1-1-1981

Approximate reasoning, logics for self-reference, and the use of nonclassical logics in systems modeling

Daniel Guy Schwartz
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds

Let us know how access to this document benefits you.

Recommended Citation

APPROXIMATE REASONING, LOGICS FOR SELF-REFERENCE,

AND THE USE OF NONCLASSICAL LOGICS

IN SYSTEMS MODELING

by

DANIEL GUY SCHWARTZ

A dissertation submitted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY
in
SYSTEMS SCIENCE

Portland State University

1981

TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the dissertation of
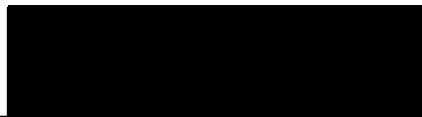Daniel Guy Schwartz presented June 2, 1981.
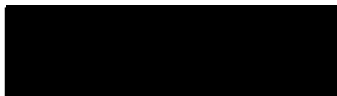
George GOLendaris, Chairman

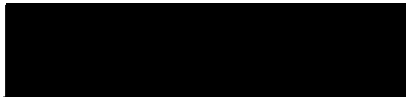Harold A. Linstone

Robert L. Stanley

Lotfi A. Zadeh

Martin Zwick

APPROVED:

Stanley E. Rauch, Dean of Graduate Studies and Research

AN ABSTRACT OF THE DISSERTATION OF Daniel Guy Schwartz for the

Doctor of Philosophy in Systems Science presented June 2, 1981.

Title: Approximate Reasoning, Logics for Self-Reference, and the

Use of Nonclassical Logics in Systems Modeling.

APPROVED BY MEMBERS OF THE DISSERTATION COMMITTEE:

George G. Lendaris, Chairman

Harold A. Linstone

Robert L. Stanley

Lotfi A. Zadeh

Martin Zwick

This work advances the use of nonclassical logics for developing qualitative models of real-world systems. Abstract mathematics is "qualitative" inasmuch as it relegates numerical considerations to the background and focuses explicitly on topological, algebraic, logical, or

other types of conceptual forms. Mathematical logic, the present topic, serves to explicate alternative modes of reasoning for use in general research design and in model construction.

The central thesis is that the theory of formal logical systems, and particularly, of logical systems based on nonclassical modes of reasoning, offers important new techniques for developing qualitative models of real-world systems. This thesis is supported in three major parts.

Part I develops a semantically complete axiomatization of L. A. Zadeh's theory of approximate reasoning. This mode of reasoning is based on the conception of a "fuzzy set," by which means it yields a realistic representation of the "vagueness" ordinarily inherent in natural languages, such as English.

All axiomatizations of this mode of reasoning to date have been deficient in that their linguistic structures are adequate for expressing only the simplest fuzzy linguistic ideas. The axiomatization developed herein goes beyond these limitations in a two-leveled formal system, which, at the inner level, is a multivalent logic that accommodates fuzzy assertions, and at the outer level, is a bivalent formalization of segments of the metalanguage. This system is adequate for expressing most of the basic fuzzy linguistic ideas, including: linguistic terms, hedges, and connectives; semantic equivalence and entailment; possibilistic reasoning; and linguistic truth.

The final chapter of Part I applies the theory of approximate reasoning to a class of structural models for use in forecasting. The result is a direct mathematical link between the imprecision in a model

and the uncertainty which that imprecision contributes to the model's forecasted events.

Part II studies the systems of logical "form" which have been developed by G. Spencer-Brown and F. J. Varela. Spencer-Brown's "laws of form" is here shown to be essentially isomorphic with the axiomatized propositional calculus, and Varela's "calculus for self-reference" is shown to be isomorphically translatable into a system which axiomatizes a three-valued logic developed by S. C. Kleene. No semantically complete axiomatization of Kleene's logic has heretofore been known.

Following on Kleene's original interpretation of his logic in the theory of partial recursion, this leads to a proof that Varela's concept of logical "autonomy" is exactly isomorphic with the notion of a "totally undecidable" partial recursive set. In turn, this suggests using Kleene-Varela type systems as formal tools for representing "mechanically unknowable" or empirically unverifiable system properties.

Part III is an essay on the theoretical basis and methodological framework for implementing nonstandard logics in the modeling exercise. The evolution of mathematical logic is considered from the standpoint of its providing the opportunity to "select" alternative modes of reasoning. These general theoretical considerations serve to motivate the methodological ones, which begin by addressing the discussions of P. Suppes and M. Bunge regarding the role of formal systems in providing "the semantics of science." Bunge's work extends that of Suppes and is herein extended in turn to a study of the manner in which formal systems (both classical and nonclassical) can be implemented for mediation between the observer and the observed, i.e., for modeling. Whether real-world systems in fact obey the laws of one logic versus another must remain moot, but

models based on alternative modes of reasoning do satisfy Bunge's criteria for empirical testability, and therefore do provide viable systems perspectives and methods of research.

This work is dedicated with love and appreciation to

GURU MAHARAJ JI,

without whose grace I could not possibly have gained

the presence to see it through.

But if a man meditate on Me and Me alone, and will
worship Me always and everywhere, I will take upon Myself
the fulfillment of his aspiration, and I will safeguard
whatsoever he shall attain.

Lord Krishna

From The Baghavad Gita (Swami 1977, p. 90)

PREFACE

The underlying theme of this work originated approximately ten years ago during my master's degree program in mathematics at Simon Fraser University, British Columbia, Canada. From a course in mathematical logic came the idea that, if one could write all of the laws of physics in a semantically complete formal logical system, then the semantic models of that system would represent all and only those possible universes in which the given laws hold true.

I mentioned this idea once to Professor R. Dass, an applied mathematician working in relativity theory. He replied that any model of the world is scientifically acceptable as long as it is internally consistent, agrees with the known data, and is explainable to other mathematicians. This was positive encouragement that the idea was worthy of development, but the opportunity for carrying it further did not arise at that time.

During these same years came my first encounter with G. Spencer-Brown's *Laws* *of* *form*, which I read with avid interest, but then put aside as being little more than an anomaly of logical studies and as having only peripheral significance. My main interests turned toward multivalued logics, modal logics and particularly, the subject of paradox. These studies bore out a speculation that, if a meaningful rendering of paradox could be developed, then Dr. Dass's requirement of consistency could be dispensed with, or at least, be replaced with a more general one.

These ideas remained foremost with me as a beginning student in systems science at Portland State University, 1975. There the quest began anew for a possible line of development. The first concrete idea arose serendipitously, in the office of my faculty advisor, Professor George Lendaris. While waiting for him to finish a telephone conversation, I took an issue of The International Journal of General Systems from his bookshelf and began turning pages.

That issue contained Francisco Varela's paper, "A calculus for self-reference," in which I immediately recognized exactly the same notations as used in the book by Spencer-Brown. Curiosity quickly manifested in activity, and a thorough review of both Spencer-Brown's book and Varela's paper led to the next most natural question: What is the relationship of these systems to contemporary systems which use the standard linguistic forms?

An exploration of this question was carried out under the tutelage of Professors Robert Rempfer and Robert Stanley of the Portland State University Department of Mathematics. This resulted in a paper which was ultimately accepted for publication (Schwartz 1981a) in the same journal as carried Varela's paper and which now appears in this dissertation as the essential content of Part II. Francisco Varela was one of the reviewers of the published version, and his remarks were most helpful in producing what appears here as Chapter 9.

The work in Part I of this dissertation arose through a continued association with George Lendaris, who introduced me to Professor L. A. Zadeh's theory of fuzzy sets and approximate reasoning. By the time the work on Varela's calculus was complete, it had become clear that this calculus could be taken also as a formalization of a small collection of

Zadeh's basic ideas. This suggested expanding the language of the "isomorphic copy," of Varela's system into a richer language, adequate to express a more extensive portion of Zadeh's theory. This work began as a project for a systems research workshop, with both George Lendaris and Professor Martin Zwick, and culminated in a semantically complete axiomatization of the essential aspects of approximate reasoning. The proof of semantic completeness given here is patterned after a similar proof given for a somewhat simpler class of "free-variable theories" in my master's thesis (Schwartz 1973), and its success is due largely to some special algebraic techniques developed by myself and my master's thesis advisor, Professor S. K. Thomason.

I was most fortunate in having the opportunity to present these ideas to Lotfi Zadeh himself, in an informal seminar at Berkeley, California, 1980. It gave me a special inspiration when he said that he was impressed with my work, and I am grateful for his comments which led to the correctin of several oversights in an earlier version. For the work in this part, Professor Stanley must also be credited for many helpful suggestions on improving the written presentation.

The concluding chapter of Part I came out of some studies with Professor H. A. Linstone, in technology assessment. Dr. Linstone extended a challenge to show that fuzzy sets had a practical viability, and the reponse that has manifested here is an application of approximate reasoning to the problem of forecasting under uncertainty. This same chapter comprises a talk given at the 1981 meeting of the Society for General Systems Research, and a shorter version appears in that meeting's proceedings (Schwartz 1981b).

The work in Part III arose in an attempt to formulate earlier philosophical ideas and was intended originally as a short introduction to a dissertation consisting mainly of Parts I and II. When Martin Zwick introduced me to the works of Mario Bunge, however, the task quickly developed into a research project of its own. The outcome has been to provide the entire dissertation with a solid methodological base.

Here I would like to formally express my sincerest appreciation to all of these fine gentlemen for the part they have played in my development and in the production of this work. Their assistance and constant encouragement have obviously been essential to my success. And this applies especially to my advisor, George Lendaris, who helped see me through more difficulties and setbacks than I could possibly document here.

Also I wish to mention other mentors who in various ways have been invaluable sources of encouragement and inspiration. These are Professors Devendra Sahal, Magaroh Maruyama, and Ervin Laszlo. The latter in particular should be thanked for helping with my first professional publication (Schwartz 1977) and thereby encouraging me that my ideas could be of interest. None of these, of course, should be held accountable for any errors or inaccuracies that this dissertation might still contain.

To Mary and Greg Valdez of the Word Rite Center, Tualatin, Oregon must go credit for their very elegant production of the typed draft. This was produced on their IBM OS/6 word processor, with a math-symbol attachment, and has gone through several editings and revisions.

Last, I must also express my recognition of Guru Maharaj Ji, to whom this work is dedicated. I was initiated into Guru Maharaj Ji's knowledge and meditation in Portland, Oregon on November 4, 1974, and it was in fact the experience of his knowledge which inspired me to return to the university. Through my efforts to practice his knowledge as he prescribes, I gained the requisite clarity to carefully distinguish between my studies and the ultimate purpose of my life. This has given me the very precious opportunity to engage in my mundane pursuits with a true sense of detachment, and oftentimes, an ineffable joy. Each person finds a purpose in the objects of his devotion, but it is only in devotion to the Truth that the true purpose is revealed.

TABLE OF CONTENTS

CHAPTER

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Each stage of society's advance bears forth its own requirements
in the art of managing complexity. Systems science represents a contem-
porary answer to the complexity of the contemporary world. It overlays
the multitude of academic disciplines and technical specializations with
a broad-based interdisciplinary (or transdisciplinary) approach to
problem solving in a multifaceted environment (the systems approach).
It studies systems as "whole entities" and strives to lay down principles
which describe their structures and dynamic behaviors more or less
independently of their specific constituting parts (general systems
theory). And it provides a technical language for cross-disciplinary
communication among scientists, artists, engineers, economists, sociolo-
gists, etc., for pooling talents as required by a given problem setting.

Mathematical systems theory represents that part of the systems
research movement which provides mathematically succinct analytical
tools for general systems analysis. Systems may be stable or adaptable;
fragile or resilient; they might seek goals, oscillate, wander, or
evolve; they may be hierarchical, or self-referential, or autopoietic
(self-perpetuating); they may be deterministic or stochastic; our percep-
tions of them may be complete and precise, or only partial and fuzzy;
and so on. This newly emergent plethora of "systemic" concepts presents
the need for new mathematical techniques appropriate for providing them
with rigorous formulation. As with traditional science, so also with

systems science, mathematics remains the handmaiden of conceptual precision.

Qualitative mathematics may be defined as mathematics used to represent nonmeasurable system properties. Ill-defined problems such as arise in the so-called "soft" sciences like economics, policy research, and futures studies, are not always amenable to the traditional mathematics--e.g., differential equations and statistics--which rests on objective numerical measurement of system parameters. Important parameters of many systems are often difficult, or impossible, to provide with accurate representation--e.g., corporate morale--yet they must be accounted for in any reasonable systems analysis. Recent developments in mathematical systems theory now show that several branches of abstract mathematics offer effective methods for meeting this requirement. The usefulness of abstract mathematical theories for providing a qualitative methodology stems from their ability to leave numerical considerations in the background and to focus explicitly on geometrical, algebraic, logical, or otherwise conceptual forms. As illustrations, consider the following:

1) Topology studies geometrical shapes with respect to their properties that remain after stretching and bending. A contemporary example of quality representation coming out of this area is Thom's theory of "structural stability and morphogenesis" (catastrophe theory) which applies differential topology to the study of sudden, discrete transformations induced by the gradual application of a continuous force (Thom 1975). Another is "Q-analysis," Atkins topological "language of structure," which defines a system in terms of interdependencies existing among the system parameters (cf. Cavallo 1979).

2) Abstract algebra, topology's mathematical counterpart, studies the general properties of algebraic functions and relations. As examples in this area we have Petri's "general net theory," which combines graphical and computational tools for modeling systems in terms of informational flows (also cf. Cavallo 1979), Sahal's use of the algebraic theory of dimensions to define a concept of "self-similitude" for evolving systems (Sahal 1976 and 1978), and an application of category theory to the study of autonomy and self-organization in living systems (Varela and Goguen 1978). Another instance is the class of "structural models," which combines both algebraic and topological ideas. In general, a structural model is a collection of variables together with their interrelations (Lendaris 1980), and it is represented geometrically as a digraph in which the interrelations are exhibited in a visual display.

3) Mathematical logic explicates various aspects of human reasoning within a precise formalism. The leading example of quality representation in this area is L. A. Zadeh's theory of "fuzzy sets and approximate reasoning." This theory has an advantage over many forms of qualitative mathematics in that it formally represents "quality" as a property of the system observer's use of language and his manner of reasoning with linguistic ideas, rather than of the system being observed. Hence, it provides a convenient tool for modeling and manipulating linguistic evaluations. Moreover, approximate reasoning can capture some of the same properties as other modeling methods. For example, Zwick, Schwartz, and Lendaris (1978) show that Thom's "cusp" and "butterfly" catastrophes can be represented as properties of fuzzy sets; and where this dissertation Part II discusses the formal representation of paradoxical system properties, Zadeh (1979) shows that analogous results are obtainable in

a "possibilistic logic" based on fuzzy sets. Such a methodology, more-
over, should not in any way be confused as one which is itself inexact
or imprecise. While it is possible to develop a mathematics of imprecise
ideas, precision at some explicit level is always inherent in the very
nature of all mathematical methods. An "imprecise mathematics" would be
a contradiction of terms.

The work undertaken in this dissertation falls within the category
of mathematical logic. Specifically, it addresses the application of
nonclassical logics in systems research. The central thesis may be
encapsulated as follows: that the theory of formal logical systems, and
particularly, of logical systems based on nonclassical modes of reasoning,
offers important new techniques for building qualitative models of
real-world systems. The present work is here claimed to advance this
thesis a further step beyond its current stage. Moreover, the work is
in this respect largely foundational; other than the structural models
discussed in Chapter 6, no new models are produced. The primary effort
has mainly been to develop the basis for using nonclassical logics in
systems analysis, and thereby, to indicate more clearly how models based
on such logics may be designed. This is accomplished in three major
parts.

Part I develops a semantically complete formal axiomatization of
the basic ideas appearing in Zadeh's theory of approximate reasoning.
In particular, the development includes both a multivalent logic, which
takes as its truth values the points in the interval $[0,1]$, and a "fuzzy
logic," which takes for its truth values certain fuzzy subsets of $[0,1]$.
The primary motivation for this part of the work is the same as for the
axiomatization of other mathematical theories. It lays down in concise

terms exactly those ideas which the given theory entails, and it provides a means by which one can then examine that theory for its consistency, completeness, decidability, and other properties as a system of mathematical deduction. Semantic completeness, particularly, is in this respect essential, for establishing that the resulting formalization indeed represents the relevant aspects of the informal theory. Part I furthermore develops a new class of structural models for use in systems forecasting. In this case, approximate reasoning is used to provide a formal link between the "imprecision" of a structural model as a representation of reality and the "uncertainty" of the model's forecasted results.

Part II studies the logical systems developed by Spencer-Brown (1969) and Varela (1975), which provide special symbolisms and rules of operation for the logical "forms" that underly both classical and self-referential reasoning. The work in this dissertation establishes the exact relation of these systems to formal logical systems which employ the standard linguistic notations. Based on a connection with recursive function theory, this in turn yields insights into the potential use of self-referential reasoning for modeling "mechanically unknowable" or empirically unverifiable real-world system properties, such as the notion of a "perfect system" (Wienberg 1975), or certain formulations of "time" (Gale 1968).

