdash \Sigma_2$ to be a *constraint* on a representation, as opposed to an accidental fact about it. By differentiating the two classes of source sequents $\Sigma_1 \vdash \Sigma_2$ with reference to S_n and S_s, we can make a clear distinction between the nomic and stipulated constraints. None of these elements of a representation system are made explicit in the logic framework.

(2) The logic framework models a constraint on the targets of a representation system in terms of a semantic entailment relation from sets of well-formed representations to well-formed representations. This relation is determined by the set of model structures that stand for the targets of the representation system being modeled. Typically, a set Γ of well-formed representations is said to semantically entail a well-formed representation δ iff every structure that makes each member of Γ true makes δ true. Thus, constraints on the targets of a representation system are specified in terms of the types of well-formed representations, rather than in terms of the types of their targets. Although this feature greatly contributes to the simplicity of the logic framework, it blurs the distinction between the types that classify representations and the types that classify targets. This blurring is harmless and even economical when we discuss the completeness and soundness of the system, where we are interested in the constraints on targets *as translated* into the relationships among the types of representations. In this dissertation, however, we are interested in the correspondence between the constraints on representations and the constraints on targets, so we need to make it completely clear (a) what kind of objects classify representations and what kind of objects are constraints on them, and (b) what classifies targets and what kind of objects are constraints on them. This is why we posited a separate set of objects Θ that classify targets, where constraints on targets are defined as special sequents $\Theta_1 \vdash \Theta_2$ defined out of Θ. As we will see shortly, this lets us model a correspondence

between the constraints on representations and the constraints on targets as a correspondence between certain source sequents and certain target sequents. The logic framework just does not have enough structures to let us compare the two different kinds of constraints.

What is the limitation of our framework, compared to the logic framework? One obvious omission is that it has no component that models the rules of manipulation often associated with a representation system, while the set of inference rules in the logic framework model them to a certain extent. Thus, as it is presented in this dissertation, our framework cannot deal with the efficacy and inefficacy of a representation system that wholly or partly depend upon the associated inference rules.

It does not mean, however, that we cannot extend our framework to overcome this limitation. In fact, Shimojima (1996) and Barwise and Shimojima (1995) have developed the versions of the present framework that can model the manipulation rules associated with a representation system in a general fashion. Those extended versions have an element called the "operation rules" that models the manipulation rules of the system being modeled, and let us capture the efficacy and inefficacy of representation system that arise from the constraints on operations (as opposed to the constraints on representations). We, however, do not incorporate this feature in the present framework, because it would carry our discussion too far into afield, while the present, simpler framework lets us discuss just enough about the match and mismatch between the constraints on representations and those on targets. As the Constraint Hypothesis asserts, we can largely explain the variance of inferential potentials of different representations with reference to that much of the representation systems.

Our framework lets us specify the syntactic stipulations in a representation system, and its extended versions let us talk about the manipulation rules, the soundness, and the completeness of the system. However, just because our frameworks have more structures, they are less efficient for these purposes than the logic framework, which do the same job with much simpler mathematical tools. Our framework is designed to capture different aspects of a representation system, and for those particular purposes, it is superior to the logic framework.

4.3 Free rides, over-specificity, and self-consistency

As we will see in the rest of this dissertation, the conceptual framework developed above lets us identify many important properties of representation systems pertaining to their efficacy and inefficacy in reasoning. In this section, we will focus on the properties of providing free rides and content specificity that we discussed in chapters 2 and 3, plus the property called "self-consistency" that we newly introduce in this chapter. We will offer formal characterization of each property on the basis of the conceptual framework we have just developed.

The Constraint Hypothesis claims that the varieties and degrees of efficacy of a representation system in reasoning are largely explained with reference to particular ways in which the constraints on the representations match or fail to match with the constraints on the targets. In the dissertation, we use the following notion of "constraint projection" as the core from which we reconstruct various patterns of match and mismatch of constraints:

Definition (Constraint projection) Let $\Sigma_1 \vdash \Sigma_2$ be a source sequent and $\Theta_1 \vdash \Theta_2$ be a target sequent of a representation system $\mathcal{R}$. We say that the system $\mathcal{R}$ *projects the source constraint* $\Sigma_1 \vdash \Sigma_2$ *to the target sequent* $\Theta_1 \vdash \Theta_2$ iff:

1. Σ_1 and Σ_2 indicate Θ_1 and Θ_2 respectively,

2. $\Sigma_1 \vdash \Sigma_2$ is a constraint on the representations of $\mathcal{R}$.

Note that when a representation system has this property, we cannot use a representation of the system to present the chunk of information Θ_1 in isolation, not at least by realizing the states of affairs Σ_1. We have to present at least one piece of information in Θ_2 at the same time. To wit, suppose we want to present the information Θ_1 by realizing the states of affairs Σ_1 in a representation s' in the system $\mathcal{R}$. By the clause 2 above, there holds a structural constraint $\Sigma_1 \vdash \Sigma_2$ on the representations of $\mathcal{R}$. As a representation in $\mathcal{R}$, s' respects this constraint, and hence s' must support one of the states of affairs σ_i in Σ_2, which indicates a pieces of information θ_i in Θ_2. Thus, s' necessarily presents at least one member θ_i of Θ_2.

Of course, the additional information θ_i may or may not be a consequence of Θ_1. Whichever it may turn out to be, the property of constraint projection makes the system ineffective for the purpose of using a representation simply to record the information that has been explicitly assumed or obtained at a stage of problem solving.

Free ride and content specificity

We will use this property of constraint projection as the basis from which we reconstruct several interesting properties of the system pertaining to its efficacy and inefficacy. One of such properties is the following:

Definition (Misprojection of constraint) Let $\Sigma_1 \vdash \Sigma_2$ be a source sequent and $\Theta_1 \vdash \Theta_2$ and a target sequent of a representation system $\mathcal{R}$. We say that the system $\mathcal{R}$ *misprojects the source constraint* $\Sigma_1 \vdash \Sigma_2$ *to the target sequent* $\Theta_1 \vdash \Theta_2$ *in* $\mathcal{R}$ iff:

1. $\mathcal{R}$ projects $\Sigma_1 \vdash \Sigma_2$ to $\Theta_1 \vdash \Theta_2$;

2. For every $\theta_i \in \Theta_2$, $\Theta_1 \vdash \theta_i$ does not hold on the targets of $\mathcal{R}$.

This property of the system not only prevents us from presenting the information Θ_1 in isolation (by means of the states of affairs Σ_1), but also guarantees that the extra information that accompanies Θ_1 does *not* follow from the information Θ_1. So, this property not only prevents using a representation as a simple record of the obtained or assumed information Θ_1, but also forces an unwarranted piece of information to the representation if we express it by means of Σ_1.

The following property is even stronger: it forces an unwarranted information no matter what states of affairs we use to express the given information:

Definition (Over-specificity) Let Θ_1 be a chunk of target information of a representation system $\mathcal{R}$, i.e. $\Theta_1 \subseteq ran(\Rightarrow)$. The system $\mathcal{R}$ is *over-specific* in presenting the information Θ_1 iff: for each Σ_1 that indicates Θ_1, there are Σ_2 and Θ_2 such that $\mathcal{R}$ misprojects $\Sigma_1 \vdash \Sigma_2$ to $\Theta_1 \vdash \Theta_2$.

Thus, if a system is over-specific in presenting the information Θ_1, then no representation in the system can present the chunk of information Θ_1 in isolation from an unwarranted piece of information in Θ_2, *no matter* what states of affairs we may use to indicate them. We saw this inefficacy of representation systems in the examples of Harry's memory maps, Berkeley's picture of a triangle, Dennett's image of a tiger, Hyperproof diagrams, a geometry diagram for the Pythagorean theorem, and Euler diagrams. Many of the representation systems discussed in chapter 3 are over-specific about the chunk of information Θ_1 by the fact that there is *only one* set of states of affairs Σ_1 that indicates Θ_1, while there are sets Σ_2 and Θ_2 that satisfy the clauses 1 and 2 above.

The other important finding is that the over-specificity of these systems are due to *nomic* constraints such as geometrical and topological constraints on the formation of representations, as opposed to *stipulative* constraints such as syntactic formation rules. We therefore should sharpen the characterization of the constraint specificity of those systems in the following way:

Definition (Natural over-specificity) Let $\Theta_1 \subseteq ran(\Rightarrow)$. The system $\mathcal{R}$ is *naturally over-specific* in presenting the information Θ_1 iff: for each Σ_1 that indicates Θ_1, there are Σ_2 and Θ_2 such that:

1. $\Sigma_1 \vdash \Sigma_2$ is a nomic constraint on the representations of $\mathcal{R}$;
2. $\mathcal{R}$ misprojects $\Sigma_1 \vdash \Sigma_2$ to $\Theta_1 \vdash \Theta_2$.

Thus, the fragment of first order language cited in Stenning and Oberlander (1995) is not naturally over-specific, although it is over-specific in the general sense. As we will discuss in the next chapter, no linguistic system is naturally over-specific, if it is ever over-specific. On the other hand, many graphical systems are naturally over-specific.

What about the free ride phenomenon? We can characterize a system's capability of providing a free ride in the following way:

Definition (Free ride) Let Θ_1 be a chunk of target information and θ be a piece of target information of a representation system $\mathcal{R}$, i.e. $\Theta_1 \subseteq ran(\Rightarrow)$ and $\theta \in ran(\Rightarrow)$. The system $\mathcal{R}$ *provides a free ride from* Θ_1 *to* θ iff there are Σ_1 and σ such that:

1. $\mathcal{R}$ projects a source constraint $\Sigma_1 \vdash \sigma$ to the target sequent $\Theta_1 \vdash \theta$;
2. $\Theta_1 \vdash \theta$ holds on the targets of $\mathcal{R}$.

This property of the system guarantees that whenever we present the information Θ_1 by means of certain states of affairs (named "Σ_1" above), we thereby present an additional piece of information θ that follows from the information Θ_1. This makes the system effective for the purpose of updating the representation so that it presents as many pieces of information that is deducible from what has been assumed or obtained. The system of Harry's memory maps, the system of Hyperproof diagrams, the system of WHISPER's block diagrams, the system of geometry diagrams, the system of Venn diagram, and the system of Euler diagrams cited in chapter 2 are all effective in this way.

As is clear from the above characterization, a structural constraint on representations plays a crucial role in a system's capacity of providing each free ride. In particular, the free rides in the examples we witnessed in chapter 2 were ensured by *nomic* constraints such as geometrical and topological constraints on the representations. The sense of "automaticity" of the free rides come from this fact. Stipulations such as syntactic formation rules were partly responsible in some instances, but not fully. We thus introduce the following more specific notion of free ride:

Definition (Natural free ride) Let $\Theta_1 \subseteq ran(\Rightarrow)$ and $\theta \in ran(\Rightarrow)$. The system $\mathcal{R}$ *provides a natural free ride from* Θ_1 *to* θ iff there are Σ_1 and σ such that:

1. $\Sigma_1 \vdash \sigma$ is a nomic constraint on the representations of $\mathcal{R}$;

2. $\mathcal{R}$ projects $\Sigma_1 \vdash \sigma$ to $\Theta_1 \vdash \theta$;

3. $\Theta_1 \vdash \theta$ is holds on the targets of $\mathcal{R}$.