Part III explores the theoretical foundations and methodological framework for applying nonclassical logics in mathematical modeling. This includes a consideration of how the theory of formal logical systems is related to the "semantics of science" in the sense discussed by Bunge (1973). The principal conclusion is that, since nonclassical logics are

semantically "meaningful," they lead to empirically testable models of real-world systems and therefore yield methodologically correct approaches to model specification.

The philosophical remarks on the use of formal systems (Part III) was initially what inspired an interest in the calculus for self-reference (Part II), and the results of Part II provided the technical insights which led to the axiomatization of approximate reasoning (Part I). Thus the above arrangement of the three parts is somewhat the reverse of the chronology in which the essential ideas evolved. The chosen order of presentation is appropriate, however, since the later developments have ultimately become the core of the work.

Each part of the work is essentially self-contained in its concern with a specialized body of ideas, but all three parts are tied by a common purpose: the use of formal logics in improving our understanding of the world. Here follows a brief exposition of the context in which these studies arise and a summary of the main results.

§1.1. Overview Of Part I

During the last seventy or eighty years, the study of natural languages has developed more or less independently of mathematical logic, even though these studies have enjoyed common philosophical underpinnings. Linguistics has of course emulated mathematics to an extent, in striving for rigorous, analytically precise models of the verbal modes of human communication. But mathematics was not correspondingly inspired to expand the expressive capabilities of formal languages beyond the requirements of a two-valued "true-false" level of discourse. This was due primarily to the interest in analyzing the foundations of

mathematical thought, for which the two-valued logic was generally
adequate.

It was in response to these shortcomings of conventional mathema-
tics that approximately fifteen years ago L. A. Zadeh proposed the idea
of "fuzzy sets." The seminal paper (Zadeh 1965) described how impreci-
sion and value judgments in the use of natural languages can be modeled
in a rigorously defined algorithmic semantics through a simple generali-
zation of the classical definition of a set of elements. Whereas a
classical set S may be represented by a membership function μ_S by letting
$\mu_S(x)=1$ if x is "in" the set S, and letting $\mu_S(x)=0$ if x is "not in" S,
the idea of a "fuzzy set" is obtained by allowing μ_S to range over the
entire interval from 0 to 1. Thus one obtains the idea of a "degree" or
"grade" of membership in S. As an example: If T is a fuzzy set of
"tall persons," then a person who is six feet in height might have .9 as
a grade of membership in T, while a person who is five feet six inches
in height might have a grade of membership of only .4 in T. The assign-
ment of such grades of membership is by and large the product of a
subjective interpretation in the context of a given "universe of
discourse"--e.g., a building which is only six feet in height in the
fuzzy set of "tall buildings" would have a very low grade of membership.
Usually the grades of membership are given in terms of an explicitly
defined function on the interval [0,1].

Thus formulated, fuzzy sets lead naturally to an interpretation of
"linguistic hedges"--i.e., modifiers like "very," "more or less," "almost,"
and "quite"--as well as the Boolean connectives "not," "or," and "and,"
as operations on membership functions of fuzzy sets. The result is an
intuitively plausible algorithmic interpretation for simple propositions

like "John is <u>tall</u>," "May has <u>dark</u> hair," and "May is <u>much</u> <u>younger</u> than John," as well as the somewhat more complex propositions such as "<u>Most</u> Frenchmen are not <u>blond</u>," "It is <u>quite</u> <u>possible</u> that <u>many</u> <u>wealthy</u> Americans have <u>high</u> blood pressure," and "It is <u>probably</u> <u>quite</u> <u>true</u> that <u>most</u> X's are <u>much</u> <u>larger</u> than <u>most</u> Y's" (from Zadeh 1978a). Further, by means of the fuzzy interpretations of the Boolean connectives, one obtains a well-defined and intuitively plausible logic for reasoning with fuzzy linguistic assertions.

After several preliminary pioneering papers on this idea, there began a rapidly accelerating acceptance of fuzzy sets into a multitude of technical disciplines. After only one decade of research, Gaines and Kohout (1977) summarized the literature on fuzzy sets in a bibliography which listed some 1,150 entries with approximately 750 being specifically concerned with the theory and applications of this one idea. Now, a bibliography by Kandell and Yager (1979) pushes the total to 1,800 entries.

Applications of fuzzy sets have been investigated along such topics as pattern recognition (Zadeh's initial concern), cybernetic controls, automated production systems, quality control, multiple-criteria decision making, fuzzy classification schemes, data base management, dynamic systems modeling, and natural language processing. In general, the theory of fuzzy sets offers many opportunities for improving both the basic machine capability and the man-machine interface.

Papers by Zadeh have focused primarily on the theoretical developments. The idea of linguistic hedges was explored (Zadeh 1972). The idea of a "fuzzy logic," which takes as its truth values the fuzzy subsets of [0,1]--rather than the points in [0,1] as is characteristic

of most multivalent logics--began to take shape in 1975 (Zadeh 1975a and 1975b) and was later studied as a "fuzzification" of the Lukasiewicz infinitary logic $L_{\aleph_1}$ by Bellman and Zadeh (1977). Zadeh's more recent works (1978a and 1978b) have developed the theory of "linguistic possibility," or "possibilistic reasoning," which provides translation rules for transforming natural language expressions into a computable formalism (PRUF).

The philosophical aspects of "fuzzy reasoning" have been explored by Gaines (1976). Related books and edited collections of essays deserving of mention are: Kaufman (1975), Moisil (1975), Zadeh and Fu (1975), Gupta, Saridis, and Gaines (1977), Dubois and Prade (1978a), Kandel and Lee (1979), Gupta, Ragade, and Yager (1979), and Negoita (1980).

Now that the theory of approximate reasoning has crystalized into an established body of knowledge, an obvious next step is to provide this theory with a complete formal axiomatization. Giles (1976) took a preliminary step in this direction by showing that the axioms for Lukaseiwicz's infinitary logic $L_{\aleph_1}$ also serve as axioms for a multivalent interpretation of the fuzzy Boolean connectives. The language of $L_{\aleph_1}$ is essentially identical with the language of the classical propositional calculus, however, and therefore does not have the capability for expressing such concepts as "semantic equivalence," "semantic entailment," "linguistic variable," "hedge," "linguistic truth," and so on, which are essential to the theory of approximate reasoning. Axiomatizing these more complex ideas evidently requires a formalism with substantially richer linguistic capabilities.

The work in Part I of this dissertation makes a further step in this direction. A class of formal logical systems is defined wherein

each system represents one or more linguistic variables over a single universe of discourse. These systems are unique in their use of a linguistic-axiomatic structure which operates simultaneously at two distinct but interrelated semantic levels. An "inner level" accommodates fuzzy linguistic assertions such as "John is young and very tall" in a multivalent interpretation, and an "outer level" accommodates composite expressions like "'John is young and very tall' is semantically equivalent with 'John is not old and not at all short'" in a bivalent interpretation. In effect the outer level is a formalization of part of the metalanguage that is used for discussing linguistic assertions, and bivalency stems from the tacit assumption that at the metalevel two linguistic assertions either are or are not semantically equivalent.

Once defined, the multivalent axiomatization of approximate reasoning is easily modified to become a semantically complete axiomatization of a "fuzzy logic" per se, which uses "fuzzy numbers" in [0,1] as truth values. There remain, however, many fuzzy-set theoretic ideas which are not represented in these formalisms. In particular should be mentioned the fuzzy quantifiers like "most," "almost all," and "few." Further, while the present axiomatization captures the abstract idea of a linguistic hedge, it is still not sufficiently rich to provide hedges with explicit arithmetical definitions. Nevertheless, given the results of Part I, it is evident that many of these limitations can be dealt with through further extensions of these two-leveled systems.

The principal accomplishments represented by this part of the dissertation may thus be summarized as follows: (i) a rigorous formalization of the general concept of "linguistic variable" in the context of semantic equivalence between linguistic assertions, and (ii) a precise

axiomatization of how these linguistic ideas are interrelated under both multivalent and fuzzy-logical interpretations. Part I contains five chapters, as described below.

Chapter 2 provides the basic terminology. This includes all the relevant concepts from the theory of formal logical systems, together with all needed ideas from the theory of fuzzy sets and approximate reasoning. In its treatment of formal systems, this chapter also serves the purposes of dissertation parts II and III.

Chapter 3 gives an explicit definition of "linguistic theory" (or "theory of a linguistic variable"), which incorporates the abovementioned two-leveled linguistics into the form of a rigorously defined class of formal logical systems. Accompanying this is a definition of the semantics for linguistic theories, amounting to a precise set-theoretic description of the concept of linguistic variable.

The main result is established, in full detail, in Chapter 4. The given class of linguistic theories is complete with respect to the given semantics. This result is what ensures that the axiomatization is indeed adequate to capture the essential characteristics of fuzzy linguistic variables and the multivalent theory of approximate reasoning. In other words, semantic completeness ensures that all and only the semantically "true" fuzzy linguistic propositions are formally derivable from the axioms in question.

Miscellaneous extensions and generalizations of the main results are considered in Chapter 5. These include: linguistic theories of multiple and interrelated linguistic variables; linguistic theories having multiple universes of discourse; the full fuzzy logic, as described above; a formalization of possibilistic reasoning; and several

alternative modes of fuzzy-logical inference. The most significant result of this chapter is the extension to an axiomatization of fuzzy logic, and the proof that this axiomatization also is semantically complete.

Chapter 6 concludes Part I with an application. The structural modeling package, SPIN (McLean, et. al. 1976), is enhanced by incorporating a fuzzy-reasoning approach to "certainty accounting." The result is an explicit formal connection between the certainty in a given model as a correct description of reality and the certainty of that model's forecasted results. Briefly stated, the less "fuzzy" the model, the more certain are the results, and the more "fuzzy" the model, the less certain are the results. Algorithms are developed for calculating "degrees of certainty" of forecasts as degrees of truth of fuzzy-logical propositions, and sample computations are provided for a simple model with five parameters.

Here it should be mentioned that, while such models as discussed above might be construed as an approach to the method of "future discounting" discussed by Linstone (1973 and 1979), that author has pointed out to me in private communication that "uncertainty" is only one of several motivations for discounting, which in particular include matters of economics and of biological survival. Further, where Bellman (1977) suggests the "tremendous opportunities" offered by fuzzy systems theory, Linstone and Simmonds (1977 pp. 133-134) overlay this with the qualification that one must also be watchful that adopting such tools do not constrain one's view in places where new modeling techniques need to be invented. Similar remarks apply, however, to all modeling techniques, be they fuzzy or not; it is nowdays commonly understood (although

perhaps not commonly put into practice) that one should only use modeling tools where they apply. Even given the above qualification, therefore, this author sides with Bellman, but only by putting more stress on the "opportunities." It is simply a matter of time before fuzzy systems prove their worth in forecasting as well as the other fields where they are now being applied.

§1.2. Overview Of Part II

Historically, paradox has been the bane of the mathematical disciplines. The reason? In any classical, Aristotelean logic, paradox inevitably leads to a completely trivial system in which all propositions become formally provable. Thus, if Aristotelean logic is to be vindicated, paradox must be eliminated.

A case in point is Russell's discovery of a paradox in Frege's theory of sets (Russell 1902). Because both Frege and Russell aimed to promote classical logic as the irrevocable basis of all mathematical reasoning, including that employed in set theory, a resolution of the Russell paradox was required within this frame. Hence arose the theory of types (Russell 1908).

Yet paradox itself is a naturally occurring form of human intellection. People oftentimes argue themselves in circles quite cheerfully-- and usually unwittingly--without showing the slightest signs of discomfort or remorse. One day a person might be a convinced optimist, and the next day be the most cynical pessimist, or he might decide once to pursue a certain objective, and then suddenly begin moving in a completely different direction. It is nowdays a cliché that human behavior is largely "irrational." How one should characterize "rational" behavior

is, of course, a subject of debate. Yet it is in any case clear that, if there is any "logic" which would be appropriate as a basis for modeling such irrationality, it is evidently not the internally consistent Aristotelean logic of the exclusive "true or false."

Bateson gave substance to this view in a study of schizophrenia (Bateson, et. al. 1956). Through internalizing a paradoxical "double bind," a person can entertain two contradictory beliefs simultaneously, with either one of them becoming available as a behavioral prescriptor. Based on this discovery, Bateson and his coworkers thus proposed that the capacity to entertain paradoxes is a factor in many psychological disorders.

But paradoxical reasoning per se need not be characterized as exclusively pathological. Many philosophies, eastern and western, ancient and modern, openly acknowledge paradoxical assumptions (cf. Smith 1976). For example, Hinduism conceptualizes Brahman as being "neither this nor that" but a fundamental reality residing beyond the realm of intellectual comprehension. And Sartre's existentialist philosophy is founded on a belief that freedom of the will is simply a "fact," but that in its "facticity," it is "absurd"; we have freedom of choice, but we are bound to choose, lest we lapse into "nothingness" (Sartre 1956). Hence, Sartre's freedom is beyond reason, but is nevertheless real; i.e., it is illogically "true." In such cases, paradox (or absurdity) plays the special role of defining limits for the human rationality, by pointing to--or at least suggesting--the existence of realities which are completely beyond the human abilities to comprehend.

Given this "naturalness" of paradoxical reasoning, therefore, it might be expected that even the mathematical sciences should begin to

introduce paradoxical ideas into their models of reality. Mathematically, paradox is distinguishable from an ordinary contradiction as being an assertion to the effect that a certain proposition is equivalent with its negation (P≡¬P), whereas a contradiction asserts the simultaneous truth of a proposition and its negation (P&¬P). While no scientific theories to date have openly promoted paradoxical (or contradictory) ideas, there are various instances which impinge on paradox in their use of semantic self-reference or other forms of logical circularity. Recent developments of this genre include the following: the idea of mutual causality (Maruyama 1963); the interdependency of the observer and the observed (von Foerster 1973 and Howe and von Foerster 1975); systemic self-modeling (Sahal 1976, 1977, and 1978); and the "perfect systems law" (Weinberg 1975). Further, Maturana (1970) analyzes cognitive self-reference as a function of circularity in the neural structure of the brain, Zwick (1978) suggests that the "measurement problem" in quantum mechanics might be a result of a Gödelian self-reference in the mathematical formalism; and Rescher (1973 and 1980) proposes a "logic of inconsistency" to fulfill an evident need for drawing meaningful conclusions from inconsistent sets of data.

On the other hand, while none of these--except perhaps Rescher--openly promote the use of paradox, it appears that several of them could be construed as entailing paradox if one chose--notably Weinberg and von Foerster. In any case, these examples suggest that science is but a step away from implementing paradox in full force as a means of conceptualizing the world that it observes. But using paradox in this manner ultimately depends on developing logically sound modes of reasoning

which formally accommodate paradox without simultaneously degenerating into completely meaningless trivialities.

Part II of this dissertation studies a contemporary approach to this issue. In the book _Laws of form_, Spencer-Brown showed how the classical bimodal logic could be reduced to a pair of simple rules governing the manipulation of some elementary conceptual forms. He furthermore studied these forms for their capacity to describe paradoxical reasoning (Spencer-Brown 1969). The result was a system of reasoning with two semantic "states": a "marked state" (denoted by ⌐) and an "unmarked state" (denoted by a blank space) in a calculus of equations between formal expressions, known as the "primary algebra." Paradox in the primary algebra has the form

$$p = \overline{p}\rceil$$

in contrast with the form

$$p \equiv \rceil p$$

which uses the standard logical notations. Spencer-Brown was successful in showing that the propositional calculus is mirrored in the primary algebra in that every tautology P is representable as a formally derivable equation of the form p=⌐, but the ability to deal with paradoxes went only as far as representing them outside the framework of his algebra as arising through "oscillation" between the two states.

This limitation of the primary algebra was resolved by Varela (1975) in a "calculus for self-reference" which extends Spencer-Brown's conception of logical form. The marked state and the unmarked state are augmented by an "autonomous state" (denoted by ⊐ and an axiomatic system is developed which incorporates equations of the form

$$p = \square$$

as abbreviations of the paradoxical forms

$$p = \overline{\overline{p}|}$$

in a meaningfully nontrivial formalism. The title, "calculus for self-reference," stems from the analysis of paradoxical equations p=p̄⌐ as being linguistic representations of a situation in which the logical status of p is defined in terms of itself through "re-entry" into the "form of indication" ⌐. Since it is a consequence of Varela's completeness theorem that the above two equations are formally derivable from one another in the calculus, Varela's axiomatization captures the sense in which p is paradoxical if and only if p is logically autonomous.

The aim of Part II of this dissertation is to ascertain exactly how the primary algebra and the calculus for self-reference correspond to logical systems which employ the standard notational conventions--i.e., the logical connectives ⌐, &, $\lor$, $\supset$, and $\equiv$. Steps in this direction which have already been taken by others include the following: (i) Orchard (1975) strengthened Spencer-Brown's results by showing that, not only tautologies, but also antitautologies and contingencies are representable in the primary algebra, (ii) Cull and Frank (1979) showed that the primary algebra is essentially just Boolean algebra, (iii) Varela (1979) points out how of his semantics of "marked," unmarked," and "autonomous" states corresponds to the semantics of Kleene's (1938 and 1952) three-valued truth-table system which uses "true," "false," and "undefined" truth values, and (iv) Kohout and Pinkava (1980) have analyzed the algebraic structure of both the Spencer-Brown and the Varela calculi to what might well be the fullest possible extent, showing that the primary algebra may be isomorphically mapped into 8 distinct conventional systems of connectives and that the calculus for self-reference has 4380

possible interpretations, mostly as π-algebras. These works thus explore primarily the algebraic and semantic characteristics of the calculi and are concerned with isomorphisms of algebraic structures, whereas the present work shifts attention toward the formal proof-theoretic characteristics and is concerned with isomorphisms of proof systems.

It is worth noting that, while Kohout and Pinkava are correct in their observation that the Spencer-Brown and the Varela calculi contribute essentially "nothing new" algebraically, this downplays the more significant feature of these calculi; namely, that they are semantically complete axiomatizations of those algebras. Studying these calculi as proof systems per se gives added insight into the relation of the primary algebra to the classical propositional calculus and of the relation of the calculus for self-reference to the Kleene system. This leads in turn to an extension of the study of paradox into the theory of recursive decidability--which formed the original basis for Kleene's truth-tables-- and it ultimately unveils an approach to using Kleene-Varela type systems as a framework for formalizing paradoxical general systems concepts. Part II contains three chapters as follows.

Chapter 7 recasts Brown's primary algebra as a rigorously defined formal logical system PA, and shows that PA is "essentially isomorphic" with a standard axiomatization PC of the classical propositional calculus. Stated more precisely, PA is exactly isomorphic with an inessential extension PC^* of PC that is obtained by adjoining a special symbol $\underline{F}$ and a defining axiom of the form $\underline{F} \equiv P \& \neg P$. Isomorphism takes the form of a translation τ between formal languages which preserves formal derivability. A consequence of this result is that equality of expressions in PA is isomorphic with logical equivalence of propositions in PC^*.

Chapter 8 considers the calculus for self-reference, similarly recast as a rigorously defined formal logical system CSR. This system is translated isomorphically into a new system K which uses Kleene's notion of "strong equivalence" (denoted by $\cong$) as the correspondent of equality in CSR. The aim is to show that K is a semantically complete axiomatization of the Kleene truth-table system mentioned above. No such formal axiomatization of Kleene's truth-table system has heretofore been known. Later sections of Chapter 7 show that, whereas PA is Aristotelean, CSR is non-Aristotelean and can be extended to include paradoxical expressions of the form $P \cong \neg P$, without degenerating into the trivial system in which all propositions become provable. Using terminology defined in §2.1, this is to say that CSR, although inconsistent, is nevertheless a "coherent" formal logical system. In this respect, "coherence," rather than consistency, is seen to be the essential condition for a nonclassical logic to be meaningful.

Chapter 9 explores some consequences of Chapter 8, showing in particular that, under Kleene's original interpretation of the three-valued logic in the theory of partial recursion, Varela's "autonomous state" becomes isomorphic with the idea of a "totally undecideable partial recursive set." The concluding section interprets this result in several contexts, showing the potential use of formally paradoxical assertions for discussing "mechanically (i.e., recursively) unknowable" system properties from within a logically coherent frame.

The views expressed in the latter chapter happen to closely resemble those of Zwick (1978), who argues that the mathematics of physical measurement theory is such that phenomena could occur which would be unexplainable in that they would be represented by undecidable

propositions. The position taken in this dissertation is at first sight weaker, since it addresses mainly the question of how the conceptualization of such phenomena is possible (i.e., it attributes unexplainability to the observer). On the other hand, Zwick's view attributes unexplainability to the mathematics and does not rule out the possibility that there might be another mathematics in which the phenomena become explainable. On this count, the present view is actually stronger in suggesting that certain things are intrinsically unexplainable, regardless of the mathematics, and therefore require an alternate logic in order to be conceived and discussed.

The results of this part are what led to the two-leveled systems developed in Part I. It is easily verified that all the axioms and inference rules of the system K remain valid when that system is reinterpreted in the theory of fuzzy sets--with Kleene's strong equivalence here being reinterpreted as semantic equivalence of linguistic assertions. Thus, in a sense, the systems defined in Part I are extensions of K.

§1.3. Overview Of Part III

Lukasiewicz, one of the chief developers of the theory of multi-valued logics, prophesied over fifty years ago that science would eventually incorporate nonclassical modes of reasoning into its models of the observable world (cf. Gaines 1976). It is easy to see that this is in principle a completely reasonable proposition.

There is no truly compelling reason why the naturally occurring systems--taking this to include societies and economies as well as the traditional objects of scientific research--should be assumed to obey the laws of classical logic. On the contrary, in much the same way as

the discovery of non-Euclidian geometries led to new conceptions of time and space, the discovery of meaningfully coherent nonclassical logics has revealed classical logic as being only one of the many possible frameworks for organizing ideas. Logic, after all, is a property of the observer, and alternative modes of reasoning may be regarded as alternative templates, or sets of spectacles, through which the world can be viewed. While it is true that a model of some aspect of reality must conform to the known data, there does not appear to be anything in the empirical realm which insists that one form of logic should take precedence over another.

The first notable instance of Lukasiewicz's idea came to fruition during the early forties, in the form of "quantum logics" (Reichenbach 1944). Today, there exists a rapidly growing plethora of multiple-valued logics, modal logics, temporal logics, interrogative logics, and so on, all of which contribute to this theme. For specific examples cf. Rescher (1969) and the bibliographies of Gaines and Kohout (1977) and of Dunn and Epstein (1977). Fuzzy logic and the logic for self-reference, studied in this dissertation, are topics which have gained special attention in recent years.

Incumbent with the development of such logics there arise foundational and methodological issues concerning their implementation for real-world systems analysis. These include the significance of formal logic as a means of generating new perspectives on the observable world, through "selecting" alternative modes of reasoning, together with the manner in which formalized nonstandard modes of reasoning enter into the relationship of the observer to the observed. Of special importance is the issue of the meaningfulness and empirical testability of models

based on nonclassical reasoning. At the heart of this is a somewhat deeper issue concerning the status of "model theory" (i.e., the theory of formal logical systems and their semantics) as providing the "semantics of science." Apparently conflicting views on this subject have been put forth by Suppes (1961), a logician, and Bunge (1973), a philosopher of science, and a careful treatment of their discussions is imperative for placing the use of nonstandard reasoning on a solid foundation.

Chapter 10 begins with a historical overview of the development of mathematical logic, which illustrates a particular view regarding the process by which logical studies have evolved. Then, the special nature of mathematical logic as providing a means of "selecting" one's logic is studied from within this frame. The formal distinction between a "symbol" and its "meaning" is seen as that which lends mathematical logic its analytical power. Chapter 10 continues with some remarks on formal logic as a form of qualitative mathematics--specifically, as a tool for "quality representation" in situations where numerical techniques do not apply. The concluding section of Chapter 10 provides a general introduction to the theory of formal logical systems and their semantics, at a level which is appropriate for the discussion that follows.

Chapter 11 then proceeds to methodological issues. Bunge's theory of modeling for the sciences is reviewed. Then follows an analysis of the abovementioned contrast between the views of Bunge and Suppes. We ultimately concur with Bunge in that an extramathematical "semantic assumption" is essential. Bunge's work is then built upon to show how, together with his semantic assumption, concepts from the theory of formal systems provide all the necessary means for a rigorous formulation

of Bunge's basic notions of "general theory," "model object," "theoretical model," and empirical "testability." This leads to a representation of the manner in which nonclassical logics may be invoked for model construction, and the manner in which such nonstandard models enter into the relation of the observer to the observed. The concluding section shows that the models thus produced are always empirically testable, as long as the underlying logic is "coherent" in the sense discussed in Part II. This seals in the affirmative the question of whether models based on nonclassical modes of reasoning can have methodological viability.

PART I

AXIOMS FOR APPROXIMATE REASONING

CHAPTER 2

BASIC TERMINOLOGY

Concepts from the theory of formal systems, as needed for all three parts of this dissertation are laid down in §2.1. Most of these are straightforward adaptations of standard ideas, with one notable exception being the property of "coherence" which is defined herein (§2.1.5.2) specifically for use in Parts II and III. The principal references have been Shoenfield (1967) and Mendelson (1964), while secondary sources of information were Smullyan (1961) and Robinson (1963). The reader who is unfamiliar with the principles of mathematical logic may find it helpful to supplement §2.1 with a reading of §10.4, which gives a more thorough explanation of the same underlying ideas, together with some examples.

Needed concepts from the theory of fuzzy sets and approximate reasoning are given in §2.2. These have been taken from two papers by Zadeh (1975b and 1978a) and as such provide groundwork for the development in Part I. A general introduction to the subject is provided by Gaines (1976).

§2.1. Formal Logical Systems And Their Semantics

We may begin with a preview of the main definitions. A formal logical system (§2.1.4) consists of a language, a set of axioms, some inference rules, and a set of theorems, where the theorems are comprised of all expressions of the language which can be generated (derived) from the axioms by means of the inference rules. For each language there is

defined the companion notion of a <u>semantics</u> (§2.1.2), consisting of a collection of interpretations in which expressions in the language may obtain meanings. An <u>interpretation</u> generally consists of a universe of discourse, assignments of meanings from within that universe for the expressions of the language, a concept of truth or validity for expressions, and a set of valuation mappings by which one determines the truth value, or validity, of expressions in terms of their meanings. The general concept of a <u>formal logic</u> (§2.1.3) is embodied in the definition of semantic interpretation.

The basic components of a formal logical system and its semantics are illustrated in Figure 2-1. The details follow below.

§2.1.1. Languages

Let $\underset{\sim}{S}$ be any uncountably infinite collection of objects (e.g., the ordinal numbers). The objects in this collection will be used as <u>symbols</u> and as such will be the essential building blocks for all the formal languages discussed in this dissertation. Formal symbols do not strictly speaking "symbolize" anything. Rather, they are just abstract, independently existing objects which may be "assigned" various meanings within the context of different semantic interpretations (cf. §2.1.2).

A <u>string</u> is a finite sequence of symbols. The <u>length</u> of a string is its number of symbol occurrences. To indicate that a sequence of symbols is being regarded as a string, the elements of the sequence are written as <u>concatenated</u>, i.e., without separating commas. Thus, if s_1, s_2, and s_3 are symbols, the notation s_1,s_2,s_3 denotes an ordinary sequence, while $s_1 s_2 s_3$ denotes a string. Concatenation is used similarly

Formal Logical
System

Formulas

Theorems

Axioms

Inference Rules

Semantic
Interpretations

Universe
of
Discourse

Truth
Values

Meanings

Truth
Valuations

Univ.
of Disc.

Tr.
Vals.

Mean's
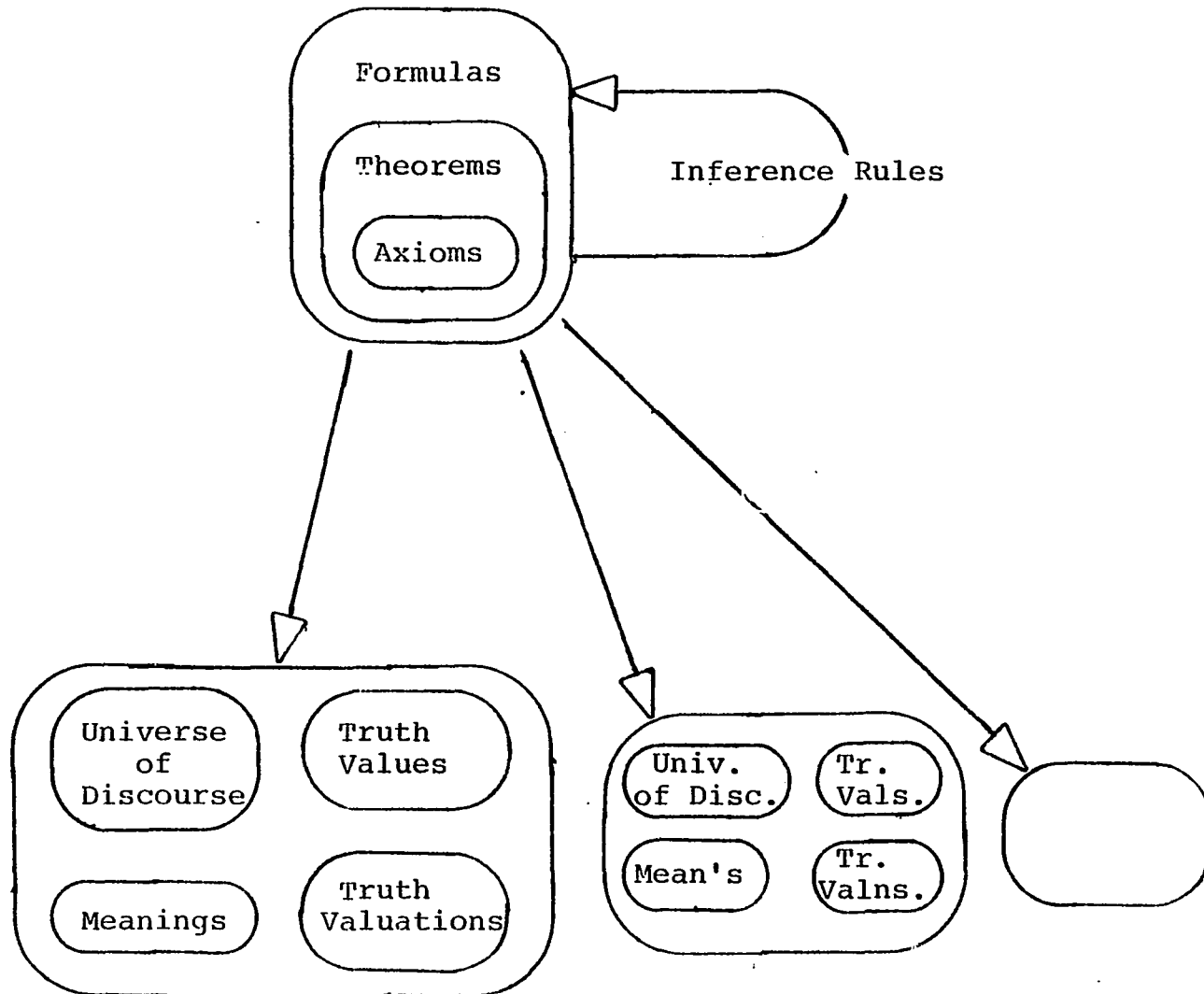
Tr.
Valns.

Figure 2-1. Formal logical systems with interpretations.

in case the s_i are strings--i.e., the concatenation to two or more strings is a string.

If s and s' are strings, and s_0 is a symbol, then the notation $s(s'/s_0)$ denotes the string that is obtained from s by <u>uniform</u> <u>substitution</u> of an occurrence of s' for each occurrence of s_0 in s. It is tacit that, if s does not contain any occurrences of s_0, then $s(s'/s_0)$ is just s.

A <u>formal</u> <u>language</u> consists of a set of <u>symbols</u>, together with a set of <u>well-formed</u> <u>strings</u>, as described below:

1) The set of <u>symbols</u> to be used specifically <u>for</u> L are assumed to be members of the collection $\underset{\sim}{S}$ given above. The only requirement on this set is that it contain a sufficient, countable or uncountable, number of members as appropriate for the intended use of L. Most languages will have a few symbols designated as <u>logical</u> <u>connectives</u>, which, in the case of standard languages, typically include a <u>negation</u> <u>symbol</u>, denoted by ⌐, and a <u>disjunction</u> <u>symbol</u>, denoted by V.

2) The <u>well-formed</u> <u>strings</u> of L are selected from the set of all strings for L. For a specific L, well-formed strings may be defined recursively or in accordance with a context-free grammar. Depending on the language, they are variously referred to as formulas, expressions, equations, equivalences, etc.

In general, languages are defined for the purpose of formalizing the essential ideas in an intended set of interpretations, i.e., a semantics. Thus, although strings per se are purely <u>syntactical</u> objects, devoid of any specific meanings, well-formed strings have appropriate <u>forms</u> so that they will be meaningful under certain interpretations of

the constituent symbols. This aspect of the creation of formal languages
is discussed more fully in §10.4.1.

§2.1.2. Semantics

A _semantics_ Σ for a formal language L is a class of interpretations
for L, where an _interpretation_ I for L is defined in four steps, as
follows:

1) A set U_I is specified as a _universe_ _of_ _discourse_, the members
of which are referred to as _individuals_. A universe of discourse provides
the _context_ in which well-formed strings of L obtain _meanings_. For
sufficiently simple languages, such as that of the propositional calcu-
lus and most of the other systems studies in Part II, universes of
discourse may be assumed to be empty, since the meanings of the logical-
connectives are always implicit in their truth-functional behavior as
defined by the interpretation's valuations (see the following). For
richer languages, such as studied in Part I, universes are typically
non-empty, and the interpretation then includes specific meaning assign-
ments as described below.

2) _Meaning_ _assignments_ are defined in terms of mappings of certain
symbols in L to entities in, or definable on, the universe U_I. For
example, an "individual constant" in L (cf. §3.1.1) is assigned as its
meaning a specific individual in U_I, and a "relation symbol" in L (also
§3.1.1) is assigned as its meaning a specific relation of individuals in
U_I. Assignments of meanings to symbols ultimately yield meanings for
expressions. For example, if a and b are individual constants, and if $=$
denotes the equality relation, then the expression a=b "means" that a is
equal to b.