We can also speak about a "compulsory" version of the property of providing free ride:

Definition (Compulsory free ride) Let $\Theta_1 \subseteq ran(\Rightarrow)$ and $\theta \in ran(\Rightarrow)$. The system $\mathcal{R}$ *provides a compulsory free ride from* Θ_1 iff for each Σ_1 that indicates Θ_1, there is some σ such that:

1. $\mathcal{R}$ projects $\Sigma_1 \vdash \sigma$ to $\Theta_1 \vdash \theta$;
2. $\Theta_1 \vdash \theta$ holds on the targets of $\mathcal{R}$.

This property guarantees that whenever we present the information Θ_1, *no matter* what states of affairs we use to indicate it, we thereby present an additional piece of information θ that follows from the information Θ_1. Many systems of representation that we considered in chapter 2 has this stronger property—again by having a *unique* set of states of affairs Σ_1 that indicate Θ_1, along with a state of affairs σ that satisfies the clauses 1 and 2 above. Thus, many of the free rides that we witnessed in chapter 2 were compulsory, as well as natural.

This compulsory nature of a free ride is a welcome addition for the purpose of presenting as many pieces of information that is deducible from the given Θ_1. However, as a strong case of constraint projection, it aggravates the negative aspect shared by every case of constraint projection. That is, it prevents any representation of the system from presenting the information Θ_1 in isolation, and thus makes the system ineffective for the purpose of simply recording only that information that is explicitly obtained. In contrast, the weaker, non-compulsory version of free ride leaves the possibility that there is an alternative set of states of affairs Σ_1' indicating Θ_1 that triggers no free ride.

On our definition, a free ride is a case in which the information we obtain for free is a consequence of the original assumptions that we put in a representation. It is, however, theoretically possible that this "free" information is *not* a consequence of the original assumptions. We may get a "wrong ride," so to speak.

Definition (Wrong ride) Let $\Theta_1 \subseteq ran(\Rightarrow)$ and $\theta \in ran(\Rightarrow)$. The system $\mathcal{R}$ *provides a wrong ride from* Θ_1 *to* θ iff there are Σ_1 and σ such that:

1. $\mathcal{R}$ projects a source constraint $\Sigma_1 \vdash \sigma$ to the target sequent $\Theta_1 \vdash \theta$;

2. $\Theta_1 \vdash \theta$ does not hold on the targets of $\mathcal{R}$.

We will leave the reader to supply an example of representation systems that provides a wrong ride. Notice that the wrong ride is a special case of misprojection of constraint, where the projected constraint $\Sigma_1 \vdash \Sigma_2$ has the singleton Σ_2.

We close our discussion of free ride and over-specificity by noting several facts, which are immediate from our definitions but of crucial importance for the development of the theory in the next chapter:

Proposition (Free ride and constraint projection) If a representation system $\mathcal{R}$ provides a free ride, it projects some source constraint.

Proposition (Over-specificity and constraint projection) If a representation system $\mathcal{R}$ is over-specific in presenting some information, it projects some source constraint.

Thus, the constraint projection by a representation system is a common element in the phenomena of free ride and content specificity. (We mentioned this element informally at the end of chapter 3.) We can similarly show that the projection of a nomic constraint is a common element in the phenomena of natural free ride and natural over-specificity:

Proposition (Natural free ride and constraint projection) If a representation system $\mathcal{R}$ provides a natural free ride, it projects some nomic constraint.

Proposition (Natural over-specificity and constraint projection) If a representation system $\mathcal{R}$ is naturally over-specific in presenting some information, it projects some nomic constraint.

In the next chapter, we will appeal to the property of projecting nomic constraints to distinguish so-called graphical representation systems from so-called linguistic representation systems.

Self-consistency

Our conceptual framework arises from our desire to investigate the constraints governing representations explicitly to capture the exact workings of a representation system in reasoning. Taking the constraints seriously lets us think about the following possibility.

Given a set of states of affairs, we can ask whether these states of affairs can ever hold together in a single situation. For example, the following states of affairs *cannot* hold together: that a set A includes a set B, that B is not empty, and that A and B are disjoint. Likewise, the following states of affairs are not compatible: that blocks b, c, and g are arranged linearly, that b is to the left of c, that c is to the left of g, that g is to the left of b. Each of these sets of states of affairs is *inconsistent*, let us say.

Suppose that there is an inconsistent set of states of affairs Θ_1 within the coverage of a representation system $\mathcal{R}$. Namely, Θ_1 is a subset of $range(\Rightarrow)$, but the members of Θ_1 do not hold together in any target of $\mathcal{R}$. This does not imply that a set of states of affairs Σ_1 that indicate Θ_1 in $\mathcal{R}$ is also inconsistent. For example, you can easily write the set of sentences, "$B \subseteq A$," "$B \neq \emptyset$," and "$A \cap B = \emptyset$." This means that, in this sentential representation system, a set of states of affairs that indicate the inconsistent set of states of affairs is not itself inconsistent. What if, however, there are some constraints on the representations of $\mathcal{R}$ that make any set of states of affairs inconsistent if it indicates Θ_1? Or more generally, what if the constraints on the representations make any set of states of affairs inconsistent if it indicates *any* inconsistent set of states of affairs? In such a system, no representation presents inconsistent information about its target. We simply cannot create such a representation because of the constraints governing the representations. The system is *self-consistent*, so to speak. More formally:

Definition (Self-consistency) A system $\mathcal{R}$ of representation is *self-consistent* iff for all $\Theta_1 \subseteq ran(\Rightarrow)$ and all $\Sigma_1 \subseteq dom(\Rightarrow)$, if:

- $\Theta_1 \vdash \emptyset$ is a constraint on the targets of $\mathcal{R}$, and
- $\Sigma_1 \Rightarrow \Theta_1$,

then:

- $\Sigma_1 \vdash \emptyset$ is a constraint on the representations of $\mathcal{R}$.

Here the constraints $\Theta_1 \vdash \emptyset$ and $\Sigma_1 \vdash \emptyset$ amount to the inconsistency of Θ_1 and Σ_1. No situation supports a member of $\emptyset$. Thus, when the target constraint $\Theta_1 \vdash \emptyset$ holds on the targets of $\mathcal{R}$, it means that no target of $\mathcal{R}$ support all members of Θ_1, i.e., Θ_1 is inconsistent with respect to the targets of $\mathcal{R}$. Similarly for the source constraint $\Sigma_1 \vdash \emptyset$.

Are there any real representation systems that exhibit self-consistency? Several people have noticed the self-consistency of particular representation systems and its importance to overall efficacy of the systems. Thus, every diagram in the system of Hyperproof "depicts a genuine possibility" (Barwise and Etchemendy 1994, p. 50; also p. 75). Indeed, Barwise and Etchemendy (1995) has proven the self-consistency of the system of Hyperproof diagrams in their model-theoretic analysis of the system (PROPOSITION 1). One of the graphical systems cited by Stenning and Inder (1995) "cannot represent inconsistent set of propositions" (p. 315). The above definition of self-consistency makes it explicit that the property is due to the matching between the constraints governing the representations and the constraints on the target of a representation system. It also makes clear what kind of matching ensures self-consistency of a representation system.

We should, however, take one more step, and distinguish the self-consistency ensured by nomic constraints from that ensured by stipulative constraints:

Definition (Natural self-consistency) A system $\mathcal{R}$ of representation is *naturally self-consistent* iff for all $\Theta_1 \subseteq ran(\Rightarrow)$ and all $\Sigma_1 \subseteq dom(\Rightarrow)$, if:

- $\Theta_1 \vdash \emptyset$ is a constraint on the targets of $\mathcal{R}$, and
- $\Sigma_1 \Rightarrow \Theta_1$,

then:

- $\Sigma_1 \vdash \emptyset$ is a *nomic* constraint on the representations of $\mathcal{R}$.

Thus, in a representation system that is naturally self-consistent, we cannot create a representation system that presents inconsistent information due to nomic constraints on the representations, not due to pure syntactic stipulations placed upon them. I have never seen a system whose self-consistency is ensured purely by syntactic stipulations, but it is important to make the conceptual difference explicit.

What makes self-consistency an important property of a representation system? How does it affect the overall efficacy of the system? Its importance comes from the fact that it gives the system "model-theoretic capacities." More specifically, a self-consistent system lets the user ascertain the consistency of a given information set.

Consequently, it lets the user prove that a piece of information does not follow from the given information. How? Because every representation of the system presents only consistent information, we can prove the consistency of a given information set just by constructing a representation that presents the information set in question. We can ascertain the non-consequence of information θ from the information set Θ_1 by proving the consistency of the the information set $\Theta_1 \cup \bar{\theta}$, namely, the consistency of the information that Θ_1 hold but θ does not hold.

We witness these procedures of checking consistency and non-consequence in everyday life. We draw the picture of a certain furniture arrangement and think that we have ascertained the possibility of that particular arrangement. We test the feasibility of a particular design of a building by constructing a rough, "mashing" model of the building. Here we trust that whatever we draw or build presents a genuine possibility about our targets. That is, we are relying on the self-consistency of the systems of pictures and of mashing models. In contrast, nobody try to prove the consistency of the furniture arrangement or the building design just by constructing the sentences in English that describe them. We simply do not believe in the self-consistency of the system of English declarative sentences.

We see the same use of the self-consistency of a representation system in more formal settings. Gelernter (1959) shows how his theorem-prover, the Geometry Machine, cuts the number of candidate subgoals of a geometry proof by drawing a geometry picture that shows the non-consequence of some subgoals. The geometry machine dramatically reduces the time of its theorem-proving procedure in this way. The logical system of Hyperproof exploits the self-consistency of its diagrammatic subsystem to formally prove consistency and non-consequence, as well as the independence of a particular type of information from the given set of information. Equipped with this "model-theoretic" capacities, the logical system of Hyperproof handles a variety of problems that can not usually handled in a proof-theoretic manner. Both the system of Hyperproof diagrams and Gelernter's system of geometry diagrams are naturally self-consistent, namely, their self-consistency is ensured by nomic constraints on their representations.

Thus, self-consistency is an important property of a representation system that accounts for a significant difference of functionality between systems with and without the property. Our conceptual framework lets us characterize the property formally, making explicit that it is a particular instance of the matching between the constraints on representations and the constraints on targets.

In fact, we can show that except for a non-trivial case, the property is due to the constraint projections made by the system.

Proposition (Self-consistency and constraint projection) If a representation system $\mathcal{R}$ is self-consistent, and there is an inconsistent subset Θ_1 of $ran(\Rightarrow)$, then there is a constraint that $\mathcal{R}$ projects.

Proof. Since Θ_1 is inconsistent, $\Theta_1 \vdash \emptyset$ is a constraint holding on the targets of $\mathcal{R}$. Let Σ_1 be the subset of $dom(\Rightarrow)$ that indicate Θ_1. (There is such a subset since $\Theta_1 \subseteq ran(\Rightarrow)$.) Since $\mathcal{R}$ is self-consistent, $\Sigma_1 \vdash \emptyset$ is a constraint on the representations of $\mathcal{R}$. But by the definition of the indication relation, $\emptyset$ indicates $\emptyset$. Thus, $\mathcal{R}$ projects $\Sigma_1 \vdash \emptyset$ to $\Theta_1 \vdash \emptyset$.

For a similar reason, the following also holds:

Proposition (Natural self-consistency and constraint projection) If a representation system $\mathcal{R}$ is naturally self-consistent, and there is an inconsistent subset Θ_1 of $ran(\Rightarrow)$, then there is a nomic constraint that $\mathcal{R}$ projects.