3) For every interpretation, a set of <u>values</u> is chosen and there is defined a set of <u>valuations</u>. The <u>values</u> may be any objects--e.g., the letters T and F, the numbers 0, ½, and 1, the members of the unit interval [0,1], or the fuzzy subsets of [0,1]. In most places, these are called <u>truth-values</u>, but an exception is in Part II, where certain values are referred to as logical <u>states</u>. <u>Valuations</u> are mappings of the well-formed strings of L, or a certain subset of well-formed strings of L, into the set of values. For simple languages, which do not require universes of discourse, valuations merely take the place of truth tables, characterizing the truth-functional behavior of logical connectives. For richer languages, valuations apply to well-formed strings which make assertions about individuals in U_I, and they yield a truth-value in terms of the assertion's meaning. For example, a string of the form a=b receives the values T or F depending on whether the individual that was assigned as a meaning for a is, or is not, the same individual as was assigned to b. To avoid possible confusion, it should be noted also that strings of the form x=y, where x and y are "individual variables" (cf. §3.1.1), are not included in the domain of the valuation mappings, since x and y vary over the universe U_I but do not have any specific individual assignments. A discussion of this appears in §10.4.1, and an example in §3.2.

4) A concept of <u>truth</u> or <u>validity</u> is specified for well-formed strings. For all languages this concept is essentially predetermined by the set of valuations. In the propositional calculus, for example, one makes use of the concept of a <u>tautology</u>, this being a well-formed string which is assigned the value 1 (or T) by all valuations (cf. §7.1). In richer languages, the appropriate concept is that of semantic <u>validity</u>,

which is determined by the valuations as a property of the meaning of a
string under the given interpretation (meaning assignments) of its
constituent symbols. For example (cf. §3.2), a string of the form a=a
would be valid, since it receives the value 1 under all valuations; and
a string of the form x=x would be valid, since, no matter what individual
constant is put in place of x, the string likewise receives the value 1
under all valuations. The idea used in the latter example makes this
notion of validity applicable to all well-formed strings of L, whether
or not they are within the domain of the valuation mappings. Note that
it is implicit in the foregoing that only well-formed strings which make
assertions about specific individuals are formally regarded as "true" or
"false." The property of a strings being valid in an interpretation I
is asserted by the notation I⊨S.

§2.1.3. Logics

A formal logic consists of a formal language L, together with a
semantics Σ for L. For any semantics, it is normally implicit in the
definition of its interpretations that all the interpretations have
their valuation mappings acting on the same logical connectives in the
same way. For example, in classical logics, the expression P receives
the value T, if and only if P receives the value F, and P receives the
value F if and only if P receives the value T. Thus, the nature of the
logic--i.e., as being two-valued, many-valued, fuzzy-valued, etc.--is
embodied in Σ; and any given L may serve as the language of several
different logics.

§2.1.4. Formal Systems

A formal logical system is normally defined for the purpose of
characterizing the properties of a given set of semantic interpretations.
In this case, it is desired to find a set of axioms and inference rules
from which one can formally derive all the well-formed strings that
happen to be true about (or valid in) all the given interpretations.
Typically, this set of interpretations is a subset Σ' of a given seman-
tics Σ for a language L, in which case L becomes the language of the
formal system and is denoted by L(F). Then, a formal logical system F
consists of a language L(F), together with some axioms, inference rules,
and theorems, as follows.

§2.1.4.1. Axioms

The only requirement placed on axioms of F is that they be well-
formed strings of L(F). Yet, since axioms are customarily selected
specifically for the above mentioned objective of characterizing a
subset Σ' of a semantics Σ for L(F), one naturally chooses only axioms
which are valid in every interpretation in Σ'.

§2.1.4.2. Inference Rules

The inference rules of F are mappings of the set of well-formed
strings of L(F) into itself. In some treatments, inference rules are
described by means of a "schema." For example, the rule of modus ponens
(§3.3.1 and §7.1) may be described by

$$\frac{\begin{array}{l} P \\ P \supset Q \end{array}}{Q}$$

which says that, from the formula P, together with a formula of the form
P⊃Q, one may infer the formula Q.

In this dissertation, an n-ary inference rule is regarded as a mapping in the sense of being a set of (n+1)-tuples of well-formed strings having certain _forms_. For example, modus ponens is herein defined as the binary rule consisting of triples having the form $(P, P \supset Q, Q)$. Each such (n+1)-tuple is an _instance_ of the rule; and if $(s_1, \ldots, s_n, s_{n+1})$ is a rule instance, then $s_1, \ldots, s_n$ are the _hypotheses_ of that instance, and s_{n+1} is the _conclusion_. Thus one captures the sense in which a certain conclusion may be _derived_ (or _inferred_) _from_ a set of hypotheses _by_ _means_ _of_ _a_ _rule_. Owing to the desire to characterize a set Σ' of interpretations for L(F), one naturally chooses for F only inference rules which always act so as to _preserve_ _validity_ in each I in Σ', i.e., so that only valid conclusions can be inferred from valid hypotheses.

§2.1.4.3. Theorems

The set of _theorems_ of F is the smallest set of well-formed strings of L(F) that contains the axioms of F and is closed under the inference rules of F. An alternative, and sometimes useful, description of this set is given by the following inductive definition:

1. The axioms of F are theorems of F,

2. If the hypotheses $s_1, \ldots, s_n$ in an instance of an inference rule of F are theorems of F, then the conclusion, s_{n+1}, is a theorem of F,

3. Nothing is a theorem of F except as required by 1 and 2.

It follows that the set of theorems of F is uniquely determined by the language, axioms, and inference rules of F.

§2.1.4.4. Inductive Definitions And Inductive Proofs

The above <u>inductive</u> <u>definition</u> of a set of theorems is typical of many definitions used throughout the study of formal systems. Oftentimes, the analogue of item 3 is omitted, in which case it is tacitly assumed. Whenever it is desired to establish that the members of such a set have a certain property, one uses a corresponding <u>inductive</u> <u>proof</u>. Such a proof may be regarded as a generalization of an ordinary proof by mathematical induction, having its own form of <u>induction</u> <u>hypothesis</u>. For example, to establish that the theorems of a system F have a certain property P, one shows:

1. The axioms of F have the property P, and

2. If s_{n+1} is the conclusion in an instance $(s_1,\ldots s_n,s_{n+1})$ of an inference rule of F, and if the hypotheses $s_1,\ldots,s_n$ have P, then the conclusion s_{n+1} also has P.

In this example, the induction hypothesis is the assumption in 2 that the property indeed holds for $s_1,\ldots,s_n$. A proof using this technique appears below, Proposition 2.1.5.1-1.

§2.1.5. Further Properties And Terms

The following sections discuss the concepts of proof, consistency, coherence, Aristotelean vs. non-Aristotelean, and extension for formal systems. These are all strictly <u>syntactical</u> properties in that they refer only to the formal structure of a system--in other words, they are independent of any underlying semantics. Let F be an arbitrary formal logical system.

§2.1.5.1. Proofs

A _proof_ of a string s in F is a finite sequence $s_1, \ldots, s_n$ of well-formed strings of L(F) such that s is s_n, and for each index i=1,...,n, either:

1. s_i is an axiom of F, or

2. s_i is the conclusion in an instance of an inference rule of F in which the hypotheses of that instance are all among the s_j for j<i.

Such a proof is said to have _length_ n. The notation F⊢s is used to assert the condition that there is some proof of s in F, and F⊬s asserts the contrary. (Note: Many authors follow Frege (1879) by writing ⊢$_F$s and ⊬$_F$s. The notation adopted here is typographically more convenient.)

Proposition 2.1.5.1-1. A well-formed string s of L(F) is a theorem of F if and only if F⊢s.

Proof. To show that, if s is a theorem, then F⊢s, we use proof by induction on the set of theorems of F.

Case 1: the string s is an axiom of F. Then the one-string sequence, s, is a proof of s in F.

Case 2: the string s is a conclusion s_{n+1} in some instance $(s_1, \ldots, s_n, s_{n+1})$ of an inference rule of F. In this case, the induction hypothesis is that F⊢s_i, for each i≦n. Let $s_1', \ldots, s_m'$ be the sequence of strings that is obtained by putting the proofs of the s_i, for i≦n, together, in the order of the indices of the s_i. Then the sequence $s_1', \ldots, s_m', s_{m+1}$ is a proof of s_{m+1} in F, by the definition of a proof. Thus, F⊢s by induction on the theorems of F.

To show that, if $F \vdash s$, then s is a theorem of F, we can use mathematical induction on the length of proofs in F. Assume that s has a proof of length n.

Case 1: n=1. Then s is an axiom, by the definition of a proof. Hence s is a theorem, by item 1 in the definition of a theorem.

Case 2: n>1; say that the given proof of s is $s_1,\ldots,s_n$, where s_n is s. The induction hypothesis asserts that all proofs in F having length less than n yield theorems of F. Thus all of $s_1,\ldots,s_{n-1}$ are theorems of F, since each initial segment of $s_1,\ldots,s_n$ is a proof in F. By item 2 of the definition of a proof, s_n is inferred from some of the s_i for i<n by means of an inference rule of F. It therefore follows by item 2 in the definition of a theorem that s_n (which is the string s) is a theorem of F. This completes the proof, by mathematical induction on n. □

In the sequel, Proposition 2.1.5.1-1 is used tacitly to equate the properties of a string's being derivable and of its being a theorem in a formal logical system. Thus, the notation $F \vdash s$ is used to assert that s is a theorem of F.

§2.1.5.2. Consistency and Coherence

The property of consistency for a formal system is based on the notion of a well-formed string's being an inconsistency, where an inconsistency is a string having a certain form, depending on the system's language. For example, in a system which is based on the usual logical connectives, inconsistencies have the form P&¬P--in which case they are referred to as contradictions. Given an appropriate definition of an inconsistency for a formal logical system, the system itself is said to

be consistent if the theorems of the system do not include any such inconsistent strings.

The following property asserts that the set of theorems has a nontrivial structure. Precisely, a formal logical system F is coherent if there exists at least one well-formed string s of L(F) such that F⊬s.

This property of a formal system is used in the present work for establishing the characterization of Aristotelean vs. non-Aristotelean systems discussed below (§2.1.5.4). Here it should be mentioned that this notion of coherence bears a connection with Rescher's (1973) "coherence theory of truth" which was developed in an effort to obtain meaningful inferences from inconsistent sets of data. Rescher's approach may be characterized as an "internal" one, focusing on a concept of semantic truth which provides the fabric upon which such inferences can be performed; and "coherence" is taken in the literal sense as a property of "hanging together"--a set of propositions is coherent if the propositions are not truth-functionally independent.

The present approach, by contrast, is an "external" approach which focuses on a property of formal logical systems. It makes no reference to propositions being, or not being, independent, and it takes "coherence" more in the sense of signifying "meaningfulness." The essential ingredient is retained, however, inasmuch as coherent systems as defined herein also allow for meaningful inferences from inconsistent premises.

The above definition is in fact more closely aligned with Rescher's most recent thinking on this problem (Rescher 1980). What is here termed "inconsistency" for a formal logic is there termed "strong inconsistency," and what is here termed "incoherence" is there termed "logical chaos."

§2.1.5.3. Extensions

An _extension_ of a formal system F may be obtained by performing one or more of the following operations:

1. Adjoin some _new_ symbols--i.e., symbols not already in L(F)--to L(F), which generates some new well-formed strings,

2. Adjoin some additional axioms to the set of axioms of F, thereby generating some new theorems,

3. Adjoin some additional inference rules to the set of inference rules of F, which also generates new theorems.

In case of operation 1, it is always tacit that the domains of the inference rules of F--considered as mappings--are extended to the new language. If S is a set of symbols, A is a set of axioms, and R is a set of inference rules, then the system obtained by adjoining these items to F may be denoted by F[S,A,R].

A _simple_ _extension_ of F is an extension that is obtained by performing operations only of the above type 2. An extension F* of F is an _inessential_ _extension_ of F if, for every well-formed string s of L(F) F*⊢s implies that F⊢s--i.e., F* is such that no strings of L(F) can be derived as theorems of F* unless they are already derivable in F. Such extensions are generally nontrivial, however, since L(F*) may contain strings provable only in F*.

§2.1.5.4. Aristotelean And Classical Systems

It will be said that a well-formed string s _follows_ _from_ a string s' _in_ F if F[s']⊢s. A system F will be termed _Aristotelean_ if, in F, _every_ well-formed string of L(F) follows from _any_ inconsistency; otherwise F will be termed _non-Aristotelean_. It follows that:

1. A system is Aristotelean if and only if all its coherent simple
   extensions are consistent,

2. A system is non-Aristotelean if and only if it has at least one
   simple extension which is inconsistent yet coherent.

Regarding this usage of terminology, it should be noted that we
here deviate from an established convention. Traditionally, "Aristotelean"
has been taken to mean "includes the law of the Excluded Middle (which
in standard languages is expressed by the formula $P \lor \neg P$), and "non-
Aristotelean" has traditionally meant "without Excluded Middle." This
distinction does not correspond to the above since, first, it is possible
to weaken the propositional calculus to a system which still contains
Excluded Middle but in which inconsistency does not imply incoherence,
and second, there are formal systems not containing Excluded Middle yet
in which inconsistency does imply incoherence. A notable example of the
latter is Kleene's axiomatization of the logic of intuitionism
(cf. Mendelson 1964, p. 43, Exercise 2(e), parts (i) and (v)).

It may be argued, however, that the present usage is equally
acceptable, in fact is possibly more appropriate, since it captures the
general spirit of Aristotle's thinking without becoming involved with
the details of a system's formal structure. The sense of Aristotelean
logic was that contradiction (inconsistency) should be forbidden, and
this for the evident reason that he viewed contradiction as being tanta-
mount to incoherence. Readers familiar with the work of E. Post (1921)
may note that the present conception of "Aristotelean" bears a resemblance
to what is now called "post-complete" and might equally have been termed
"post-contradiction-complete."

Last, we should consider also how the above distinction relates to the somewhat less precise distinction between "classical" and "nonclassical" systems. In this dissertation is employed the traditional meaning of classical as referring to those systems which embody both the linguistic structure and the full proof-theoretical strength of the classical propositional calculus, together with its immediate derivates, such as the first-order and second-order systems. Given this terminology, we have immediately that classical implies Aristotelean; but not conversely. In §8.4 it is shown, for example, that Spencer-Brown's primary algebra, although clearly nonclassical, is Aristotelean. Also, the linguistic theories of Chapters 3 through 6 are nonclassical in their use of a "two-leveled" linguistics, but are Aristotelean since they are classical at the "outer" level. Examples of nonclassical non-Aristotelean systems appearing in this dissertation are Varela's calculus and its isomorphisms (Chapters 8 and 9).

§2.1.6. Semantic Models And Semantic Completeness

A semantics Σ for a formal language L may be denoted $\Sigma(L)$; in case L is the language of a formal system F, then Σ may alternatively be denoted by $\Sigma(F)$. Let F be a formal logical system and let Σ be a semantics for L(F). An interpretation I in $\Sigma(F)$ is a model of F if every theorem of F is valid in I. This property may be expressed by the notation I⊧F. A well-formed string s of L(F) is valid in $\Sigma(F)$ if s is valid in every model of F in $\Sigma(F)$, in which case one may write $\Sigma(F)$⊧s.

Proposition 2.1.6-1. For any formal logical system F and any well-formed string s of L(F), F⊢s only if $\Sigma(F)$⊧s.

Proof. This is immediate by the foregoing definitions. □

A formal logical system F is <u>complete</u> with respect to a semantics Σ(F) if validity in Σ(F) implies formal derivability in F--i.e., if the converse of Proposition 2.1.6-1 holds. This concept of semantic completeness captures one part of what is meant by saying that a formal system <u>characterizes</u> a certain semantics. To wit, in combination with Proposition 2.1.6-1, it asserts that a well-formed string is derivable in F if and only if it is valid in every "valid interpretation"--i.e., model--of F. Usually it is only the models in a semantics that are mathematically interesting.

A further aspect of formal characterization is that the language of F must be rich enough to express the essential properties of the given semantics. Satisfaction of this requirement is usually determined by an intuitive judgment, based on one's knowledge of the structures which are to be taken as interpretations. A discussion of these and related ideas also appears in §10.4.1, and an example appears in §3.4.2.

§2.1.7. Classes of Formal Systems

When defining a formal system F one usually implicitly specifies a generic type of system, representing a class $\mathcal{C}$ of systems, all of which have certain linguistic-axiomatic properties in common. Accordingly, when defining a semantics Σ for L(F), one winds up defining a semantics for every formal system in $\mathcal{C}$, in which case, a proof of semantic completeness for F serves as a completeness proof for every system in $\mathcal{C}$. This justifies saying that the entire class $\mathcal{C}$ is <u>complete</u> with respect to the <u>semantics</u> $\mathcal{S}$, where $\mathcal{S}$ is the class of all semantics for systems in $\mathcal{C}$.

§2.1.8. Common Abbreviations

The notation "iff" is used as an abbreviation for the phrase "if and only if." Membership of an element a in a set A may be indicated by $a \varepsilon A$.

§2.2. Fuzzy Sets And Logics

Since this work is focused on the specific task of axiomatizing the logic inherent in Zadeh's fuzzy linguistics, it necessarily leaves out a major portion of what is nowdays a very rich body of mathematical knowledge. The following treatment, moreover, is intentionally brief, and might profitably be supplemented with a reading of Zadeh (1975b) and Gaines (1976). Only those ideas specifically needed for Chapter 3 and 4 are provided in detail. Other topics, such as the full fuzzy logic, linguistic truth, and fuzzy inference, are taken up again in the appropriate sections of Chapter 5.

§2.2.1. Fuzzy Sets

A fuzzy set is a collection of objects which has unsharp boundaries (Zadeh 1972). This idea is defined mathematically as a generalization of the classical idea of a subset of a given universe U. Let U be a universe of discourse--i.e., any set. Then a classical subset S of U may be represented by a membership function $\mu_S : U \to \{0,1\}$ defined by: $\mu_S(x)=1$, if $x \varepsilon S$; and $\mu_S(x)=0$, if $x \notin S$. By contrast, a fuzzy subset S of U is defined as the analagous object which is represented by a function $\mu_S : U \to [0,1]$, where [0,1] denotes the closed unit interval. In this case the value $\mu_S(x)$ is referred to as the degree or grade of membership of x in S.

As an example, assume that S is a fuzzy set of "red" objects (considered as a fuzzy subset of a universe consisting of all objects). Then a ripe red apple might be assigned the value 0.9, indicating a high degree of membership in S, while a ripe orange might have the value 0.4, indicating a moderate degree of membership in S.

Another example, appearing frequently in Zadeh's papers, is a pair of explicitly defined membership functions for the collections of "young" and "old" ages of people. Here let U be the set of ages from 0 to 100. Then a fuzzy set Y of "young" ages may be defined by

$$\mu_Y(x) = \begin{cases} 1 & \text{if } x \leq 25 \\ [1 + (\frac{x-25}{5})^2]^{-1} & \text{if } x > 25, \end{cases}$$

and a fuzzy set 0 of "old" ages may be defined by

$$\mu_0(x) = \begin{cases} 0 & \text{if } x \leq 50 \\ [1 + (\frac{x-50}{5})^{-2}]^{-1} & \text{if } x > 50. \end{cases}$$

These functions can be represented graphically as in Figure 2-2. Note that specification of the intended universe of discourse is always necessary for establishing the context in which such interpretative membership functions are to be defined. For example, the above definitions of μ_Y and μ_0 would not be appropriate in the context of ages of automobiles or of galaxies. In considering a fuzzy set S as the meaning of a natural language expression α, the grade of membership of an object in S is referred to as that object's degree of compatibility with α.

A crossover point in a fuzzy subset S of a universe U is an object x in U for which $\mu_S(x)=0.5$. The support of fuzzy subset S of U is the set of members x of U for which $\mu_S(x)$ is positive. A fuzzy singleton is a fuzzy subset whose support consists of exactly one element in U.
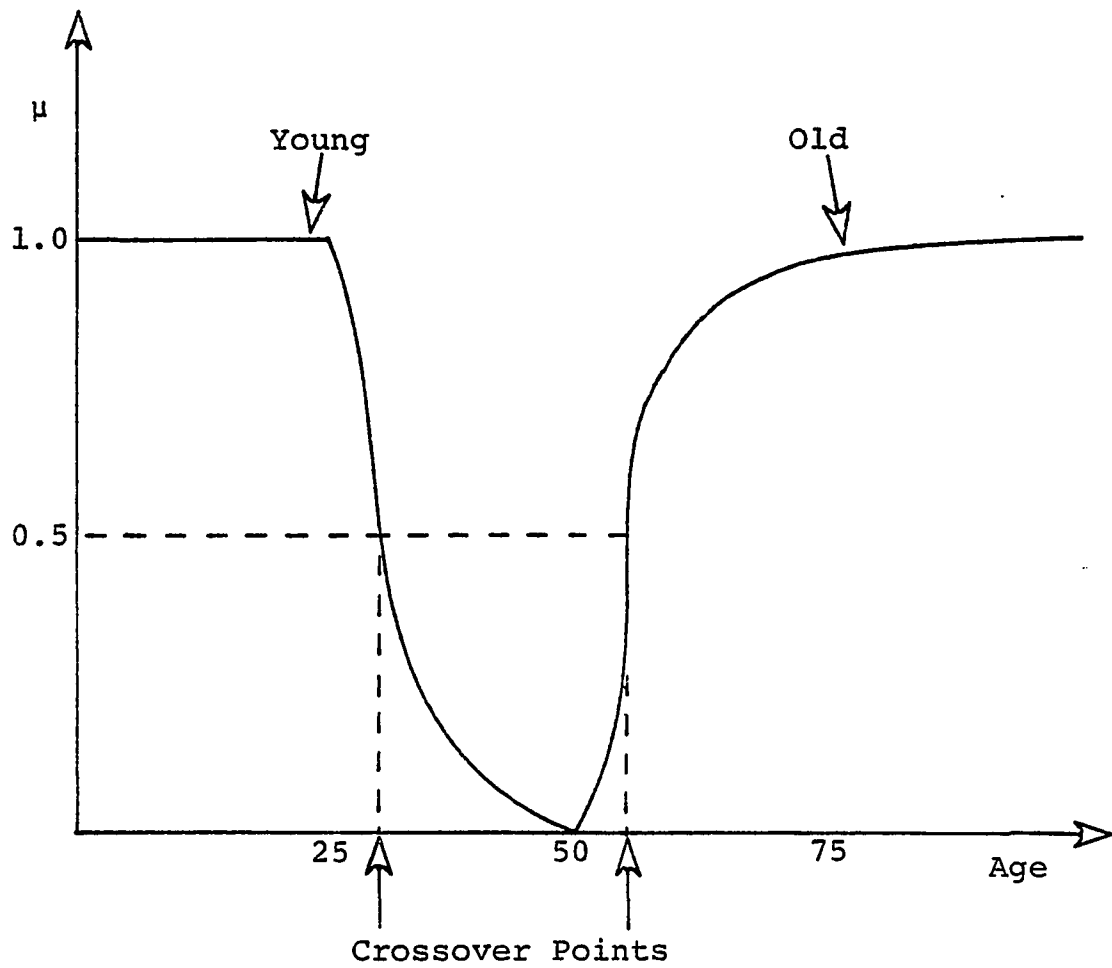
<u>Figure 2-2</u>. Compatibility functions for "Young" and "Old."

Evidently, a fuzzy subset S of U may be represented as a collection of ordered pairs (μ, x) where μ is a number in the interval $[0,1]$ and x is an object in U. If S has finite support, say $x_1., \ldots, x_n$, and for each $i=1, \ldots, n$, μ_i is the grade of membership of x_i in S, then S may be represented symbolically as a combination of fuzzy singletons by

$$S = \mu_1 x_1 + \mu_2 x_2 + \cdots + \mu_n x_n,$$

or more compactly, by

$$S = \sum_{i=1}^{n} \mu_i x_i.$$

A similar notation might be used in case the support is countably infinite. If the support is a continuum, then S may be denoted by

$$S = \int_U \mu_i x_i.$$

Thus the signs "+", "Σ" and "$\int$" are here used only as convenient notational devices and are not taken in their usual mathematical meanings.

Sometimes it is useful to write the ordered pair $\mu_i x_i$ with a separating slash mark, i.e., as μ_i / x_i. For example, where U is the set of ages from 0 to 100, a fuzzy set M of "middle" ages might be given explicitly by

$$M = 0.3/40 + 0.5/41 + 0.8/42 + 0.9/43 + \sum_{x_i=44}^{48} 1/x_i$$

$$+ 0.9/49 + 0.8/50 + 0.7/51 + 0.6/52 + 0.5/53$$

$$+ 0.4/54 + 0.3/55.$$

§2.2.2. Fuzzy Relations

Let $\alpha_1, \ldots, \alpha_n$ be variables ranging over universes of discourse $U_1, \ldots, U_n$, respectively. A <u>fuzzy</u> <u>relation</u> R <u>of</u> $\alpha_1, \ldots, \alpha_n$ (or <u>on</u> $U_1, \ldots, U_n$) is a fuzzy subset of the cartesian product $U_1 \times U_2 \times \cdots \times U_n$. Then R is

represented by an n-ary membership function $\mu_R:U_1 \times \cdots \times U_n \to [0,1]$ and may be expressed by the notation

$$R = \int_{U_1 \times \ldots \times U_n} \mu_R(x_1,\ldots,x_n)/(x_1,\ldots,x_n).$$

To illustrate (from Zadeh, 1972): if $U_1 = \{TOM,DICK\}$ and $U_2 = \{JOHN,JIM\}$, then a fuzzy relation R of "resemblance" between members of U_1 and U_2 may be defined by

$$R = 0.8/(TOM,JOHN) + 0.6/(TOM,JIM)$$
$$+ 0.2/(DICK,JOHN) + 0.9/(DICK,JIM).$$

Let R be a fuzzy relation on a pair of universes U_1,U_2, and let S be a fuzzy relation on U_2 and a third universe U_3. Then the <u>composition</u> of R and S, denoted by RoS, is a fuzzy relation on U_1,U_3 defined by

$$RoS = \int_{U_1 \times U_3} \sup_{x_2}(\min[\mu_R(x_1,x_2),\mu_S(x_2,x_3)])/(x_1,x_3),$$

where sup refers to the maximum over all possible values of the indicated variable.

Continuing the foregoing example, suppose that S is a "resemblance" relation between members of U_2 and $U_3 = \{PETE,MIKE\}$, given by

$$S = 0.6/(JOHN,PETE) + 0.9/(JOHN,MIKE)$$
$$+ 0.5/(JIM,PETE) + 0.4/(JIM,MIKE).$$

Then the composition RoS is

$$RoS = 0.6/(TOM,PETE) + 0.8/(TOM,MIKE)$$
$$+ 0.5/(DICK,PETE) + 0.4/(DICK,MIKE),$$

which may be taken as a "resemblance" relation between the members of U_1 and U_3.

§2.2.3. Fuzzy Boolean Operations And Related Definitions

In fuzzy set theory the definitions of fuzzy complement, union, intersection, and of fuzzy set equality, inclusion, and cartesian product are simple generalizations of the classical definitions.

Let A and B be fuzzy subsets of a universe U, with membership functions μ_A and μ_B. The three Boolean operations are generalized as follows. The fuzzy complement of A, denoted by -A, is defined by

$$\mu_{-A}(x) = 1-\mu_A(x) \qquad \text{for all } x\varepsilon U,$$

the fuzzy union of A and B, denoted by A∪B, is defined by

$$\mu_{A\cup B}(x) = \max[\mu_A(x),\mu_B(x)] \qquad \text{for all } x\varepsilon U,$$

and the fuzzy intersection of A and B, denoted by A∩B, is defined by

$$\mu_{A\cap B}(x) = \min[\mu_A(x),\mu_B(x)] \qquad \text{for all } x\varepsilon U.$$

The fuzzy sets A and B are equal, notation A=B, if

$$\mu_A(x) = \mu_B(x) \qquad \text{for all } x\varepsilon U,$$

and the fuzzy set A is included in the fuzzy set B, notation A⊆B, if

$$\mu_A(x) \leqq \mu_B(x) \qquad \text{for all } x\varepsilon U.$$

Now suppose that B is a fuzzy subset of some possibly different universe of discourse U'. The cartesian product of A and B, notation A×B, is the fuzzy relation on U×U' defined by

$$\mu_{A\times B}(x,x') = \min[\mu_A(x),\mu_B(x')] \qquad \text{for all } (x,x')\varepsilon U\times U'.$$

All of the foregoing reduce to the classical case if μ_A and μ_B are restricted to the set $\{0,1\}$.

In fuzzy linguistics (cf. §2.2.5), these mathematical ideas are used in various ways to provide meanings for natural language expressions. In particular, the fuzzy Boolean operations -, ∪, and ∩ provide fuzzy interpretations of the Boolean connectives "not" (negation), "or" (disjunction), and "and" (conjunction). Example: If the fuzzy sets Y and O

defined in §2.2.1 are taken as the meanings of the expressions "young" and "old," then the meaning of the expression "not young and not old" would be defined as $(-Y) \cap (-O)$.

§2.2.4. Linguistic Hedges

Linguistic hedges are modifiers, usually adjectives, such as "very," "quite," "approximately," "more or less," "somewhat," and so on. The strength of the theory of fuzzy sets as a basis for linguistic reasoning stems from the capacity for linguistic hedges to be represented as well-defined operations on fuzzy-set membership functions. A typical example is the interpretation of the "very" hedge in terms of the squaring function on [0,1]. Let A be a fuzzy subset of a universe U, and suppose that A is being used as the meaning of some linguistic expression α. Then the meaning of the expression "very α" may be given as the fuzzy subset very-A of U defined by

$$\mu_{\text{very-A}}(x) = [\mu_A(x)]^2 \qquad \text{for all } x \varepsilon U.$$

The effect of applying the "very" operator to the foregoing meanings Y and O of "young" and "old" is depicted graphically in Figure 2-3. This particular operator is an intuitively natural interpretation of "very," since, in order for an object in U to have a high degree of membership in very-Y (very-O) it must have an even higher grade of membership in Y (in O).

The most extensive treatment of hedges to date is Zadeh (1972), which defines a collection of general operations--"concentration," "dilation," "fuzzification," etc.--which may be used to provide interpretations of a wide variety of linguistic expressions. Such interpretations are always arbitrary, except that they should reasonably conform to
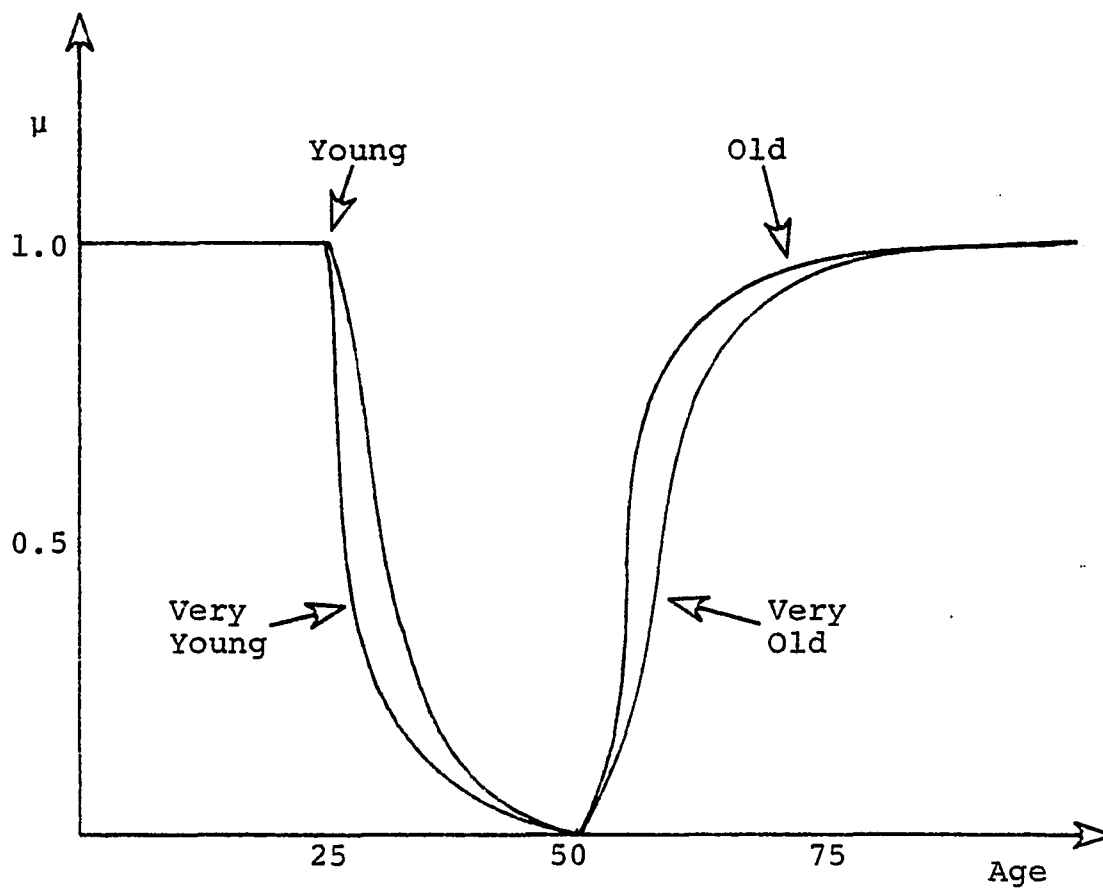
Figure 2-3. Effect of the "very" operator.

ordinary usage of the expression being interpreted. A detailed analysis

of specific hedges is foregone in this dissertation, since the axiomatiza-

tion developed herein is aimed only at capturing the general role of

hedges in linguistic reasoning and does not attempt to provide hedges

with explicit formal definitions.