Thus, except for the trivial cases in which the entire information that covered by a system is consistent, the projection of nomic constraints is a necessary condition for the natural self-consistency of the system.

4.4 Conclusion

We started this chapter by explicating the main concepts of our framework, such as situation, state of affairs, target domain, source domain, constraint, and representation system. We then compared our framework with the logic framework, to highlight the advantage and disadvantage of our framework. It should be now clear that the main thrust of our conceptual framework is the desire to spell out the constraints on representations and their interactions with the constraints on targets, which we believe to capture the exact workings of representation systems in reasoning. In the final section, we applied our conceptual framework to define what it is for a system to project a constraint, and characterized several interesting properties of representations systems on the basis of that concept. In particular, we refined our analyses of free ride and content specificity given in chapters 2 and 3, and also characterized the property of self-consistency that provide a system with certain model-theoretic capacities. Our analyses make it clear that all these properties are special instances of matching between constraints on representations and constraints on targets. Since each property has a significant impact on the overall efficacy of a representation system, our analyses in this chapter lends a strong support to the Constraint Hypothesis, which urges explicit investigation of the patterns of matching between constraints on representations and constraints on targets.

Chapter 5

Graphical and linguistic modes of representation

On the common conception, pictures, images, and diagrams are graphical representations. Sentences in a first-order language and declarative sentences in natural language are linguistic representations. This much is clear and obvious. But what distinguishes graphical representations from linguistic representations? What exactly is the boundary between graphical representations and linguistic representations? To these questions, the common-sense conception does not give us a clear answer. The conceptual boundary between graphical and linguistic sentences seems to be there, but we are ill-prepared to tell where.

This chapter is devoted to the analysis of the conceptual boundary between graphical and linguistic representation systems. For this purpose, however, we do not have to find necessary and sufficient conditions for a system to be a graphical representation, or for a system to be a linguistic representation. It is sufficient to find a property shared by all graphical systems but not by any linguistic systems, as indicated in Figure 5.1 below. (This also amounts to finding a property shared by all linguistic systems but not by any graphical systems.) If we find such a property P, we can use the presence of P for a proof of a system's not being linguistic, and the absence of P for a proof of a system's not being graphical. If we know that a system is either graphical or linguistic, then P also gives us a *positive* test for the system's being graphical or linguistic.

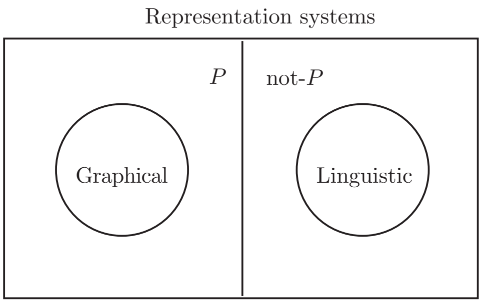


Figure 5.1

We do not demand P to be a sufficient condition for being graphical nor not-P to be a sufficient condition for being linguistic. But we do demand P to be a fundamental property of graphical systems in the sense that it accounts for other properties that are commonly attributed to graphical systems. Likewise, we demand not-P to be a fundamental property of linguistic systems in the sense that it accounts for other properties of linguistic systems. There might be a god who can name all graphical systems, while naming no linguistic systems. Although the property of being-named-by-the-God lets us draw a line between graphical and linguistic systems, we do not take it as a sufficient answer to our question. The property of being-named-by-the-God is not fundamental to graphical systems: it does not explain other properties of graphical systems at all. Likewise, the property of not-being-named-by-the-God is not fundamental to linguistic systems.

In the following, we will start out with examining the previous attempts of drawing a line between graphical and linguistic representation systems (section 5.1). After showing why the proposed contrasts are not sufficient for our purpose, we will formulate our own proposal (section 5.2). Using the concepts introduced in chapter 4, we will claim that all graphical systems project at least some non-trivial nomic constraints, while no linguistic systems do so. In other words, a graphical system has some natural constraints that prevent it from expressing certain information in a particular way without expressing certain other information, while a linguistic systems has no such natural constraints. By way of initial demonstration of this claim's plausibility, we will show how the proposed contrast handles some of the difficult borderline cases between graphical and linguistic systems. We will then move on to show that the proposed contrast is "fundamental" by using it to explain different properties commonly attributed to graphical and linguistic systems (section 5.3). We will combine our proposal with the analyses of the properties of providing free rides, content-specificity, and self-consistency given in the previous chapters, and explain major differences in efficacy between graphical and linguistic representations. Since this amounts to explaining the difference of efficacy of two important classes of representation systems in terms of constraint projection, our discussion as a whole lends a strong support to the main thesis of this dissertation, i.e. the Constraint Hypothesis.

5.1 Some proposals on the boundary

As we will see below, people use different terms to classify graphical and linguistic systems: "analogical" versus "Fregean" (Sloman), "analog" versus "propositional" (Palmer, Lindsay), "graphical" versus "sentential" (Stenning, Oberlander, Inder), "diagrammatic" versus "linguistic" (Larkin, Simon, Barwise, Etchemendy, Hammer). Moreover, it is not certain that all these pairs of terms are intend to classify the same pair of classes of representation systems. To avoid involvement in the terminological muddle, we make the following note before we start examining the analyses given by these authors. All that we do in this section is to extract the theoretical contrasts that they make between two classes of systems that they designate with their favorite terms, and see if we can *apply* their ideas to draw the line between the two classes of representation systems that we are interested in, namely, the class of graphical systems and the class of linguistic systems. Thus, some of our criticisms on their analyses should not be taken literally. The analyses may work for the pair of classes that those authors are interested in.

Digital versus analog

Following Goodman (1968), let us say that a representation system is *analog* if it uses an infinite class of states of affairs to indicate an infinite class of information and each class is dense, namely, its members can be ordered in the way that between each pairs of elements there is another element. Let us say a representation system is *digital* if it uses a discrete class of states of affairs to indicate a discrete class of states of affairs.[1] For example, the analog speedometer on an automobile affords an analog representation system about the speeds of the vehicle since there is a dense class of states of affairs (positions of the pointer) that can hold in the meter and this class indicate a dense class of possible speeds of the vehicle. The light on the dashboard that registers oil pressure affords a digital representation system because there are only two states of affairs (on and off) that indicate information (high and low) about the oil pressure. Both classes of indicating states of affairs and indicated states of affairs are discrete.[2] Of course, a single representation system can be both analog and digital with respect to different subsets of information within its coverage.

One may be tempted to use this distinction between analog and digital to draw

[1] Goodman also requires "differentiability" of these discrete members, but let us not worry about it.

[2] We borrow these examples from Dretske (1981, p. 136).

a line between graphical systems and linguistic systems. Thus, it might be proposed that a linguistic system is digital with respect to the entire set of information it covers, while a graphical system is analog with respect to the entire set of information it covers. However, Goodman (1968) has effectively cut this line of proposal. He says (p. 68):

> Diagrams, whether they occur as the output of recording instruments or as adjuncts to expository texts or as operational guides, are often thought—because of their somewhat pictorial look and their contrast with their mathematical or verbal accompaniments—to be purely analog in type. Some such as scale drawings for machinery, are indeed analog; but some others, such as diagrams of carbohydrates, are digital; and still others, such as ordinary road maps, are mixed.

Since diagrams of carbohydrates are digital with respect to the entire sets of information they cover, Goodman's observation also precludes the suggestion that a graphical system is analog with respect to *at least a part* of the information set it covers. Linguistic systems might be all digital, but the property of being digital is shared by some graphical systems.

Distinctions in terms of syntactic bases

Some authors try to contrast graphical and linguistic representations by focusing on the syntactic structures of representations on which the interpretations of the representations are determined. There are two lines of proposals of this kind.

Some authors assert that the syntactic structure of a linguistic representation can be adequately analyzed as a linear concatenation of symbols, while that of a graphical representation cannot be so analyzed. Thus, according to Larkin and Simon (1987), a sentential representation is a "data structure in which elements appear in a single sequence," while a diagrammatic representation is a "data structure in which information is indexed by two-dimensional location" (p. 72). Thus, in a sentential representation, "each element is 'adjacent' only to the next element in the list," while in a diagrammatic representation, "many elements may share the same location, and each element may be 'adjacent' to any number of other elements" (p. 107).[3] Likewise, Stenning and Inder (1995) suggest that "the essential property of pure linguistic modalities that sets them off from graphical ones is that the only inter-word relation which is interpreted is concatenation" (p. 319). Bertin (1973) seems to have the same contrast in mind when he calls a system of mathematical notations "linear" while characterizing a graphical system handles three variables, the variation of marks and the two dimensions of the plane, to present information (p. 3).

This contrast, however, fails to account for the possibility of "linear" diagrams. Consider a system of "position" diagrams, frequently used to solve a GRE-style problem concerning the seating of people on linearly arranged chairs. We use, for example, the following representation to mean that Amy is at the leftmost seat, Mary is at the second from left, nobody is at the middle seat, Kelly is at the second from right, and there may or may not somebody at the rightmost seat:

A M _ K X

Figure 5.2

We can specify the syntactic structure of this representation in terms of the positions of symbols "A," "M," "_," "K," and "X" in a one-dimensional arrangement and can determine the semantic content of the representation on the basis of that syntactic specification. In fact, we can imagine a *system* of representation all consisting of diagrams of this kind. There seems no reason not to call it a diagrammatic, and therefore graphical, system.

[3]Note that Larkin and Simon do not take this to be the "whole story" about the diagrammatic-linguistic distinction.

Also consider the following representation, meant to present the relative distances of several local cities from Indianapolis:

Figure 5.3

For example, we can read off the information that Bloomington is about 50 miles away from Indianapolis from fact that the word "Bloomington" is 25 millimeters away from the word "Indianapolis." Since the word "Louisville" appears to the right of the name "Bloomington," we are also allowed to read off the information that Louisville is further away from Indianapolis than Bloomington is. Under these interpretation rules, the above representation is clearly a diagram, even though we can specify its content with reference to only the one-dimensional arrangement of symbols. Thus, although linear, sequential concatenation may be a common syntactic property shared by all linguistic representations, it does not set them off from graphical representations.

The other line of proposal asserts that linguistic representations always use a special symbol to express a property or a relation holding in its target, while graphical systems lack such "relation symbols" and use the spatial properties and relations among "object symbols" for that purpose. For example, in distinguishing "analogical systems" from linguistic "Fregean" systems, Sloman (1971, 1995) indicates that analogical representations use "properties of an relations between parts of the representing configuration" to represent "properties and relations of parts in a complex represented configuration" (1971, p. 216) without recourse to "explicit symbols" for properties and relations (1995, p. 13). Palmer (1978) asserts, "Propositional representations are simply those in which there exist relational elements that model relations by virtue of themselves being related to object elements" (p. 294).

Note that neither Palmer nor Sloman is committed to the view that some representations can represent a property or a relation among objects without having their elements stand in a certain relationship. Even when a representation has a relational symbol, the symbol must still stand in a certain relationship to other symbols (often object symbols) in order to express the fact that objects in the represented situation stand in a certain relationship. As Wittgenstein (1921/1971) points out, it is not that the complex sign "aRb" says that a stands to b in the relation R, but that that "a" stands to "b" via "R" says that aRb (3.1432). Thus, the suggested contrast is not that one kind of representations represent a relation by a relation, while the other kind represent a relation by a relation symbol. The contrast is that one kind of representations represent a relation by a relation among object symbols *only*, while the other represent a relation by a relation among object symbols *plus* a relation symbol.