§2.2.5. Linguistic Variables

A linguistic variable is a variable whose values are expressions

in a natural or artificial language. For example "Age" becomes a linguis-

tic variable if we take its values to be expressions like "young," "not

young," "very young," "old," "not very young and not very old," and so

on. Such expressions will be referred to as linguistic terms. If χ is

a linguistic variable, its linguistic terms comprise a term set $T(\chi)$

which can be generated from a smaller set of atomic terms, or terminals,

by means of a context free grammar G. The subject of grammars is taken

up in §2.2.6.

The foregoing sections give several examples of how a linguistic

term can be given a meaning, or interpretation, as a fuzzy subset of a

universe U. If α is a term in $T(\chi)$, then its meaning will be denoted by

$I(\alpha)$, where I is referred to as an interpretation function, or a meaning

assignment function, for the linguistic variable χ. It follows that a

linguistic variable may be characterized as a quintuple $(\chi, T(\chi), U, G, I)$.

The basic components of a linguistic variable are illustrated graphically

in Figure 2-4.

It is implicit in this definition that a linguistic variable may

take n-ary linguistic terms, i.e., linguistic terms whose meanings are

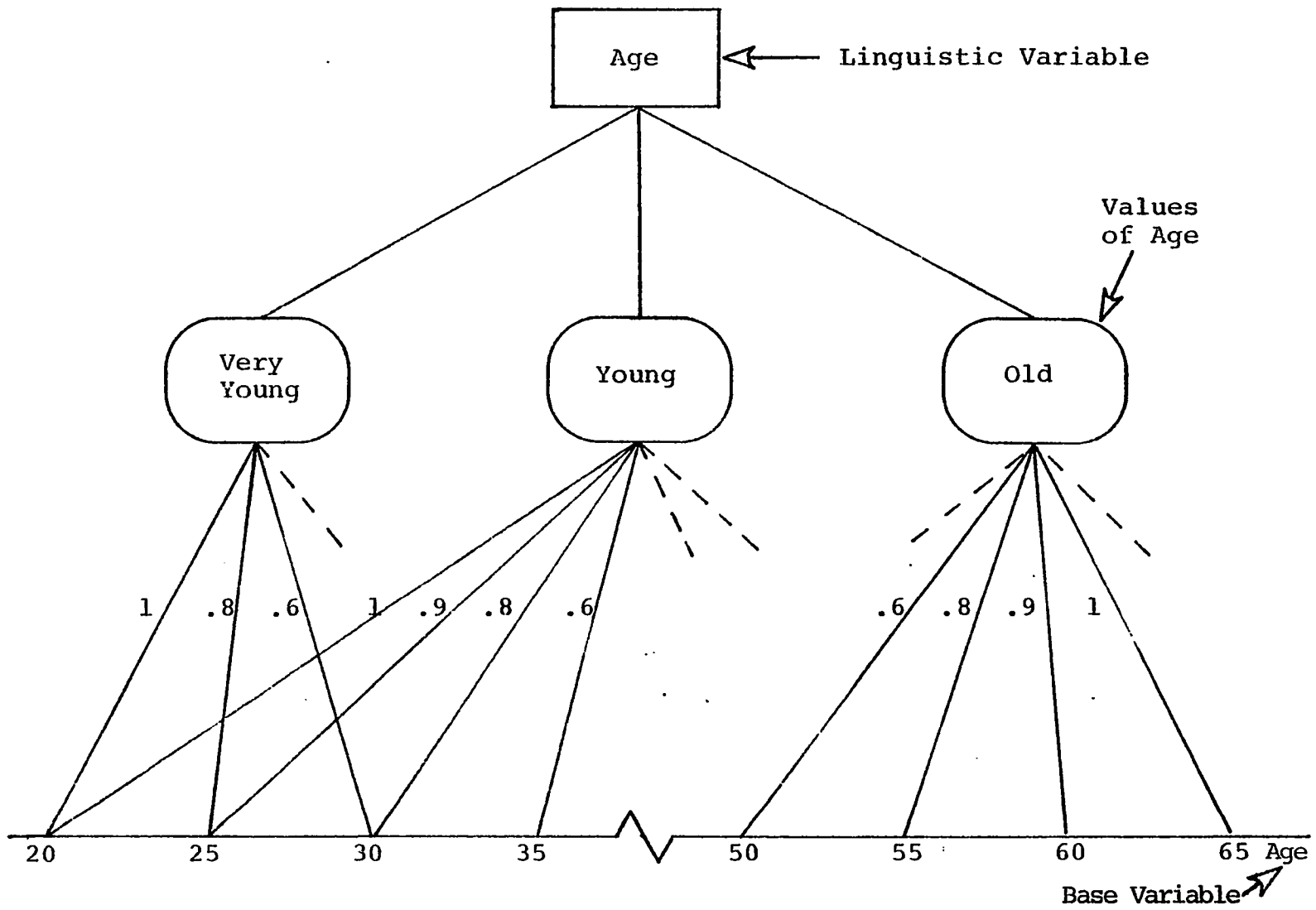n-ary fuzzy relations, as its values. In case the variables of such

Figure 2-4. Structure of a linguistic variable.

relations range over the same universe U, then one considers only relations on U, while in the more general case that they range over different universes, then U would be taken as the Cartesian product of those universes. Further, in a specific application, one usually considers several linguistic variables simultaneously over the same universe U. For example, a description of "young, tall, and agile men" would entail the three linguistic variables "Age," "Height," and "Physical Fitness" over the same universe of human males. On the other hand, it is sometimes necessary to consider different linguistic variables simultaneously over totally distinct universes. A situation in which this is done is where one formally includes a notion of linguistic "Truth" (cf. §2.2.9).

Zadeh (1978a) provides the following. Two linguistic terms α and β of some linguistic variable χ are <u>semantically equivalent</u> if the meaning assignment function I is such that $I(\alpha)=I(\beta)$, where "=" is equality of fuzzy sets as defined in §2.2.3. The term α <u>semantically entails</u> the term β if $I(\alpha)\subset I(\beta)$, where "$\subset$" is the fuzzy set inclusion relation defined in §2.2.3.

Semantic equivalence characterizes the situation of two linguistic terms representing the same semantic meaning. For example, in the foregoing, it turns out that "young" is not semantically equivalent to "not old," and "old" is not semantically equivalent to "not young," i.e., they are not semantic <u>antonyms</u>. On the other hand, they could become antonyms simply by redefining the compatability functions for "young" and "old," perhaps along the scheme of "true" and "false" given in §5.3.1. Semantic entailment characterizes one of the many possible forms of "fuzzy inference" (cf. §2.2.9). For example, it also follows from the foregoing that "very young" semantically entails "young."

§2.2.6. Grammars

The use of a grammar for generating a term set for a linguistic variable ensures that the semantic meanings of composite terms are explicitly definable in terms of the meanings of the smaller set of atomic terms. Further, the grammar approach is useful for excluding nonsense terms like "very and not young." As an example (cf. Zadeh, 1975b, Part II, p. 326) a term set $T(\chi)$ for the linguistic variable χ=Age may be generated by a grammar $G=(V_T, V_N, T, P)$, where (i) V_T is the set of atomic terms (or terminals) consisting of "young," "old," "not," "or," "and," "very," and some parentheses (and), (ii) V_N is a set of nonterminals consisting of the letters T, A, B, C, D, and E, in which case the letter T is a symbolic representative of the term set $T(\chi)$, and (iii) P is a production system given by

$T \rightarrow A$,	$C \rightarrow D$,
$T \rightarrow T$ or A,	$C \rightarrow E$,
$A \rightarrow B$,	$D \rightarrow$ very D,
$A \rightarrow A$ and B,	$E \rightarrow$ very E,
$B \rightarrow C$,	$D \rightarrow$ young,
$B \rightarrow$ not C,	$E \rightarrow$ old,
$C \rightarrow (T)$.	

A linguistic term in $T(\chi)$ is produced in G by any sequence of productions in P which starts with the nonterminal T and ends in an expression containing only terminals. For example, a production of the composite term "not very young and not very old" is

$T \rightarrow A \rightarrow A$ and $B \rightarrow B$ and $B \rightarrow$ not C and $B \rightarrow$ not D and $B \rightarrow$ not very D

and $B \rightarrow$ not very young and $B \rightarrow$ not very young and not $C \rightarrow$ not very

young and not E → not very young and not very E → not very young

and not very old.

A discussion of the use of grammars, together with the use of a "parsing tree" for analyzing complex expressions, is provided in the above reference, pp. 326-331. Note that a meaning for the term "not very young and not very old" may be defined explicitly in terms of the foregoing meanings of the atomic terms as the fuzzy set M given by

$$\mu_M(x) = \min[\mu_{-(very-Y)}(x), \mu_{-(very-O)}(x)]$$
$$= \min[1-\mu_{very-Y}(x), 1-\mu_{very-O}(x)]$$
$$= \min[1-[\mu_y(x)]^2, 1-[\mu_O(x)]^2].$$

§2.2.7. Linguistic Approximation

Suppose that the above grammar for the linguistic variable "Age" is extended to contain the term "middle aged" among its terminals, and suppose that one of the production rules is

not very young and not very old → middle aged.

Then the fuzzy set M defined in §2.2.1 would serve as a definition of I(middle aged), in which case the above production rule becomes an expression of semantic equivalence.

But now consider the term "not young and not old." The meaning of this term will be close to the meaning of "middle aged," although not exactly mathematically identical. Yet in an application of fuzzy linguistics to a discussion of human ages, it would likely cause no difficulty to just assume that the more complex composite term and the simpler atomic term are interchangeable.

This illustrates a problem which has received considerable attention in the literature: that of developing an algorithm for determining

when the meanings of two linguistic terms are sufficiently similar to regard those terms as pragmatically identical. This amounts to creating an algorithmic measure of "linguistic approximation" in terms of membership functions of fuzzy sets. Yager (1979) now appears to have resolved this problem in a simple and intuitively satisfying approach based on Zadeh's concept of a "level-set."

No attempt has been made in this dissertation to axiomatize the concept of linguistic approximation. Rather, composite terms are dealt with only as they are generated by a given grammar. Yet it is a reasonable conjecture that the results of this work can be extended to include a concept of linguistic approximation. This would involve "fuzzifying" the "outer linguistic level" (cf. Chapter 3), and might lead to a system with three linguistic levels rather than two.

§2.2.8. Multivalent Fuzzy Logic

Transition from the concept of a linguistic variable to a specific multivalent logic is accomplished by interpreting degrees of compatability as truth values. Example: if "young" is given as its meaning I(young) the foregoing fuzzy set Y, then the linguistic assertion "young(x)" is taken as expressing that "x is young" and is given as its truth value the degree of compatability $\mu_Y(x)$. Thus, the assertion "x is young" is "true" to the degree that x is a member of the fuzzy set Y.

This concept of "truth" is what forms the basis for the axiomatization of a theory of approximate reasoning as a multivalent logic in Chapter 3. Such systems of reasoning are to be distinguished from those for which Zadeh formally reserves the term fuzzy logic, however, which are characterized by having the truth values be fuzzy subsets of [0,1].

The concept of "fuzzy truth" is taken up in §2.2.9 and §2.2.10, and the manner in which the formalism of Chapter 3 may be modified to accommodate fuzzy logic per se is discussed in Chapter 5.

In the multivalent context, the foregoing fuzzy interpretations of the Boolean connectives "not," "or," and "and" can be adapted to serve as connectives between linguistic assertions. For example, if "p(x,y)" denotes the assertion "young(x) and old(y)", and if I(young)=Y and I(old)=O, then the truth values of p(x,y) are given as grades of membership of ordered pairs (x,y) in the fuzzy relation I(p) defined by

$$\mu_{I(p)}(x,y) = \min[\mu_Y(x),\mu_O(y)],$$

and if the connective was "or" rather than "and," one would use "max" in place of "min."

Next arises the question of logical inference--i.e., a fuzzy interpretation of the logical "implies." While the foregoing interpretations of the Boolean connectives are nowdays well-established (cf. Bellman and Giertz 1973) there is no widespread agreement on an appropriate form of fuzzy-logical implication. A general analysis of fuzzy inference and a list of several versions appears in Gaines (1976), and Zadeh's papers provide several other versions, including the idea of semantic entailment mentioned in §2.2.5.

A study of this topic is undertaken in this dissertation in §5.5. It may be noted in Chapter 3 that no specific form of fuzzy inference is incorporated into the class of formal linguistic theories defined therein. This leaves open the option of adjoining one of several different forms, as a given situation may demand.

§2.2.9. Types Of Fuzzy Sets, Fuzzy Logic, And Linguistic Truth

A fuzzy set of type 2 is characterized by a membership function whose values are fuzzy subsets of [0,1], rather than points in [0,1]. By contrast, fuzzy sets of the kind described in §2.1.1 are referred to as fuzzy sets of type 1. The concept of a type 2 fuzzy set has been developed into a fuzzy logic which uses fuzzy subsets of [0,1] as truth values (Bellman and Zadeh 1977) through an application of Zadeh's "fuzzification operator" to the connectives in the Lukasiewicz infinitary logic $L_{\aleph_1}$. In other works, particularly Zadeh (1975b), this form of "truth" has also been studied as a special kind of linguistic variable. The overall aim is to provide intuitively plausible interpretations for linguistic assertions like

"x is very young is not very true",

where the meaning of "true" is itself imprecise.

The word "Truth" becomes a linguistic variable by taking its term set to include expressions like "true," "not true," "very true," "false," "not very true and not very false," and so on, and by taking the universe of discourse to be the unit interval. In using these terms, however, it soon becomes evident that the foregoing interpretations of the Boolean connectives do not yield the ordinary intuitions about the meanings of composite linguistic truth values. For example, except for trivial interpretations of the terms "true" and "false," one does not have that the composite term "true and false" is even approximately equivalent to the term "false," which ordinary reasoning would require.

Because of this Zadeh has developed an alternative set of connectives, specifically for use with linguistic truth. The characteristics

of these connectives, and the manner in which they can be represented in a semantically complete formal logical system, is taken up in §5.3.

§2.2.10. Linguistic Possibility

Recent work by Zadeh has been concerned with reinterpreting the theory of fuzzy linguistics as a theory of "linguistic possibility" (eg. Zadeh 1978a and 1978b) which is an outgrowth of earlier work under the heading of "linguistic probability" (as in Zadeh 1975b, Part III). Here it is worth noting that the revised terminology is clearly more appropriate, since probabilistic logic--which uses joint probability as its interpretation of the logical "and"--is substantially different from the logic of fuzzy sets. A recent paper specifically addressing this point is Zadeh (1980). The sense in which the term "possibility" applies may be described as follows. Suppose that the linguistic term "small integer" is assigned as its meaning the following fuzzy subset of the nonnegative integers:

$$SI = 1/0 + 1/1 + 0.8/2 + 0.6/3 + 0.4/4 + 0.2/5,$$

and consider the linguistic assertion "x is a small integer." In the absence of any further information about x, this assertion may be interpreted as an expression of possibility that x is an integer in the interval from 0 to 5. Further, given the above meaning SI of "small integer," it makes sense to say that the possibility that x has a particular value v in [0,5] is computed as the degree of compatability of v with the term "small integer." Thus, given the above assertion and meaning assignment, the possibility that x=3, for example, would be 0.6. In this manner, the fuzzy set SI becomes reinterpreted as a possibility distribution for the variable x.

Zadeh has used this interpretation as the guiding intuition for developing a set of <u>translation</u> <u>rules</u> for transforming natural language expressions into an algorithmic language. Thus the theory of fuzzy sets is brought one step closer to machine intelligence applications in natural language processing and in question answering systems.

The manner in which some of these ideas are formalizable in the axiom systems of this dissertation is taken up in §5.4. In general, most of the ideas used in possibilistic reasoning are coextensive with concepts in the theory of fuzzy linguistics and linguistic truth.

§2.9.11. Fuzzy-Logical Quantifiers

The standard logical quantifiers, "for all" and "there exists," also have generalizations for use in the context of fuzzy sets. As well, fuzzy linguistics offers the opportunity to express the implicitly fuzzy quantifiers of natural discourse, like "most," "many," "few," "some," "not very many," "almost all," etc. The semantics of fuzzy quantifiers has been developed in several of Zadeh's papers, and it plays an important role in the use of linguistic possibility for natural language translation.

Fuzzy quantification theory is not considered in this dissertation because its introduction into the axiomatization problem would substantially increase the difficulty of establishing semantic completeness. Now that the completeness results of Chapter 4 have been developed, however, it appears that extension to an axiomatization of fuzzy quantification is certainly feasible, and would constitute a natural topic for follow-on research.

CHAPTER 3

LINGUISTIC THEORIES

This chapter defines a class of formal systems wherein each system
is a formal theory of some collection of linguistic variables over a
single universe of discourse. The first step toward this definition is
to develop a rigorously defined formal linguistics which captures the
idea of a collection of linguistic terms as being generated by one or
more grammars, together with the idea of semantic equivalence between
linguistic terms. Formalization of semantic equivalence leads naturally
to a further introduction of negations, disjunctions, etc. of proposi-
tions that express such equivalence relations. Thus we arrive at a
linguistics having two distinct "levels" as discussed in §1.1 and §2.2.8.
Explicit definition of the formal languages which comprise this linguis-
tics is given in §3.1.