Can we apply this contrast to draw a line between graphical and linguistic representations? Unfortunately, we can imagine a linguistic system that does not use relation symbols. Consider a first-order language L with only one predicate symbol "$Left\_of$." In L, the following strings of symbols are all well-formed sentences:

$$Left\_of(a, b),\ (\forall x)Left\_of(c, x)$$
$$(\exists x, y)(Left\_of(x, b) \vee Left\_of(a, y))$$

Imagine modifying L slightly into another language L', where instead of using a predicate symbol "$Left\_of$," we use a concatenation of individual terms to mean what "$Left\_of$" means in L. Thus, in L', the following strings of symbols are all well-formed sentences:

$$ab,\ (\forall x)cx$$
$$(\exists x, y)(xb \vee ay)$$

This change does not seem to make our language "non-linguistic," although L' uses no relation symbol. If someone objects that "∀," "∃," "∨" can be taken as relation symbols in a broad sense, then we can think of a quantifier-free, connector-free version of L'. Although it is a extremely poor language, it is still a system of representation, and it is linguistic. This illustrates that the use of relational symbol is not essential in our conception of linguistic system.

Homomorphism

It has been often suggested that graphics, especially pictures, "resembles" what they represent. Sometimes, the resemblance is used to account for the very relation of representation between graphics and their targets, namely, to characterize what it is for a graphic to represent its target. Goodman (1978, p. 4) attacks this view rather forcefully, saying that the relation of resemblance is reflexive and symmetric while the relation of representation is neither reflexive nor symmetric. He also note that two automobiles that come off an assembly line do not necessarily represents each other, although they are very similar. Thus, the resemblance theory seem to fail as an account for the representation relation between graphics and their targets.[4]

Still, the resemblance theory may be true as a theory about the difference between graphics and linguistic representations. Namely, whatever the relation between graphics and their targets might be, it might be still true that graphics resemble what they represent to a greater extent than linguistic representations resemble their targets. This theory is certainly not subject to Goodman's criticism. Thus, Barwise and Etchemendy write (1990b, p. 22):

[4]Note that Goodman's notion of "representation" is different from ours in that it only means the representation relation between a certain type of graphical representations and their targets. However, if his claim is true about the representation relation in his sense, then it seems obvious that it is true about other kinds of representation relation.

> Another advantage of diagrams...is that a good diagram is isomorphic, or at least homomorphic, to the situation it represents, at least along certain crucial dimensions....By contrast, the relationship between the linguistic structure of a sentence and that of its content is far more complex. It is certainly nothing like a homomorphism in any obvious way.

To apply this idea to draw a line between graphical and linguistic representation systems, we need spell out the isomorphism or homomorphism that holds between graphics and their targets in the way that excludes linguistic representations. Barwise and Hammer (1995) is apparently undertaking this task. According to their characterization, a representation system is more or less "diagrammatic/homomorphic" in virtue of having more or fewer of the features listed below, and also in virtue of having stronger or weaker versions of them (pp. 44–45):

1. Objects in the target situations, "target objects," are denoted by objects in the representations, "icon tokens," with different types of objects represented by different types of tokens.[5]
2. If a representation s is true of a target situation t, then:
   (a) If icon tokens in s stand in some relevant relationship R, then there holds in t a relationship R', represented by R, among the target objects that they denote.
   (b) The converse holds as well.
   (c) If a grammatical relationship R among icon tokens has some structural property (such as transitivity, asymmetry, irreflexivity, etc.), then this same property must hold of the target relation R' represented by R.
   (d) The converse holds as well.

[5]Barwise and Hammer assume that types, properties, and relations of target objects are represented by properties and relations of icon tokens.

(e) If a token k of some type T has some special property P in s, then the target object k' is of the corresponding type T' having the corresponding property P' in t.

(f) The converse holds as well.

3. Every representation is true in some target situation.

On this definition, the system of Venn diagrams is rather diagrammatic (homomorphic). The circles that appear in a Venn diagram denotes classes in the target situations (feature 1). The relation of disjointness between two classes is denoted by the relation of having-the-overlapping-region-shaded, which is symmetric just as disjointness is symmetric (features 2c, 2d). The system of Euler diagrams would be more diagrammatic than the system of Venn diagrams because, in addition to the fact that each circle stands for a class in the target, set-inclusion and set-disjointness are denoted by circle-inclusion and circle-disjointness, and while set-inclusion is transitive and set-disjointness is symmetric, circle-inclusion is transitive and circle-exclusion is symmetric. And it is impossible for an Euler diagram to present an inconsistent set of information (feature 3). Perhaps, photographs will be classified as highly diagrammatic (or homomorphic); sentences of a first order languages will be one of the least diagrammatic (or homomorphic).

In general, Barwise and Hammer's homomorphism criterion seems to do well in capturing our intuition about the degrees in which systems are more or less graphical. It, however, does not exactly fit our purpose. First, it is not clear how the criterion is related to other properties commonly attributed to graphical representation systems, especially, the kinds and degrees of their efficacy. Intuitively, the representations of highly homomorphic systems on this criterion will let us "see through" the structure of what they represent. But it has not been specified exactly what is nice about this "seeing through" and exactly which of the listed features is responsible for this capacity.

Barwise and Hammer note that the "close relationships" between such representations to the structures that they represent "allows one to deductively establish facts usually obtainable only model-theoretically" (p. 51). Presumably, this capacity stems from feature 3 in the list. Feature 2a and 2b may be also responsible. Since they do not elaborate this point, we cannot guess any further. But notice that feature 3 is the property of self-consistency discussed in the last chapter, and the features 2a and 2b are special cases of the constraints on representations matching with the constraints on their targets. This suggests that the consideration of constraint matching may provide a more principled criterion for a system's being graphical that explains its semantic capacity and other properties concerning its efficacy. As Hammer (1995, p. 11) and Barwise and Etchemendy (1995, p. 214) later suggest, we perhaps should take the above list as the "hallmarks" of "good" graphical representations, rather than a criterion for being graphic.

Content specificity

No matter how we may decide to interpret Barwise and Hammer's homomorphism criterion, it is not meant to provide a "cut-and-dried, definitive definition" of diagrammaticality, but to provide a metric with which we measure the degree in which a given system is more or less diagrammatic. Unlike Barwise and Hammer, Stenning and Oberlander (1995) propose a strong contrast between graphical representation systems and linguistic representation systems. The proposal is centered around the notion of "specificity" of representation systems. After characterizing specificity as the feature of a representation system that "compels specification of information, in contrast to systems that allow arbitrary abstraction" (p. 99), they identify specificity "as the feature distinguishing graphical and linguistic representations, rather than low-level visual properties of graphics" (p. 98).[6] They also claim that although the specificity of a representation system leads to the expressive weakness of the system, it also "aid processibility" (p. 98) of the information represented in the system. From this consideration, they think that the specificity "helps explain why graphical techniques, such as Euler's circles, for teaching abstract reasoning are so widespread, and presumably effective" (p. 99).

It is our everyday experience that graphical representations tend to be specific in their information content. We have seen plenty of examples in chapter 3. It seems also intuitively true that the information represented in a graphical mode is easier to process. So, if one could offer a clear explanation of (a) why graphical systems tend to be more specific than linguistic systems, and (b) how the specificity of a graphical system makes the information easy to process in a way linguistic systems do not, then the theory would be indeed attractive, and perhaps, we can use the notion of specificity to draw a line between graphical and linguistic representation systems.

Unfortunately, Stenning and Oberlander have not offered an explanation of either kind. They admit, "Of course, in neither our view nor Levesque's is this property [of specificity] confined to visual representation" (p. 107). (Here we can take "visual representations" as opposed to linguistic representations.) Thus, to use specificity as the base of distinguishing graphical representations from linguistic representations, we need a description of the way in which graphical systems has this property as distinguished from the way a linguistic system has it. Stenning and Oberlander have little to say on this point. When they discuss the system of Euler circles (ECs), they say that the expressive power of graphical systems such as ECs are "inherently weak"

[6]Perhaps, we can consider their notion of "specificity" to be the same as the notion of content specificity that we characterized in chapter 3, although it is not certain how much they would agree with our analysis.

(p. 132) and has "natural limit" (p. 119) while there is no "inherent reason" why linguistic systems have any of the limitations that graphical representations exhibit (p. 123). Although the words "inherent" and "natural" seem to point to the direction in which our own theory goes in contrasting graphical systems and linguistic systems, the scarcity of explanation on the part of Stenning and Oberlander does not allow us to be certain that they are really in agreement with our theory.

Even if graphical systems and linguistic systems do not differ in the ways they have the specificity property, it might be that the specificity of graphical systems leads to the tractability of inferences in the way that the specificity of linguistic systems do not. Then, we could use this difference as the base for distinguishing graphical systems from linguistic systems. Unfortunately, Stenning and Oberlander's explanation of how specificity of a representation system aids processibility uses linguistic systems as model cases, and base their account on the fact that the syntactic constraints on their systems are "very similar to" those on linguistic systems that Levesque has shown to afford tractable inferences (p. 107). Thus, they have not shown the difference, if any, between the ways specificity of graphical and linguistic systems aid processibility.

Availability of expressive limitation

Stenning and Inder (1995) try to supplement this second lack, in terms of the notion of "cognitive availability of the limits of expressive power" (p. 304). On the basis of the works of Levesque and of Stenning and Oberlander just mentioned, Stenning and Inder assume that a representation system with a limited expressive power affords more tractable inferences. But they take an additional step, and note that how much a user knows about the scope of a given representation system is a "critical determinant of cognitive properties of the system" for the user (p. 314). Even when a system is expressively weak, and has a potential for easier processibility of the information represented, “exploiting this fact relies on being aware of it” (p. 318); “availability determines whether the weakness of the representation can be exploited” (p. 304).

Stenning and Inder then use the difference in cognitive availability of the limits of expressive power to contrast graphical systems and linguistic systems: “the difference between the graphical and linguistic systems lies in the discoverability of the limitations on expression and the necessary methods of exploiting them in inference” (p. 325). They move on to consider particular graphical systems to illustrate how a graphical system makes its limits on expressiveness more available to the user than a linguistic system does. According to them, once the user understands the core of the interpretation of graphical representations, then the user can infer “quite intricate meta-logical properties” about the system, concerning what limitations are there on the expressiveness of the system. According to Stenning and Inder, the inferences of this sort rely on the “diagrams’ geometry/topology” (p. 318) and arise from “graphical constraint” (p. 334). In the case of a linguistic system, however, the inferences of this sort do not arise due to the paucity of syntactic structure of linguistic representations.

I believe that Stenning and Inder is right in pointing out that there is a difference in the ways we infer meta-logical facts about the expressive capacity of a graphical system and that of a linguistic system, and that this difference affects the overall usability of graphical and linguistic systems. Stenning and Inder, however, have not shown *how* this difference of cognitive availability of expressive capacities account for differences in usability. Take the example of inferential tractability that they cite frequently. Assuming that the limits of expressibility of a system is fully available to a user, how does the user go about exploiting his or her knowledge about the limitation to make efficient inferences? Since Stenning and Inder are silent about this point, there is no way to assure that this knowledge is really crucial in all, many, or at least some, instances in which a graphical system appears to afford more efficient inferences than a linguistic system does.

To summarize, nothing in Stenning and Oberlander's theory accounts for the different ways in which graphical systems and linguistic systems have the limitations on expressive power. Stenning and Inder try to explain the difference between graphical and linguistic systems in terms of cognitive availability of the limitations of expressive power to the user. Indeed, there seems to be such a difference between some pairs of graphical and linguistic systems. But since Stenning and Inder are silent about how this difference in cognitive availability affect the usability of graphical and linguistic systems, it is not clear that the proposed difference is a fundamental one that explains any of the observed difference of efficacy of graphical and linguistic systems. Of course, our argument does not show that their proposal is false. But it does show that the proposal is so under-developed that it is not of use for our purpose.

Inherent constraints on representations

We saw earlier that Palmer (1978) tries to contrast "analog" and "propositional" representations in terms of the presence and absence of relation symbols with semantic significance. We saw that this contrast of analog-propositional cannot be applied to distinguish graphical representations from linguistic representations due to the existence of counterexamples. Apart from this proposal based on "surface manifestation" of analog and propositional representations, however, Palmer has another proposal about their distinction. In fact, this second proposal of his is very close to our own proposal on the graphical-linguistic distinction to be developed later.

Palmer sees a representation and its target as two worlds, the "representing world" and the "represented world" (p. 262). Each world comprises objects that are related in particular ways. Objects in the represented world are denoted by objects in the representing world, and the ways the latter objects are related in the representing world model the ways the corresponding objects are related in the represented world. Thus, he is assuming some semantic correspondence at the level of relations, namely, from the relations holding in the representing world and the relations holding in the represented world. Not all relations holding in the representing world correspond to a relation in the represented world; not all relations holding in the represented world are "preserved" in the representing world. However, Palmer's definition of "representation" requires that the representing world must preserve at least some relations of the relations holding in the represented world. Apparently, Palmer excludes from his definition of representation any "false" representation—he excludes any representing world that preserves a relation absent from its target world.

In Palmer's words, the contrast between propositional and analog representations consists in the fact that "whatever structure there is in a propositional representation exists solely by virtue of the extrinsic constraints placed on it" (p. 296) while "whatever structure is present in an analog representation exists by virtue of the inherent constraints within the representing world itself" (p. 297). Here, "intrinsic constraints" means the structural constraints on representing relations, such as irreflexivity, asymmetry, transitivity of the *above* relation, and "interdimensional constraints" such as the determination of the area of the rectangular from the lengths of their sides (p. 273). In contrast, "extrinsic constraints" means constraints "imposed from outside" (p. 271) on the representing world in order to make its structure "conform to the represented world" (p. 296). Thus, in our own conceptual framework, Palmer's intrinsic constraints are instances of nomic constraints on representations and his extrinsic constraints are generalizations of our stipulative constraints on representations.

Palmer's notion of "exist by virtue of inherent (or extrinsic) constraints" may not be immediately clear. For, generally, representations are artifacts, something that we (cognitive agents) create in one way or another. So, the structure in a representing world is usually the result of our operations, for the purpose of presenting information about the represented world. At least partly, then, the structure exists by virtue of our own operation. What Palmer seems to assume when he uses the phrase "exists by virtue of inherent (or extrinsic) constraints" is that the results of our operations of recording information about a target world can be affected by the intrinsic or extrinsic constraints imposed on the representing world. Then, Palmer's proposal can be translated in the following way: all the structures in an analog representation is the result, affected by nomic constraints on representations, of our operations of presenting information in the representation, while all the structures in a propositional representation is the result, affected only by stipulative constraints on representations, of our operations of representing information.

Palmer's account on the contrast between analog and propositional representations has several limitations. First, the content of the proposal is not sufficiently clear. In particular, he does not give any general account of what it is for an intrinsic (or extrinsic) constraint to hold on a representing world and what it is for a structure in a representation to exist by virtue of such a constraint. Accordingly, our own reconstruction of his proposal, shown above, is not transparently clear. Secondly, he does not really argue for his proposal. In particular, he does not motivate it by showing what other properties of analog and propositional representations can be accounted for in terms of the proposed contrast. Aside from the details, however, the insight behind his proposal seems quite right. The reader will notice the same basic intuition behind our account of graphic and linguistic representation systems. Unlike Palmer, however, we will endeavor to make the content of our proposal clear within the conceptual framework explicitly defined in chapter 4. We will also try to explain, on the basis of our proposal, other properties commonly attributed to graphical and linguistic systems.

5.2 Proposal

Recall the definition of constraint projection in a representation system and the definition of nomic constraint on representations, both given in Chapter 4. Using these two notions, we propose the following contrast between graphical representation systems and linguistic representation systems:

Main Proposal All diagrammatic representation systems project some non-trivial nomic constraints. No linguistic systems do so.

Put informally, the proposal is that all graphical systems have some non-trivial nomic constraints on the structures of their representations that prevent them from presenting certain information in particular ways without presenting certain other, while no linguistic systems have such constraints on the structures of representations.[7]

In the following, we will give some evidence that the proposed contrast actually holds. The next section will show that this contrast is in fact fundamental, in the sense that it accounts for other properties of graphical and linguistic representation systems that are commonly observed.

One important test for the proposal of conceptual distinction is whether it can handle borderline cases well. We consider two sets of such cases. Consider the following diagram, cited in the last section:

[7]Recall from chapter 4 that a constraint $\Gamma \vdash \Delta$ is *non-trivial* if Δ is not a subset of Γ. Henceforth we will use the word "constraint" to mean "non-trivial constraint," suppressing the word "non-trivial."

Figure 5.4

Some of the proposals discussed in the last section incorrectly classifies the above as linguistic. How does our proposal classify the system behind this representation? Consider the following three states of affairs holding in this representation:

(σ_1) The word “Bloomington” is 25 millimeters to the right of the word “Indianapolis.”

(σ_2) The word “Louisville” is 57 millimeters away to the right of the word “Indianapolis.”

(σ_3) The word “Louisville” appears to the right of the name “Bloomington.”

These states of affairs constitutes a non-trivial nomic constraints governing the representations of the system: the first two states of affairs jointly entail the third. But on the semantic convention we adopt for the representation, these states of affairs indicate the following pieces of information:

(θ_1) Bloomington is about 50 miles away from Indianapolis.

(θ_2) Louisville is about 114 miles away from Indianapolis.

(θ_3) Louisville is further away from Indianapolis than Bloomington is.

Thus, on our definition of constraint projection, the system projects the non-trivial nomic constraint $\{\sigma_1, \sigma_2\} \vdash \sigma_3$ to the target sequent $\{\theta_1, \theta_2\} \vdash \theta_3$. Because of this projection, our proposal classifies the present system as non-linguistic. Assuming that the system is either graphical or linguistic, then, our proposal tells us that it is indeed a graphical representation system.

Let us consider some other borderline cases. Here is a simplified calendar describing my schedule on the next three days:

Meet B at 11am Mon	Eat with G at noon Mon	Go theater at 8pm Mon
Teach Logic at 2pm Tue	Hold office hour at 4pm Tue	
Attend G's talk at 10am Wed	Meet F at 1pm Wed	Attend K's talk at 7pm Mon

Figure 5.5

Each sentence in this calendar describe an event to happen. In addition, the semantic convention associated with the calendar allows the spatial relationships among sentences to mean something: it allows the *left-of* relation between sentences to indicate the temporal order of events described by the sentences, and allows vertical positions of sentences to indicate the precedence of the days on which the described events happen.

Is this system graphical or linguistic or anything else? My intuition is that the system is not purely linguistic, and rather graphical, and I believe that the reader would agree. In fact, this system projects a non-trivial nomic constraint. For example, the state of affairs that the sentence "Meet B at 11am Mon" is left of the sentence "Eat with G at noon Mon" and the state of affairs that the sentence "Eat with G at noon Mon" is left of "Go theater at 8pm Mon" jointly entail the state of affairs that "Meet B at 11am Mon" is left of "Go theater at 8pm Mon." Let these states of affairs be σ_4, σ_5, and σ_6 respectively. The states of affairs σ_4, σ_5, and σ_6 indicate the following pieces of information respectively:

(θ_4) My meeting J at 11 AM on Monday is earlier than my eating with G at Monday noon.

(θ_5) My eating with G at Monday noon is earlier than my going to the theater at 8 PM on Monday.

(θ_6) My meeting J at 11 AM on Monday is earlier than my going to the theater at 8 PM on Monday.

Thus, the system can be said to project the non-trivial nomic constraint $\{\sigma_4, \sigma_5\} \vdash \sigma_6$ to the target sequent $\{\theta_4, \theta_5\} \vdash \theta_6$. Likewise, the state of affairs that the sentence “Eat with G at noon Mon” is above the sentence “Teach Logic at 2pm Tue” and the state of affairs that the sentence “Teach Logic at 2pm Tue” is above the sentence “Attend G’s talk at 10am Wed” jointly entail the state of affairs that the sentence “Eat with G at noon Mon” is above the sentence “Attend G’s talk at 10am Wed.” Each state of affairs has a semantic value via the indication relation of the system, so this constraint is also projected by the system.

Now imagine impoverishing the semantics of the system slightly, so that the *left-of* relation between sentences no longer indicate the temporal precedence of the described events or anything. Under this modified system, the representation in Figure 5.5 no longer means the same thing as it does under the original system. Would we still call the representation graphical? Yes, but to the lesser extent. On our account, we still think that it is a graphical representation because the system still projects non-trivial nomic constraints (the second constraint considered in the last paragraph, for example), but our intuition is weaker since the system project a smaller number of nomic constraints than the original systems (the first constraint in the last paragraph is no longer projected).

What if we impoverish the semantics further, and does not allow the *above* relation between sentences to indicate the precedence of events at the month level? Do we still call the system graphical? Our intuition is rather clear—the system is sentential, not graphical. Our analysis accounts for this intuition. The only way to present information in this system is to add a sentence in your representation, and adding one sentence does not generally entail adding another sentence. Thus there is no natural constraint on the structures of representations that prevent us from presenting certain information without presenting other. There is no non-trivial nomic constraint that the system projects, in other words. Thus, the system is not graphical. Assuming that the system is either graphical or linguistic, our criterion say that it is a linguistic system.

Thus, our proposed contrast of graphical and linguistic systems handles borderline cases pretty well. Of course, this discussion still does not establish the truth of our proposal. Yet we have at least shown that it is rather promising.

Indeed, the projection of nomic constraint is a weak property than it may first appear, and I believe that the first statement in our proposal hard to refute. One might, however, attempt to refute the second statement, saying that some linguistic systems do project nomic constraints. We will consider two of those possible attempts.

Consider the following set of first-order sentences written on a piece of paper:

$$Dodec(a) \qquad \neg Cube(b) \qquad Dodec(a) \rightarrow Small(b)$$

$$Large(k) \& (\exists x)(Tet(x)) \qquad (\exists y)(Small(y)) \vee (\forall x)(Same\_size(x, k))$$

Figure 5.6

Assume that we can interpret the sentences on the obvious semantic rules. Then the above set of sentences constitutes a representation that describes the spatial arrangement of a limited number of blocks.