In §3.2 we complete the formalization of the concept of a linguis-
tic variable by providing an explicitly defined semantics for the given
class of formal languages. This captures the manner in which fuzzy
subsets of a universe of discourse serve as the meanings of linguistic
terms, as well as the manner in which a truth value in [0,1] for a
linguistic assertion is computed from the meanings of the linguistic
terms which occur in that assertion. Further, this semantics incorpor-
ates the interpretation of semantic equivalence as expressing equality
between fuzzy sets. Thus we arrive at the required interpretation of
the inner linguistic level as being multivalent and of the outer linguistic

level as being bivalent: multivalency comes from having the membership

functions of fuzzy sets range over [0,1], while bivalency stems from the

fact that two linguistic terms either are, or are not, semantically

equivalent (cf. §2.2.8). Even though the definition of the semantics in

§3.2 logically follows the definition of the linguistics in §3.1, it was

the desire to formalize the ideas embodied in the semantics which served

as the intuitive guide for developing the linguistics into its given

form. The two in tandem formalize the full idea of a linguistic variable

as defined in §2.2.5.

The axioms and inference rules for a class of formal systems are

described in §3.3. The intuitive guide in this case is the desire to

develop enough proof-theoretic strength to ensure semantic completeness,

which is ultimately established in Chapter 4. The axioms for the outer

level are adaptations of the axioms for the classical propositional

calculus, while the axioms for the inner level are based on a standard

definition of a de Morgan lattice (cf. §4.3).

The formal systems themselves are defined in §3.4, as a class of

linguistic theories, or simply theories. (This usage of the word "theory"

is retained throughout the remainder of Part I, and should not be confused

with its more common denotation of a "first-order theory.") The idea of

a semantic model of a theory is also discussed in §3.4, and a few elementary

results are established--such as consistency of the "minimal theory"--

leading up to the work on the completeness theorems in Chapter 4.

§3.1. Languages

The development in this section follows the pattern described in

§2.1.1. As will be seen, "individual variables" are thought to range

over some (unspecified) universe of discourse, "individual constants"
represent arbitrary (fixed) members of that universe, and "fuzzy relation
symbols" correspond to atomic linguistic terms--eg., a binary relation
symbol would be used to express the "similar to" relation of §2.2.2, and
a unary relation symbol would be used to represent the term "young" of
§2.2.1.

Further, the idea of a linguistic term here becomes embodied in
the formal definition of a "linguistic assertion." Note that in the
informal theory of fuzzy linguistics, a composite linguistic assertion
of the form "$\alpha(x)$ and $\beta(x)$" may be written alternatively as "(α and
β)(x)"--i.e., using a composite linguistic term--but that an assertion
of the form "$\alpha(x)$ and $\beta(y)$" cannot be so replaced. Thus the idea of a
linguistic assertion leads to a somewhat more complex collection of
expressions than the collection of linguistic terms discussed in §2.2.8,
but it includes the original idea of a linguistic term in a proper
subcollection.

Last, the linguistic variable per se does not receive an explicit
representation in these languages; but a given linguistic variable is
nonetheless implicitly formalized through a rigorous representation of
that variable's collection of linguistic terms. The manner in which
further ideas from the informal theory become similarly represented will
be illustrated in some examples.

§3.1.1. Symbols

The following collections of <u>symbols</u> are assumed to be mutually
disjoint subcollections of the uncountably infinite collection of symbols
which was given in §2.1.1.

1. <u>Individual variables</u>: a countably infinite set of symbols, denoted metalinguistically by x,y,z, etc.

2. <u>Individual constants</u>: an uncountably infinite set of symbols, denoted metalinguistically by a,b,c, etc.. This set is assumed to include as a proper subset an uncountably infinite set of <u>individual names</u>, which may receive the alternate denotations $\underline{i},\underline{j}$, etc. (Individual names are individual constants that are used to "name" the specific individuals in the universe of a given semantic interpretation, cf. §3.2.1.)

3. <u>Equality symbol</u>: one symbol, denoted metalinguistically by =.

4. <u>Fuzzy relation symbols</u>: for each $n \geq 0$ an uncountably infinite set of n-ary relation symbols, denoted metalinguistically by α,β,γ, etc. The 0-ary relation symbols are assumed to include a special symbol, κ_r, for each number $r\varepsilon[0,1]$. (In most cases, κ_r "names" the unique fuzzy subset S of any universe U whose membership function has $\mu_S(x)=r$ for all $x\varepsilon U$, cf. §3.2.1.)

5. <u>Multivalent (or fuzzy) connectives</u>: three symbols denoted metalinguistically by $\sim$ (not; negation), $\lor$ (or; disjunction), and $\land$ (and; conjunction).

6. <u>Hedge (or special operator) symbols</u>: for each $n \geq 1$, an uncountably infinite set of symbols, denoted metalinguistically by ϕ, ψ, etc..

7. <u>Equivalence symbol</u>: one symbol, denoted metalinguistically by $\cong$.

8. <u>Punctuation marks</u>: three symbols, denoted metalinguistically by , (comma), and (and) (parentheses).

9. <u>Bivalent (or classical) connectives</u>: two symbols, denoted metalinguistically by $\neg$ (not; negation) and $\lor$ (or; disjunction).

Individual variables and individual constants (hence also individual names) are collectively called individual terms and will have the common notations t, t', etc.

The logical symbols are the individual variables, the equality symbol, the multivalent connectives, the equivalence symbol, the punctuation marks, and the bivalent connectives. These will be common to all languages as defined in this section. The other symbols are proper symbols, any of which might or might not be used in a specific language (cf. §3.1.3).

§3.1.2. Expressions

The linguistic assertions have the general denotations p,q,r, etc. and are defined to consist of:

1. Atomic linguistic assertions: expressions of the form $\alpha(t_1,\ldots,t_n)$, where α is an n-ary fuzzy relation symbol and $t_1,\ldots,t_n$ are individual terms (hence including expressions of the form α where α is 0-ary),

2. Composite linguistic assertions: all expressions which can be generated from the atomic linguistic assertions by means of a context-free grammar (§2.2.6), subject only to the requirement that every such assertion has one of the following four forms: $\sim p$, $(p \lor q)$, $(p \land q)$, and $\phi(p_1,\ldots,p_n)$, where $p,q,p_1,\ldots,p_n$ are linguistic assertions and ϕ is a hedge symbol.

The formulas have the general denotation, P,Q,R, etc. and are defined to consist of:

1. Atomic formulas: all equations of the form $(t=t')$, where t and t' are individual terms, and all equivalences of the form $(p \cong q)$, where p and q are linguistic assertions,

2. <u>Composite formulas</u>: all expressions of the forms ¬P and (P∨Q) where

P and Q are formulas.

Some abbreviations to be used are:

      P&Q for ¬(¬P∨¬Q)   (bivalent conjunction)

      P⊃Q for ¬P∨Q   (bivalent implication)

      P≡Q for (P⊃Q)&(Q⊃P)   (bivalent logical equivalence)

Parentheses are left unwritten when not needed for readability--e.g.,

(P∨Q) may be shortened to P∨Q. Note also that, as in the above defini-

tions of & and ⊃, the connective ¬ has a lower priority than any of the

others, so that ¬P∨Q means the same as (¬P)∨Q, rather than ¬(P∨Q). In

chains of implications we may assume association to the right, so that

$P_1⊃\cdots⊃P_n$ means the same as $(P_1⊃\cdots⊃(P_{n-1}⊃P_n)\ldots)$. Similar remarks

apply for composite linguistic assertions.

An expression is <u>closed</u> if it contains no occurrences of individual

variables; otherwise it is <u>open</u>. The notation $e(e_1,\ldots,e_n/x_1,\ldots,x_n)$,

where $e,e_1,\ldots,e_n$ are expressions, denotes the uniform substitution as

defined in §2.1.1. Sometimes the items following the "/" are omitted,

in which case $e(e_1,\ldots,e_n)$ indicates that $e_1,\ldots,e_n$ occur in e.

§3.1.3. Definitions

A <u>language</u> L is comprised of:

1. Symbols: the logical symbols, together with an empty or nonempty set

   of each kind of proper symbol,

2. Grammars: zero, one, or more context-free grammars which meet the

   requirement mentioned in §3.1.2,

3. Assertions: all atomic linguistic assertions which can be made from

   symbols of L, together with all composite assertions that can then

   be generated by the grammars of L,

4. Formulas: all atomic and composite formulas that can be made up of linguistic assertions of L.

The minimal language will be the (unique) language that contains no proper symbols and no grammars. Thus the minimal language contains no linguistic assertions, but it does contain formulas--e.g., there would be a formula of the form $\neg(x=y)$. A specific language of one or more linguistic variables is obtained from the minimal language by adjoining an appropriate set of proper symbols, together with one or more grammars. Typically, one grammar will suffice, but there is no technical difficulty with having several.

It shall be required of every language that the cardinality of its set of symbols be not greater than $\aleph_1$, the first uncountable cardinal number. This assumption is needed in Chapter 4.

§3.1.4 Example

To formalize the linguistic variable "Age" discussed in §2.2 in a formal language L_{Age}, let the linguistic terms "young" and "old" be represented by the unary fuzzy relation symbols α and β, and let the hedge "very" be represented by the unary hedge symbol ϕ. Then the assertion "x is very young" is represented by the formal linguistic assertion $\phi(\alpha(x))$. Composite linguistic assertions may be generated by essentially the same grammar as given in §2.2.6: (i) let the terminals be the atomic linguistic assertions of L, the multivalent connectives $\sim$, $\vee$, and $\wedge$, the hedge symbol ϕ, and the parentheses (and), (ii) let the nonterminals be the letters T, A, B, C, D, and E (as before) together with the letters $\underline{t}$ and $\underline{t}'$, (iii) rewrite the production system of §2.2.6 by everywhere replacing "young" and "old" with $\alpha(\underline{t})$ and $\beta(\underline{t}')$, replacing

"not", "or", and "and" with $\sim$, $\vee$, and $\wedge$, and replacing "very" with ϕ,

and for each individual term t of L add the pair of production rules.

$$\underline{t} \rightarrow t \text{ and } \underline{t}' \rightarrow t.$$

This leads to a language of the kind defined in §2.1.3. As an example

of a formula in this language:

$$\neg(\alpha(x) \cong \sim\beta(x)) \equiv \neg(\beta(x) \cong \sim\alpha(x))$$

expresses the proposition (which might or might not be true) that "young"

is not semantically equivalent with the antonym of "old" if and only if

"old" is not semantically equivalent with the antonym of "young."

The special 0-ary relation symbols κ_r may be used to express such

assertions as "the individual a is .7 young" in an equivalence of the

form

$$\alpha(a) \cong \kappa_{.7}.$$

This is based on the interpretations of the κ_r defined formally below.

§3.2. Semantic Interpretations

The semantics to be used in this chapter and the next shall consist

of all possible interpretations of the following kind for all languages

of the kind described in §3.1. This section follows the pattern of

§2.1.2.

§3.2.1. Definition

An interpretation I for a language L has:

1. A universe, U_I. For each individual in U_I, there is assumed to be a

   unique individual name which serves as the name of that particular

   individual. The notation "$\underset{\sim}{i}$" then denotes the name of the individual

   whose corresponding denotation is i. The language that is obtained

from L by adjoining all the names of the individuals in U_I is denoted by L(I).

2. <u>Meaning assignments</u>:

   a. To each individual constant a of L, assignment of an individual I(a) in U_I. For each individual name $\underset{\sim}{i}$ of L(I), it is always understood that $I(\underset{\sim}{i})=i$, i.e., that I assigns to each name the unique individual of which it is the name.

   b. To each n-ary fuzzy relation symbol α of L, assignment of an n-ary fuzzy relation $I(\alpha)$ in U_I. This is equivalent to specifying a membership function $\mu_{I(\alpha)}:U_I^n \to [0,1]$. In particular, for each special 0-ary fuzzy relation symbol κ_r (if indeed there are any among the proper symbols of L), one might define $I(\kappa_r)$ by $\mu_{I(\kappa_r)}(i)=r$ for all individuals $i \varepsilon U_I$; this is to be regarded as a typical assignment, however, and not mandatory for every interpretation I.

   c. To each n-ary hedge symbol ϕ in L, assignment of an n-ary operation $I(\phi)$ on fuzzy relations in U_I such that, if $\iota_1,\ldots,\iota_n$ are well-defined fuzzy relations in U_I, then $I(\phi)(\iota_1,\ldots,\iota_n)$ is a well-defined fuzzy relation in U_I. This is equivalent to specifying a function $F_{I(\phi)}$ such that

   $$\mu_{I(\phi)(\iota_1,\ldots,\iota_n)}(\ldots)=F_{I(\phi)}(\mu_{\iota_1}(\ldots),\ldots,\mu_{\iota_n}(\ldots)).$$

3. <u>Valuations</u>: all mappings I:{closed linguistic assertions and formulas of L(I)}$\to[0,1]$ such that

   a. $I(\alpha(t_1,\ldots,t_n)) = \mu_{I(\alpha)}(I(t_1),\ldots, I(t_n))$

   b. $I(\sim p) = 1-I(p)$

   c. $I(p \vee q) = \max[I(p),I(q)]$

   d. $I(p \wedge q) = \min[I(p),I(q)]$

e. $I(\phi(p_1,\ldots,p_n)) = F_{I(\phi)}(I(p_1),\ldots,I(p_n))$

f. $I(t=t') = \begin{cases} 1 \text{ if } I(t) = I(t') \\ 0 \text{ if not} \end{cases}$

g. $I(p \cong q) = \begin{cases} 1 \text{ if } I(p) = I(q) \\ 0 \text{ if not} \end{cases}$

h. $I(\neg P) = 1 - I(P)$

i. $I(P \lor Q) = \max[I(P), I(Q)]$.

An <u>I-instance</u> of an expression e of L(I) is a closed expression of the form $e(i_1,\ldots,i_n/x_1,\ldots,x_n)$ where $i_1,\ldots,i_n$ are individuals in U_I. A formula P of L(I) is <u>valid</u> <u>in</u> I (Notation: I⊨P) if $I(P')=1$ for every I-instance P' of P.

Note that an interpretation I for a language L is uniquely determined by specifying a universe U_I and assigning a specific meaning in that universe to each proper symbol of L.

§3.2.2. Remark

Throughout Chapters 3 and 4 we consider only the case that the relations being formalized in a particular language L are relations on the same universe of discourse (c.f. §2.2.5). The situation of multiple universes is easily accommodated, however, as shown in Chapter 5.

§3.2.3. Example

Continuing the example of §3.1.4, let U_I be the set of ages from 0 to 100; let $\mu_{I(\alpha)}$ and $\mu_{I(\beta)}$ be any membership functions μ_Y and μ_O such that "young" and "old" are antonyms (cf. §2.2.5) and let $I(\phi)$ be the "very" operator of §2.2.3, defined by $F_{I(\phi)}(I(p))=(I(p))^2$. This specifies one of the (infinitely many) possible semantic interpretations of L_{Age}. In this particular interpretation, μ_Y and μ_O will be related by $\mu_Y=1-\mu_O$,

so that all formulas of the forms $\alpha(t) \cong \sim\beta(t)$ and $\beta(t) \cong \sim\alpha(t)$ are valid in I--thus capturing the fact that in this case "young" is semantically equivalent with "not old" and "old" is semantically equivalent with "not young." It follows by the definition of "$\equiv$" that the formula $(\alpha(x) \cong \sim\beta(x)) \equiv (\beta(x) \cong \sim\alpha(x))$ is valid in I.

If the symbols a and $\kappa_{.7}$ are adjoined to L, and $I(\kappa_{.7})$ is defined by $\mu_{I(\kappa_{.7})}(i) = .7$, for all $i \varepsilon U_I$, then if a is such that $\mu_Y(I(a)) = .7$, we have that the formula $\alpha(a) \cong \kappa_{.7}$ is valid in I.

§3.2.4. Remark

The assignments and valuations of an interpretation I for a language L induce the assignment of a unique fuzzy relation in U_I to each open linguistic assertion of L, according to:

1. For $m \geq 0$, $\mu_{I(\alpha(a_1,\ldots,a_m,x_1,\ldots,x_n)}(i_1,\ldots,i_n)$ is a restriction of $\mu_{I(\alpha)}$, considered as a function of $j_1,\ldots,j_m,j_{m+1},\ldots,j_{m+n}$, to the variables $j_{m+1},\ldots,j_{m+n}$ (and hence is a uniquely defined membership function),

2. $\mu_{I(\sim p)} = 1 - \mu_{I(p)}$,

3. $\mu_{I(p \vee q)} = \max[\mu_{I(p)}, \mu_{I(q)}]$

4. $\mu_{I(p \wedge q)} = \min[\mu_{I(p)}, \mu_{I(q)}]$,

5. $\mu_{I(\phi(p_1,\ldots,p_n))} = F_{I(\phi)}(\mu_{I(p_1)},\ldots,\mu_{I(p_n)})$,

where it is understood that, for example, if $\mu_{I(p)}$ is a function of two variables i_1, i_2 and if $\mu_{I(q)}$ is a function of i_2, i_3, then $\mu_{I(p \vee q)}$ will be a function of the three variables i_1, i_2, i_3.

This establishes that every open linguistic assertion may be thought of as an n-ary linguistic term, for some number n.

§3.3. Axioms and Inference Rules

This section describes the basic proof-theoretic mechanism for the linguistic theories to be defined in §3.4.