Now, one may note the appearance of the complete sentences "$Dodec(a)$" and "$\neg Cube(b)$" in the list, and say that the above representation presents the conjunctive information that a is a dodecahedron and b is not a cube. One may then argue that the presentation of this information is "forced" in the sense that when we put the sentences '$Dodec(a)$" and "$\neg Cube(b)$" in the list to present each conjunct, we inevitably present their conjunction in the list. Thus, the underlying representation system does project some nomic constraint, although it is clearly a linguistic system.

This objection conflates the information presented in a representation with the information inferable from the information presented in the representation. Although the conjunctive information is easily inferable from the two conjuncts of it that are indeed presented in this list, the conjunction itself is not a piece of information presented in the list. If the list had the entry "$Dondec(a)\&\neg Cube(b)$," then the conjunctive information in question would count as a piece of information presented by the representation. Since such a sentence does not appear in the list, the list does not present the information in question. Ask yourself if the following information is presented in the list:

(θ_7) b is small.

(θ_8) b is not cube and k is large.

(θ_9) If there is no small tetrahedron, the a is large.

These pieces of information are all inferable from, but not presented in, the above list, just like the conjunctive information that we are concerned with. The only difference is that these are less immediate conclusions than the conjunction is. For a similar reason, just because the complete sentence "$Large(k)\&\exists x(Tet(x))$" appears in the list, one cannot claim that the information that k is large is presented in the list. The claim is no more justified than the claim the list presents the information that some small tetrahedron exists because it contains the sentence "$\exists y(Small(y)\&Tet(y)) \vee \forall x(Same-size(x,k))$."

Alternatively, one may attempt to refute the second statement in our proposal by citing certain pragmatic phenomena surrounding the use of natural language sentences. Suppose I utter a sentence, "She is insane," using the pronoun "she" "deictically," to refer to Jane, whose odd behavior we are observing. We can consider this utterance a representation that presents the information that Jane is insane. Let us call this information "θ_{10}." Roughly, the utterance presents the information θ_{10} because the following state of affairs holds in it:

(σ_7) I utter the words "she," "is," and "insane" while pointing to Jane.

Now, the information that Jane is insane is not the only information that we gain from this utterance. We also gain information that Jane is a female. Call this information "θ_{11}." We gain the information θ_{11} because I used the word "she" in the utterance, or more specifically, because the following state of affairs holds in my utterance:

(σ_8) I utter the word "she" while pointing to Jane.

Obviously, the state of affairs σ_7 entails the state of affairs σ_8. Since σ_7 and σ_8 indicate θ_{10} and θ_{11} respectively, one might contend that the system of spoken English projects the nomic constraint $\sigma_7 \vdash \sigma_8$ to the target sequent $\theta_{10} \vdash \theta_{11}$. Since the system is a paradigmatic case of linguistic system, this might be considered a clear counterexample to our claim that no linguistic systems project nomic constraints.

Our response to this objection is that although my utterance presents the information θ_{10}, it does *not* present the information θ_{11}, not at least in the sense we assigned to the term "present" in chapter 4. To see this, suppose that Jane is not insane at all, despite her apparently odd behavior. Then, obviously, my utterance is *false*. What it *asserts* turns out to be inaccurate about the situation it represents. For this reason, we say that θ_{10} is a part of the information that my utterance presents in the sense we defined.

Now suppose Jane is in fact insane, but Jane is really not a female, despite his distinctively feminine look. Does this finding make my utterance false? Not quite. Under this supposition, my utterance is *inappropriate*, rather than false. It is inappropriate in the sense that it is based on the wrong *assumption* about Jane's gender. What is wrong is not the assertion made by the utterance, but the assumption under which I made the utterance. Thus, the information θ_{11} is not a part of my utterance's assertion, but a part of its underlying assumption. You extract this information from utterance simply because you trust me that I am not making a mistake about the gender of Jane.

Since our definition of "present" requires that a representation is false if a part of the information it presents is inaccurate about the target it represents, we cannot consider the information θ_{11} a part of what my utterance presents. More specifically, the state of affairs σ_8 does not indicate θ_{11} in the same sense in which σ_7 indicates θ_{10}. Thus, the system of spoken English does not project the nomic constraint $\sigma_7 \vdash \sigma_8$ to the target sequent $\theta_{10} \vdash \theta_{11}$.

Generally, not every information that we extract from somebody's utterance is the information that the utterance *presents* on the basis of the semantic rules of the language in question. Some may be underlying assumptions of the speaker that we have no strong reason not to believe. Among such assumptions are what semanticists call "presuppositions" that the listeners "accommodate" (Lewis 1979). Some may be a conversational implicature that we calculate out on the basis of the Gricean principles of cooperative conversations (Grice 1975). Also, these phenomena may not be limited to the use of natural language sentences, but relevant to the communicative use of representations in general, linguistic or non-linguistic. In fact, it would be important to study the possible impacts of the phenomena of implicature and presupposition accommodation on the overall efficacy of a representation systems as a communication tool. However, the particular notion of constraint projection that we use to distinguish graphical systems from linguistic systems is defined with reference only to the information that representations *present*. Thus, an example concerning the information that we extract through an implicature calculation or a presupposition accommodation can be neither a support or a counterexample to our proposal.

5.3 Applying the proposal

In this section, we apply our proposal on the graphical-linguistic boundary to account for the properties commonly attributed to the two classes of representation systems. In particular, we will be concerned with three major differences in their efficacy, often pointed out in the literature.

Presence and absence of free ride

As we have seen in chapter 2, Sloman (1971), Funt (1980), Larkin and Simon (1987), Lindsay (198?), and Barwise and Etchemendy (1990b) point to the phenomenon of free ride that accounts for the efficacy of certain representation systems in their dynamic use. In chapter 4, we characterized the capacity of a representation system of providing free rides in the following way:

Definition (Natural free ride) Let Θ_1 be a chunk of target information and θ be a piece of target information of a representation system $\mathcal{R}$. The system $\mathcal{R}$ *provides a natural free ride from* Θ_1 *to* θ iff there are Σ_1 and σ such that:

1. $\Sigma_1 \vdash \sigma$ is a nomic constraint holding on the representations of $\mathcal{R}$;
2. $\mathcal{R}$ projects $\Sigma_1 \vdash \sigma$ to $\Theta_1 \vdash \theta$;
3. $\Theta_1 \vdash \theta$ holds on the targets of $\mathcal{R}$.

Now, these writes not only point to the capacity of a representation system to provide natural free rides, but also indicate it as a property that distinguishes certain graphical representation systems from linguistic systems. Thus, the presence of free rides put inferences with diagrams "in stark contrast to sentential inferences" (Barwise and Etchemendy 1990b, p. 22); it accounts for "the major efficiency difference between the sentential and diagrammatic representations" (Larkin and Simon 1987, p. 96); the capacity for free rides is the "crucial property" that distinguishes visual images from propositional representation (Lindsay 1988, p. 112–113).

Our proposal on the graphical-linguistic contrast provides a simple account for the intuitions expressed by these authors. The projection of a nomic constraint is a necessary condition for providing a natural free ride. Linguistic systems do not provide any natural free rides because they do not project any nomic constraints. Since the projection of a nomic constraint is not a sufficient condition for providing a nomic free ride, not all graphical systems provide natural free rides. But when they do, that capacity is therefore a feature that clearly distinguishes them from linguistic systems. The graphical systems discussed in chapter 2, i.e. the systems of Hyperproof diagram, the system of WHISPER's block diagrams, and the system of geometry diagrams, are all of this kind.

Over-specificity versus expressive flexibility

Recall that we characterized the over-specificity of representation systems in the following way:

Definition (Natural over-specificity) Let Θ_1 be a chunk of target information of a representation system $\mathcal{R}$. The system $\mathcal{R}$ is *naturally over-specific* in presenting the information Θ_1 iff: for each Σ_1 that indicates Θ_1, there are Σ_2 and Θ_2 such that:

1. $\Sigma_1 \vdash \Sigma_2$ is a nomic constraint on the representations of $\mathcal{R}$;

2. $\mathcal{R}$ misprojects $\Sigma_1 \vdash \Sigma_2$ to $\Theta_1 \vdash \Theta_2$;

If a system has this property, no representations of the system cannot present the information Θ_1 in isolation, without presenting one of the alternative pieces of information Θ_2 that do not follow from Θ_1. This property therefore limits the expressive generality of the system.

Like the capacity of providing natural free rides, natural over-specificity is often identified as the property that distinguishes certain graphical systems from linguistic system. Thus, while a description of a man may well "fail to mention whether or not the man is wearing a hat," a picture of this man "has to go into details" (Dennett 1969, p. 135); it is "unreasonable use of the word 'image'" to speak of an image that does not exhibit content specificity (Pylynshyn 1973, p. 11); "the expressive generality of a system is often incompatible with its capacity for being diagrammatic" (Barwise and Hammer 1995, p. 47); Fregean systems are superior to analogical systems because "the structure (syntax) of the expressive medium need not constrain the variety of structures which can be represented or described" (Sloman 1971, p. 217). Also, Stenning and Oberlander (1995) explicitly designates specificity as "the feature distinguishing graphical and linguistic representations" (p. 98) although their notion of specificity is not confined to natural specificity in our sense.

Again, our proposal on the contrast between graphical and linguistic systems account for the intuitions expressed in the literature. Since the projection of a nomic constraint is a necessary condition for being content-specific, no linguistic system exhibits natural over-specificity. Since the projection of a nomic constraint is not a sufficient condition for content specificity, not all graphical systems are naturally over-specific. However, if a graphical system does exhibit natural over-specificity, it is a property that definitely distinguishes it from any linguistic system.

Thus, our account explains why natural over-specificity is often considered to be a distinguishing character of certain graphical systems. However, people who note this expressive weakness of graphical systems often point to its connection to the inferential efficiency that the graphical systems provide. More generally, people notice some "trade-off" between the expressive generality and the inferential efficiency. Thus, Sloman (1971) notes that the price of the expressive generality of a Fregean system is the lack of capacity of dealing efficiently with specific problem-domains; Lindsay (1988) discusses the trade-off between the "applicability" of a system and its power of reducing "computational complexity of inference" (p. 130); Stenning and Oberlander (1995) emphasizes that specificity of a representation system aids "processibility" of the information represented in the system (p. 98). Because of this trade-off, a graphical system tends to be over-specific but to afford efficient inferences, while a linguistic system tends to be expressively flexible but inferentially inefficient.

The references to this trade-off in the literature is frequent, but often cursory. It is therefore possible that different people point to different phenomena under the same title. And perhaps, there is no single, simple explanation for it. However, our proposal on the graphical-linguistic representations offers an account of one definite cause of this trade-off. We have seen that, due to the lack of projection of nomic constraints, linguistic systems do not afford natural free rides but, at the same time, they are not over-specific in presenting any information set in its coverage (unless we have syntactic stipulations that force such specificity). This partly explains the tendency of a linguistic system to be expressibly flexible *but* not supportive to efficient inferences.

On the other hand, all graphical systems project nomic constraints, and hence they tend to provide natural free rides. Also, due to the nomic constraints that it projects, representations in a graphical system cannot present some chunk of information in a certain way without presenting one of the alternative pieces of information. No matter whether these pieces of information are consequences of the information to be represented, this property makes the graphical system inflexible in the selection of information set to be presented and the way to present it. If none of the alternative pieces of information are consequences of the information to be represented, it is a case of over-specificity. This partly explains the tendency of a graphical system to be expressively inflexible *but* supportive to efficient inferences.

On this analysis, the currency in the trade-off between expressive generality and inferential efficiency is the projections of nomic constraints. As you get more of them, you get more inferential efficiency, but enjoy less expressive generality. As you get less projections of nomic constraints, you get more expressive generality, but less inferential efficiency.

Model-theoretic capacities

In the last chapter, we analyzed the capacity of representation systems to help ascertain the consistency of given information and the non-consequence of a piece of information from given information. To capture this "model-theoretic" capacity of certain representation systems, we have introduced the following notion of natural self-consistency:

Definition (Natural self-consistency) A system $\mathcal{R}$ of representation is *naturally* *self-consistent* iff for all $\Theta_1 \subseteq \Theta$ and all $\Sigma_1 \subseteq \Sigma$, if:

- $\Theta_1 \vdash \emptyset$ is a constraint on the targets of $\mathcal{R}$, and
- $\Sigma_1 \Rightarrow \Theta_1$,

then:

- $\Sigma_1 \vdash \emptyset$ is a *nomic* constraint on the representations of $\mathcal{R}$.

In comparing graphical and linguistic representation systems, people often point to the natural self-consistency or the model-theoretic capacities as a distinguishing character of certain graphical systems. Sloman (1995) points out the difficulty (impossibility) of drawing an object that is both round and square, compared with the ease of construction of the assertion "$round(obj)\&square(obj)$" (p. 12); diagram are of service "precisely because" it is not possible to construct a diagram that depicts two congruent triangles with unequal perimeters (Lindsay 1988, p. 122); certain representations of graphical systems "allow one to deductively establish facts usually obtainable only model theoretically" (Barwise and Hammer 1995, p. 51).

Combined with the above analysis of self-consistency of representation systems, our proposal on the graphical-linguistic systems can account for this observation. We have seen that the projections of constraints are a necessary condition for a system's being self-consistent; since a linguistic systems do not project nomic constraint, it is not naturally self-consistent, unless the entire set of information within its coverage is consistent. Since the projections of nomic constraints are not a sufficient condition for self-consistency, not all graphical systems are naturally self-consistent. But for those systems with this property, it is a definite property that distinguish them from any linguistic systems. The system of Hyperproof diagrams and the system of geometry diagrams used by Gelernter's thereom prover are such systems. They therefore afford the model-theoretic procedures for checking consistency and non-consequence in the ways linguistics system do not.

5.4 Conclusion

We devoted this chapter to the analysis of the conceptual boundary between graphical and linguistic representation systems. We began with examining various attempts at contrasting the two classes of systems. We found some of the proposals untenable due to counterexamples, others not fundamental enough to explain other properties commonly attributed to graphical and linguistic systems, and still others not entirely false but under-developed. We then moved on to formulate our own proposal, and claimed that graphical systems project at least some non-trivial, nomic constraints, while linguistic systems project no such constraints. We gave initial arguments for the truth of our claims by dismissing apparent counterexamples and demonstrating how it handles some of the obstinate borderline cases. We then discussed how the proposed contrast explain the properties commonly attributed to graphical and linguistic systems. By coupling the proposed contrast with the analyses given in the previous chapters, we explained why the capacity of providing free rides, the over-specificity, and the self-consistency are often considered distinguishing characters of graphical systems as opposed to linguistic systems. We also explained one definite cause of the trade-off of expressive generality and inferential efficiency often observed in graphical and linguistic representation systems. We hope that our discussions have convinced the reader of the plausibility of our analysis of the boundary between the graphical and linguistic systems of representation.

Our discussions in this chapter also lends a strong support to the Constraint Hypothesis that we are advocating throughout this dissertation. For, if our contrast of graphical and linguistic representation systems is correct and our subsequent applications of the contrast are not misguided, it follows that the projection of nomic constraints is a crucial property of graphical systems that explains both of their efficacy and inefficacy, and that the absence of such projection is a crucial properties of linguistic systems that explain both of their efficacy and inefficacy. The matching and mismatching of the constraints on representations and the constraints on their targets captures the difference in inferential potentials of two rather important classes of representation systems.

Of course, we *have not* explained every difference in efficacy present between linguistic systems and graphical system in terms of the projection of nomic constraints. However, we have not exhausted the explanatory potential of the proposed contrast either. It is an open question whether the proposed contrast can explain every difference in efficacy of representation systems due to their graphic and linguistic nature. There may turn out be a property P, shared by all graphical systems but by no linguistic systems, that are more fundamental than the projection of nomic constraints. No matter how it may turn out, we believe that the explicit identification of this property has been an important step in the analysis of the conceptual boundary between graphical and linguistic systems.

Chapter 6

Conclusion

In closing the dissertation, we will review the main claims that we have made and give our own estimates of how they stand after our discussions. We will also indicate possible future directions of the research initiated here.

Beside the Constraint Hypothesis, we have made the following four main claims (or sets of claims).

1. **About free ride.** The property of providing a free ride accounts for the system's ability of supporting efficient and valid updates of the information contents of representations. A system has this property if and only if a constraint of the form $\Sigma_1 \vdash \sigma$ holds on the representations, where Σ_1 and σ indicate the information Θ_1 and θ and the constraint $\Theta_1 \vdash \theta$ holds on the targets in fact.

2. **About content specificity.** The property of over-specificity accounts for the system's lack of expressive generality, namely, its inability to present a certain set of information in isolation, without thereby presenting a piece of information that does not follow from the information set in question. A system has

this property if and only if a constraint of the form $\Sigma_1 \vdash \Sigma_2$ holds on the representations, where Σ_1 and Σ_2 indicate the information Θ_1 and Θ_2 and none of the alternative pieces of information in Θ_2 follow from Θ_1, i.e., $\Theta_1 \vdash \theta$ is not a constraint on the targets for any member θ in Θ_2.

3. **About self-consistency.** The property of self-consistency accounts for certain model-theoretic capacities of the system, i.e. its ability of letting the user ascertain the consistency of a given information set, as well as the non-consequence of a piece of information from a given information set. A system has this property if and only if whenever a constraint of the form $\Theta_1 \vdash \emptyset$ holds on the targets and a set of states of affairs Σ_1 indicates Θ_1, a constraint $\Sigma_1 \vdash \emptyset$ holds on the representations.

4. **About the linguistic-graphic boundary.** All graphical systems project non-trivial, nomic constraints, while no linguistic systems do so. That is, every graphical system has a nomic constraint that prevents us from presenting certain information in a particular way without presenting certain other information, while no linguistic system has such constraints. The presence and absence of projection of the nomic constraints explain several major difference in efficacy between graphical and linguistic systems, such as inferential efficiency due to natural free rides, the lack of expressive flexibility due to natural content-specificity, and certain model-theoretic capacities due to natural self-consistency.

Whenever somebody makes claims in research, there are several questions to ask to evaluate the claims. Are the claims formulated clearly? Are they significant? Do the arguments make the claims plausible? Let us test our own claims on each of these criteria.

I believe that our claims pass the first test. We have spent chapter 4 to define our conceptual framework as explicitly as possible, and one main purpose for this work to formulate our claims as precisely as possible, so that they can not escape into unclarity when confronted with possible counter-examples. I believe that the contents of the claims have become transparent as the result, at least for those who take pains to digest the conceptual framework presented in that chapter.

Are the claims significant? Granted, none of the phenomena of free ride, content-specificity, and self-consistency are newly reported or discovered in this dissertation. As documented in the body of the dissertation, several people have noticed these phenomena about particular representation systems, and given some analyses on their impacts upon the overall efficacy of the relevant representation systems in question. However, the analyses that we have given to them are new, and have certain merits over the previous analyses. The primary merit is of course that our analyses make it clear that the property of providing free rides, content-specificity, and self-consistency are special instances of constraint-matching (or mismatching) between representations and targets, and that our analyses spell out exactly what matching and mismatching of constraints are responsible for each property. As for the issue of the boundary between graphical and linguistic systems, few previous analyses have made such a bold and clear claim about the boundary as our analysis does (except, perhaps, Palmer's analysis). Assuming the truth of the claim, then, it would be a significant contribution to the settlement of the long term issue, intuitively stated as: "What distinguish diagrams and pictures from descriptions?"

Have we established the truth of our claims? For the first three sets of claims, we believe that we have given sufficient illustrations to convince the reader of the accuracy of our analyses. The truth of our claims about free ride, content-specificity, and self-consistency seem thus conclusive. The issue of the boundary between graphical and linguistic representation systems is a different story. The issue has important implications in many other fields of philosophical and non-philosophical inquiries, and almost everyone seems to have his or her own view about it. Thus I can imagine that one still has some quibble about our analysis of the issue. Indeed, we probably have not exhausted and countered every example that one may think does not fit our distinction. Also, we have not shown that the proposed contrast accounts for *every* fundamental difference in efficacy between graphical and linguistic systems. However, having shown that our analysis handles certain borderline cases, avoids some apparent counter-examples, and accounts for several major differences in efficacy of graphical and linguistic systems, we may claim that our discussion gives a rather strong support for the analysis.

Thus, the four sets of claims are of their own interest and importance. They settle several issues that have been unsettled, sometimes conclusively and sometimes in plausible ways. However, these claims are not all that we have been arguing for. The main hypothesis of this dissertation is the following:

> **Constraint Hypothesis** Representations are objects in the world, and as such they obey certain structural constraints that govern their possible formation. The variance in inferential potential of different modes of representation is largely attributable to different ways in which these structural constraints on representations match with the constraints on targets of representation.

How, then, does the Constraint Hypothesis stand after all these discussions? How much support have they lent to the Hypothesis? Simply put, each of the four sets of claims cited above gives a piecemeal argument for the truth of the Hypothesis. While the properties of providing free ride, being content-specific, being self-consistent, and being graphical or linguistic account for important differences in inferential potential of different representation systems, all these properties have been shown to be due to particular ways in which the constraints on representations match or mismatch with the constraints on targets. The support is a strong one.

On the cool estimate, however, the Constraint Hypothesis still remains to be an hypothesis. While the properties of representation systems that we covered in this dissertation are ones that have significant impacts on the overall efficacy of representation systems, they are not the only ones with such important ramifications. And, in order to establish the truth of the Constraint Hypothesis, we need to show that *most* of those properties are attributable to particular matching and mismatching between constraints on representations and constraints on targets. Are we going to ever be able to show this? The answer is not known. We can, however, indicate a future direction of research in which we may add one more piece of evidence for the Hypothesis.

One particularly important area beyond the scope of this dissertation is the various kinds of efficacy associated with the *static* use of representations. It is a commonplace in the field of statistical analysis that certain modes of representation facilitate the analysis of the represented data in a way other modes of representation do not. The representation of data in one mode is "clearer," or "more illuminating" than the representations of the data in another mode. Here the analysis of the data does not generally involve explicit operations upon the representation. Typically, you just look at the representation, and retrieve some information more quickly and accurately, or gain some new insights about the represented data more easily. The examples of such differences of efficacy are abundantly given by Bertin (1973) and Tufte (1983; 1990). We can think of Bertin as referring to particular efficacy of this static sort when he characterizes the relative efficiency graphics as their ability to give a "correct and complete answer to a given question" within a shorter observation time" (p. 139).

Outside the field of statistical data analysis, Larkin and Simon (1987) point out the difference in amount of search arising from the different ways in which information are indexed in representations. Searching a representation does not necessarily involve physical operations upon the representation, so we can classify this search efficacy as the matter of a static use of representations. Also, it is often said that a picture is worth a thousand of words, which points to the ability of a graphical representation to present information in a much more compact form than a linguistic representation can. Tufte (1983) partly attributes this ability to "multi-functional elements" appearing in graphical representations. The saying also seems to refer to the so-called "behold" proofs of mathematical theorems, which somehow lets the user "see" the point of a proof just by inspecting a representation typically given in a graphical form. Nelsen (1993) collects many interesting examples of this kind.

Thus, there are lots of interesting data to be explained in the area of static use of representations. Does the Constraint Hypothesis apply to it? Namely, could we explain the data as matters of different ways in which constraints on representations match or mismatch with constraints on targets? I think so. In the absence of actual demonstration, however, I of course can not expect the reader to believe me. Whether or not the reader believe in this further point, I believe that we have said enough to motivate the reader to take a serious look at constraints on representations themselves and their interactions with constraints on their targets. It is for this purpose that we put forward the Constraint Hypothesis at this time, when the research on the efficacy of representation is rapidly growing into an independent field of research. This is the end of the dissertation, but certainly not the end of our story.

Bibliography

Barwise, J. (1989) *The Situation in Logic.* Stanford, CA: CSLI Publications.

Barwise, J. (1991) Information Links in Domain Theory. *The Mathematical Foundations of Programming Semantics Conference.* Eds. S. Brooks, M. Main, A. Melton, M. Mislove and D. Schmidt. Berlin: Springer Verlag.

Barwise, J. (1993) Constraints, Channels, and the Flow of Information. *Situation Theory and Its Applications, Volume 3.* Eds. P. Aczel, D. Israel, Y. Katagiri and S. Peters. Stanford, CA: CSLI Publications.

Barwise, J. and Etchemendy, J. (1990a) Information, Infons, and Inference. *Situation Theory and Its Applications.* Eds. R. Cooper, K. Mukai and J. Perry. Stanford, CA: CSLI Publications.

Barwise, J. and Etchemendy, J. (1990b) Visual Information and Valid Reasoning. *Visualization in Mathematics.* Ed. W. Zimmerman. Washington, DC: Mathematical Association of America.

Barwise, J. and Etchemendy, J. (1994) *Hyperproof.* Stanford, CA: CSLI Publications.

Barwise, J. and Etchemendy, J. (1995) Heterogeneous Logic. *Diagrammatic Reasoning: Cognitive and Computational Perspectives*. Eds. J. I. Glasgow, N. H. Narayanan and B. Chandrasekaran. Menlo Park, CA and Cambridge, MA: AAAI Press and MIT Press.

Barwise, J. and Hammer, E. (1995) Diagrams and the Concept of Logical System. *What Is a Logical System?*. Ed. D. Gabbay. Oxford: Oxford University Press.

Barwise, J., Gabbay, D. and Hartonas, C. (1995) On the Logic of Information Flow. *Language, Logic and Computation: The 1994 Moraga Proceedings*. Eds. D. Westerståhl and J. Seligman. Stanford, CA: CSLI Publications.

Barwise, J. and Perry, J. (1980) The Situation Underground. *Stanford, CA Working Papers in Semantics: Volume I*. Eds. J. Barwise and I. A. Sag.

Barwise, J. and Perry, J. (1983) *Situations and Attitudes*. Cambridge, MA: The MIT Press.

Barwise, J. and Seligman, J. (1993) The Rights and Wrongs of Natural Regularity. *Philosophical Perspectives*. Ed. J. Tomberlin. Atascadero, CA: Ridgeview.

Barwise, J. and Seligman, J. (1996) Information Flow in Distributed Systems. Unpublished.

Barwise, J. and Shimojima, A. (1995) Surrogate Reasoning. *Cognitive Studies: Bulletin of Japanese Cognitive Science Society* **2** (4): pp. 7–27.

Berkeley, G. (1710) *A Treatise Concerning Principles of Human Knowledge*. Indianapolis, IN: The Bobbs-Merrill Company, Inc., 1957.

Bertin, J. (1973) *Semiology of Graphics: Diagrams, Networks, Maps.* Translated by William J. Berg. Madison, WI: The University of Wisconsin Press, 1983.

Braine, M. D. S. (1978) On the Relation between the Natural Logic of Reasoning and Standard Logic. *Psychological Review* **85**: 1–21.

Dennett, D. C. (1969) *Content and Consciousness.* London: Routledge and Kegan Paul.

Devlin, K. (1991) *Logic and Information.* Cambridge, UK: Cambridge University Press.

Dretske, F. I. (1981) *Knowledge and the Flow of Information.* Cambridge, MA: MIT Press.

Funt, B. V. (1980) Problem-Solving with Diagrammatic Representations. *Diagrammatic Reasoning: Cognitive and Computational Perspectives.* Eds. J. Glasgow, N. H. Narayanan and B. Chanrasekaran. Menlo Park, CA and Cambridge, MA: AAAI Press and MIT Press, 1995.

Gelernter, H. (1959) Realization of a Geometry-Theorem Proving Machine. *Computers and Thought.* Eds. E. A. Feigenbaum and J. Feldman. Mew York: McGraw Hill.

Goodman, N. (1968) *Languages of Art: an Approach to a Theory of Symbols.* Indianapolis, IN: the Bobbs-Merrill Company, INC.

Grice, H. P. (1975) Logic and Conversation. *Speech Acts.* Eds. P. Cole and J. Morgan. New York: Academic Press.

Hammer, E. (1995) Diagrams, Logic and Representation. Ph.D. Thesis. Indiana University.

Harel, D. (1988) On Visual Formalisms. Technical Report 87-126. Department of Copmputer Science, Carnegie Mellon University.

Hartley, R. V. L. (1928) Transmission of Information. *Bell System Technical Journal* **7**: 535–563.

Hohauser, S. (1982) *Architectural and Interior Models*. Second Edition. New York: Van Nostrand Reinhold Company.

Hume, D. (1739) A Treatise of Human Nature. *The Philosophical Works of David Hume*. Boston: Little, Brown and Company, 1854.

Johnson-Laird, P. N. (1983) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*. Cambridge, MA: Harvard University Press.

Johnson-Laird, P. N. and Byrne, R. M. J. (1991) *Deduction*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Kosslyn, S. M. and Pomerantz, J. R. (1977) Imagery, Propositions, and the Form of Internal Representations. *Cognitive Psychology* **9**: 52-76.

Larkin, J. H. and Simon, H. A. (1987) Why a Diagram Is (Sometimes) Worth Ten Thousand Words. *Diagrammatic Reasoning: Cognitive and Computational Perspectives*. Eds. J. Glasgow, N. H. Narayanan and B. Chanrasekaran. Menlo Park, CA and Cambridge, MA: AAAI Press and MIT Press, 1995.

Levesque, H. J. (1988) Logic and the Complexity of Reasoning. *Journal of Philosophical Logic* **17**: 355–389.

Lewis, D. K. (1979) Scorekeeping in a Language Game. *Semantics from Different Points of View.* Eds. R. Bäuerle, U. Egli and A. v. Stechow. Berlin: Springer Verlag.

Lindsay, R. K. (1988) Images and Inference. *Diagrammatic Reasoning: Cognitive and Computational Perspectives.* Eds. J. I. Glasgow, N. H. Narayanan and B. Chandrasekaran. Menlo Park, CA and Cambridge, MA: AAAI Press and MIT Press, 1995.

Luengo, I. (1995) Diagrams in Geometry. Ph.D. Thesis. Indiana University.

Nelsen, R. B. (1992) *Proofs without Words: Exercises in Visual Thinking.* Washington, DC: The Mathematical Association for America.

Palmer, S. E. (1978) Fundamental Aspects of Cognitive Representation. *Cognition and Categorization.* Eds. E. Rosch and B. B. Lloyd. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Pylyshyn, Z. W. (1973) What the Mind's Eye Tells the Mind's Brain: A Critique of Mental Imagery. *Psychological Bulletin* **80**: 1–24.

Rips, L. (1994) *The Psychology of Proof.* Cambridge, MA: The MIT Press.

Shannon, C. (1948) *The Mathematical Theory of Communication.* Urbana-Champaign, IL: The University of Illinois Press, 1949.

Shepard, R. N. and Cooper, L. A. (1982) *Mental Images and Their Transformations*. Cambridge, MA: MIT Press.

Shimojima, A. (1996) Operational Constraints in Diagrammatic Reasoning. *Logical Reasoning with Diagrams*. Eds. J. Barwise and G. Allwein. Oxford: Oxford University Press.

Shin, S.-J. (1991a) A Situation-Theoretic Account of Valid Reasoning with Venn Diagrams. *Situation Theory and Its Applications, Volume 2*. Eds. J. Barwise, J. M. Gawron, G. Plotkin and S. Tutiya. Stanford, CA: CSLI Publications.

Shin, S.-J. (1991b) Valid Reasoning and Visual Representation. Ph.D. Thesis. Stanford, CA University.

Sloman, A. (1971) Interactions between Philosophy and AI: the Role of Intuition and Non-Logical Reasoning in Intelligence. *Artificial Intelligence* **2**: 209–225.

Sloman, A. (1995) Musings on the Roles of Logical and Nonlogocal Representations in Intelligence. *Diagrammatic Reasoning: Cognitive and Computational Perspectives*. Eds. J. Glasgow, N. H. Narayanan and B. Chanrasekaran. Menlo Park, CA and Cambridge, MA: AAAI Press and MIT Press.

Stenning, K. and Inder, R. (1995) Applying Semantic Concepts to Analyzing Media and Modalities. *Diagrammatic Reasoning: Cognitive and Computational Perspectives*. Eds. J. Glasgow, N. H. Narayanan and B. Chandrasekaran. Menlo Park, CA and Cambridge, MA: AAAI Press and MIT Press.

Stenning, K. and Oberlander, J. (1995) A Cognitive Theory of Graphical and Linguistic Reasoning: Logic and Implementation. *Cognitive Science* **19**: 97–140.

Tufte, E. R. (1983) *The Visual Display of Quantitative Information.* Cheshire, CN: Graphics Press.

Tufte, E. R. (1990) *Envisioning Information.* Cheshire, CN: Graphics Press.

Westerståhl, D. (1990) Parametric Types and Propositions in First-Order Situation Theory. *Situation Theory and Its Applications.* Eds. R. Cooper, K. Mukai and J. Perry. Stanford, CA: CSLI Publications.

Wiener, N. (1948) *Cybernetics.* Second Edition. New York: The MIT Press, 1961.

Wittgenstein, L. (1921) *Tractatus Logico-Philosophicus.* Second Edition. Translated by D. F. Pears and B. F. McGuinness. London: Routledge and Kegan Paul, 1971.

Curriculum Vitae

Atsushi Shimojima was born in Akashi (Japan) in 1962. He obtained his B¿A. (philosophy) in 1985 and his M.A. (philosophy) in 1988 from Doshisha University. With the five-year Fulbright grant for graduate students, he came to the Philosophy Department of Indiana University, where he was awarded the Graduate Academic Excellence Award in 1992 and the Graduate Ewing Essay Award in 1994. From 1992 to 1995, he taught introductory philosophy and logic as an associate instructor, and was a research assistant to the Visual Inference Laboratory from 1995 to 1996. After his graduation, he will continue his research at ATR Media Integration and Communications Research Laboratories in Japan.