§3.3.1. Logical Axioms And Logical Inference Rules

These shall consist of the following categories of formulas and mappings of formulas, where "formula" here refers to the entire class as defined in §3.1:

1. Hilbert and Ackermann's (1928) axiomatization of the propositional calculus: all formulas having the forms

   a. $(P \lor P) \supset P$

   b. $P \supset (P \lor Q)$

   c. $(P \lor Q) \supset (Q \lor P)$

   d. $(P \supset Q) \supset ((R \lor P) \supset (R \lor Q))$

   together with <u>modus ponens</u>: from P and $P \supset Q$ infer Q.

2. Axioms for equality of individuals: all formulas having the forms

   a. $x = x$

   b. $x_1 = y_1 \supset (x_2 = y_2 \supset (x_1 = x_2 \supset y_1 = y_2))$

   c. $x_1 = y_1 \supset (\cdots \supset (x_n = y_n \supset \alpha(x_1, \ldots, x_n) \cong \alpha(y_1, \ldots, y_n)) \ldots)$.

3. Axioms for equality of linguistic terms: all formulas having the forms

   a. $p \cong p$

   b. $p_1 \cong q_1 \supset (p_2 \cong q_2 \supset (p_1 \cong p_2 \supset q_1 \cong q_2))$

   c. $p \cong q \supset \sim p \cong \sim q$

   d. $p_1 \cong q_1 \supset (p_2 \cong q_2 \supset (p_1 \lor p_2 \cong q_1 \lor q_2))$

   e. $p_1 \cong q_1 \supset (p_2 \cong q_2 \supset (p_1 \land p_2 \cong q_1 \land q_2))$

   f. $p_1 \cong q_1 \supset (\cdots \supset (p_n \cong q_n \supset \phi(p_1, \ldots, p_n) \cong \phi(q_1, \ldots, q_n)) \ldots)$.

4. <u>Substitution rule</u>: from P infer $P(t_1,\ldots,t_n/x_1,\ldots,x_n)$.

5. Axioms for a de Morgan lattice (Rasiowa 1974, p. 38): all equations having the forms

   a. $p \vee q \cong q \vee p$ $p \wedge q \cong q \wedge p$ (commutativity)

   b. $p \vee (q \vee r) \cong (p \vee q) \vee r$ $p \wedge (q \wedge r) \cong (p \wedge q) \wedge r$ (associativity)

   c. $p \vee (q \wedge p) \cong q$ $p \wedge (p \vee q) \cong p$ (absorption)

   d. $p \vee (q \wedge r) \cong (p \vee q) \wedge (p \vee r)$ $p \wedge (q \vee r) \cong (p \wedge q) \vee (p \wedge r)$ (distributivity)

   e. $\sim\sim p \cong p$ (involution)

   f. $\sim(p \vee q) \cong \sim p \wedge \sim q$ $\sim(p \wedge q) \cong \sim p \vee \sim q$ (de Morgan's Laws)

6. Linear-ordering criterion: all formulas having the forms

$$(p \vee q \cong p) \vee (p \vee q \cong q) \quad \text{and} \quad (p \wedge q \cong p) \vee (p \wedge q \cong q).$$

The latter pair of axiom forms are so named because in the context of any semantic interpretation they assert that the truth values in [0,1] form a linearly ordered set. These axioms play a crucial role in the completeness proofs of Chapter 4.

§3.3.2. Inference Rules In General

In §2.1 was mentioned the idea of an inference rule being "validity preserving." We may now provide this idea with a precise definition. An inference rule will be <u>regular</u> if every instance $(H_1,\ldots,H_n,C)$ satisfies: for every interpretation I of any language L which contains $H_1,\ldots,H_n$, and C, if $I \vDash H_i$ for all $i=1,\ldots,n$, then $I \vDash C$.

It is easily verified that modus ponens and the substitution rule are regular. An important use of the concept of regularity is established in Proposition 3.4.2-3.

§3.4. Theories and Models

We are now in a position to define the desired class of formal logical systems. This follows the methods of §2.1.4.

§3.4.1. Linguistic Theories

A theory T of zero, one or more linguistic variables is comprised of:

1. A language, L(T): a formal language of the kind described in §3.1.1,

2. Axioms: (i) the set of all logical axioms (§3.3.1), that are formulas of L(T), together with (ii) an empty or nonempty set of specially chosen formulas of L(T), to serve as proper axioms of T,

3. Inference rules: (i) the logical inference rules--modus ponens and the substitution rule--together with (ii) an empty or nonempty set of specially chosen proper rules of T,

4. Theorems: the axioms of T, together with all formulas of L(T) that can be derived from these axioms by means of the inference rules of T.

It follows that a theory is specified by selecting a specific language L(T), and specific sets of proper axioms and proper rules. Proper axioms are normally used to express how certain proper symbols of L(T) are interrelated (see the example in §3.4.3) and proper rules might be used to express special modes of inference (cf. Chapter 5). The minimal theory will be the theory whose language is the minimal language, and which has no proper axioms or rules.

The criteria for a formula P of L(T) to have a proof in T (Notation: T⊢P) are given in §2.1.5.1.

Proposition 3.4.1-1. A formula P of L(T) is a theorem of T if and only if T⊢P.

Proof. By Proposition 2.1.5.1-1. ☐

It is well-known that a formula of L(T) becomes a formula of the propositional calculus if we treat atomic formulas as if they were propositional variables. This leads to the following.

Proposition 3.4.1-2 (Tautology Theorem). If a formula P of L(T) is tautology of the propositional calculus, then T⊢P. If P is a tautological consequence of $P_1,\ldots,P_n$, and if T⊢P_i, for all i=1,...,n, then T⊢P.

Proof. (Details are omitted; see Shoenfield 1967, p. 27). The first assertion follows from the fact that the axioms and inference rules of T include those described in §3.3.1-1. The second assertion is a corollary. ☐

A contradiction in T will be a formula of L(T) having the form P&¬P. Then consistency for T is as defined in §2.1.5.2.

It happens that, for any formulas P and Q of L(T), Q is a tautological consequence of P&¬P. By the tautology theorem, this implies that a theory T is consistent if and only if there is at least one formula Q of L(T) such that T⊬Q. Thus T is consistent if and only if it is coherent (c.f. §2.1.5.2). This means that the formal theories defined in this section, when viewed at the "outer level," are Aristotelean in the sense defined in §2.1.5.4.

§3.4.2. Semantic Models

The definition of a model of a theory T and of validity in Σ(T) is given by §2.1.6. We may reitterate Proposition 2.1.6-1 in this context.

Proposition 3.4.2-1 (Validity Theorem). If T⊢P, then Σ(T)⊨P.

Proof. Immediate, by the definition of model. □

This gives the following.

Proposition 3.4.2-2. A theory T is consistent if it has a model.

Proof. Suppose that T is inconsistent, say T⊢P&¬P. Since a formula of the form P&¬P cannot be valid in any interpretation of L(T) (a consequence of the definition of "&"), Proposition 3.4.2-1 implies that T has no models. Thus if T has a model, it must be consistent. □

Proposition 3.4.2-3. If the proper inference rules of a theory T are regular rules, then an interpretation I for L(T) is a model of T if and only if every proper axiom of T is valid in I.

Proof. Since axioms of T are theorems of T, it is obvious that, if I⊨T, then I⊨P for every proper axiom P. Suppose that an interpretation I is such that I⊨P for every proper axiom P. It is easily verified that I⊨P for every logical axiom of T. It has been noted that the logical rules are regular (§3.3.2). Thus, all axioms of T are valid in I, and all inference rules of T always act so that validity is preserved. It follows by the definition of "theorem" that if P is a theorem of T, then I⊨P. □

Proposition 3.4.2-4. The minimal theory is consistent.

Proof. Let T be the minimal theory. By Proposition 3.4.2-3, every interpretation I of L(T) is a model of T. Hence T is consistent by Proposition 3.4.2-2. □

§3.4.3. Example

A formal theory T_{Age} of the linguistic variable "Age" may be specified as follows. Let $L(T_{Age})$ be the language L_{Age} of §3.1.4.

Let T_{Age} have as its only proper axiom, the formula $\alpha(x)\cong\sim\beta(x)$. Then the interpretation I of §3.2.2 will be a model of T, by Proposition 3.4.2-3.

Since the squaring function, used to define the "very" operator, is not definable in terms of 1-, max, and min, there is no formula of $L(T_{Age})$ which explicitly defines ϕ as such. However, an approximation of the "very" operator can be introduced through proper axioms of the form

$$\neg(\phi(p)\cong p)\&(\phi(p)\vee p\cong p)$$

which asserts that the truth value of $\phi(p)$ is strictly less than the truth value of p. Since this property is satisfied by the squaring function on [0,1], it follows that the same interpretation I of $L(T_{Age})$ would be a model of the extended system that is obtained by adjoining all formulas of the above form as further proper axioms. This also shows, that the extended system may have different models, based on different meaning assignments for ϕ. Further, different models of T_{Age} may be obtained by altering the definitions of $I(\alpha)$ and $I(\beta)$, subject to the requirement that $\mu_{I(\alpha)}=1-\mu_{I(\beta)}$. Last, different models of T_{Age} may be obtained by moving to completely different universes of discourse. This means that the theory of a given linguistic variable may sometimes be used to capture the logic of that variable over several different universes.

As one adds further proper axioms, however, the class of models becomes more restricted. It is possible, for example, to require that the universes of all models have exactly 100 elements. This may be accomplished in the case of T_{Age} by adjoining some individual constants

$a_1, a_2, \ldots, a_{100}$ to $L(T_{Age})$, and adjoining as proper axioms a formula of the form

$$x=a_1 \vee x=a_2 \vee \cdots \vee x=a_{100}$$

together with all formulas

$$\neg(a_i=a_j) \qquad \text{for } i \neq j.$$

The models of this extension of T_{Age} would then differ essentially only in their variations on the interpretations of α, β, and ϕ.

One may furthermore delimit the relevant set of truth values by means of the 0-ary fuzzy relation symbols κ_r. For example, to limit the set to just five values $\{0.0, 0.25, 0.5, 0.75, 1.0\}$, include the symbols $\kappa_{0.0}$, $\kappa_{0.25}$, $\kappa_{0.5}$, $\kappa_{0.75}$, $\kappa_{1.0}$ as proper symbols of $L(T)$, and add as proper axioms the formulas

$$\kappa_{0.0} \vee \kappa_{0.25} \cong \kappa_{0.25}$$
$$\vdots$$
$$\kappa_{.75} \vee \kappa_{1.0} \cong \kappa_{1.0},$$

which rank order the five values, together with the formulas

$$\sim\kappa_{0.0} \cong \kappa_{1.0}$$
$$\sim\kappa_{0.25} \cong \kappa_{0.75}$$
$$\sim\kappa_{0.5} \cong \kappa_{0.5}$$

which describe their interrelations. Then for every n-ary relation symbol α of $L(T)$ add the formula

$$\alpha(x_1,\ldots,x_n) \cong \kappa_{0.0} \vee \cdots \vee \alpha(x_1,\ldots,x_n) \cong \kappa_{1.0}.$$

It will follow that every semantic model of T treats each κ_r as if it were the number r in $[0,1]$, and effectively uses only these five numbers as truth-values.

CHAPTER 4

SEMANTIC COMPLETENESS

The results established in this chapter show that the logical

axioms and inference rules of §3.3 exactly capture the logic which is

inherent in the multivalent interpretations of the fuzzy Boolean connec-

tives (§2.2.5 and §2.2.8) and the concept of semantic equivalence as

equality of fuzzy sets (§2.2.5). Following Shoenfield (1967, pp. 41ff),

semantic completeness may be expressed in two forms:

First Form: For any theory T and formula P of L(T), if $\Sigma(T) \models P$ then

$T \vdash P$.

Second Form: For any theory T, if T is consistent then T has a

model.

The former is the converse of the validity theorem (Proposition

3.4.2-1), and the latter is the converse of Proposition 3.4.2-2. Both

of these forms will be established here for linguistic theories, using

an adaptation of some algebraic methods which were developed by myself

and my master's thesis advisor, S. K. Thomason, for establishing semantic

completeness for a class of "free-variable theories" (Schwartz 1973).

This in turn is an adaptation of some methods of Rasiowa and Sikorsky

(1963).

As motivation for this approach, let us contrast it with the

standard approach for "first-order theories" as recorded in Shoenfield

(1967). Shoenfield's first step is to reduce the first form to the

second by means of a "Reduction Theorem for Consistency": If $T \nvdash P$, and if

P' is the "universal closure" of P, then the simple extension $T[\neg P']$ is consistent. Given this theorem, one then argues as follows. Suppose that $T\nvdash P$. Then $T[\neg P']$ has a model I by the second form of the completeness theorem. Because P' is an axiom of $T[\neg P']$, P' is valid in I. It follows that P is not valid in I. But I is also a model of T. So P is not valid in some model of T. Hence $\Sigma(T)\nvDash P$. Thus, if $\Sigma(T)\vDash P$, then $T\vdash P$.

To establish the second form, Shoenfield employs the methods of Henkin, which involves four preliminary results: (i) every first-order language L has a "cannonical interpretation" I which is built out of the syntactic elements of L by defining the individuals U_I to be equivalence classes of individual constants of L (assuming that L indeed contains constants), (ii) for every theory T there is an inessential extension (in Shoenfield, a "conservative" extension) T_c which contains a special constant for each closed instantiation $\exists x P$ in L(T), called a "Henkin Theory" (cf. Shoenfield for a precise definition), (iii) if a Henkin Theory T is proof-theoretically complete--i.e., for every closed P, either $T\vdash P$ or $T\vdash \neg P$--then the cannonical interpretation for L(T) is a model of T, (iv) "Lindenbaum's Theorem," every consistent theory has a syntactically complete simple extension.

Given these results, one proves the second form of the completeness theorem as follows. Suppose that T is consistent. Let T_c be the inessention extension of T given by (ii). T_c is consistent because the extension is inessential. Let T' be a syntactically complete simple extension of T_c, as given by (iv). T' happens to be a Henkin theory. So, by (iii), the canonical interpretation I of L(T') is a model of T' (this I exists, i.e., has a nonempty universe, because Henkin theories

contain individual constants). It follows that the restriction of I to L(T) is a model of T. Thus, if T is consistent, then it has a model.

Now the free-variable theories developed in Schwartz (1974) bear a resemblance to linguistic theories as studied in this dissertation in that they do not contain quantifiers and they admit proper inference rules (which happen to not be of any use in first-order systems). It turns out that, because of the use of proper rules, the above approach does not apply for such systems. For without prior explicit knowledge of these rules, one has no assurance that individual constants can be introduced into the language of T without upsetting the consistency of T. This difficulty is fatal to several different steps in the foregoing proof. For example, it happens that a universe U_I can be built out of equivalence classes of individual variables (rather than constants), but "new" constants are still needed to serve as surrogate existential quantifiers in establishing the appropriate analogues of the reduction theorem for consistency and of Lindenbaum's theorem.

These considerations suggest two possible alternate approaches: (i) place sufficiently many restrictions on proper inference rules to permit the introduction of new individual constants, or (ii) use an algebraic approach which altogether bypasses the need for such constants. The latter approach was chosen in Schwartz (1974) and has turned out to serve well also in this dissertation for dealing with the considerably more complex situation of two distinct linguistic levels together with a multivalent interpretation.

The general strategy of this method may be outlined as follows. We appropriate the theory of Boolean algebras (§4.1) and consider the "Lindenbaum-Tarski algebra" Γ_T of equivalence classes of formulas in the

language of a given theory T (§4.2). If T is consistent, then the equivalence classes of theorems of T are contained in one or more "ultra-filters" Δ in Γ_T. Any such Δ can be used to define a universe which uses equivalence classes of individual terms (i.e., variables or constants) as individuals (§4.3). There are furthermore definable equivalence classes of linguistic assertions, which turn out to form a separate algebra, having the form of a de Morgan lattice (§4.3). Then—the main step—one can define a "cannonical embedding" Φ_Δ which embeds this lattice in [0,1] in such a way that the lattice ordering is preserved (§4.3). This leads to the definition of the cannonical interpretations I_Δ for the language L(T) of any consistent T (§4.4). Some basic properties of the I_Δ are established in §4.5. Then the two completeness results are established in §4.6 as follows: (i) the second form is proved by showing that any cannonical interpretation of a consistent T is a model of T, (ii) the first form is proved by using algebraic principles to show that if T⊬P, then Δ can be chosen in such a way that I_Δ is a model of T in which P is not valid, from which it follows that T⊢P if P is valid in every model of T.

It is noteworthy that in (ii) the reduction theorem for consistency is absorbed by the use of Boolean algebra "Fact 4" (§4.1), and that in (i) Lindenbaum's Theorem is absorbed in the use of an ultrafilter, which is provided by Boolean "Fact 5." This exhibits the algebraic approach used here as having a versatility that is not shared by the standard syntactical proof methods.

§4.1. Boolean Algebras

The contents of §4.1.1 and §4.1.2 are adapted from Rasiowa and Sikorski (1963). All page numbers indicate places in that book where the relevant items are discussed.

§4.1.1. Definitions

A __Boolean algebra__ (p. 68) is a set Γ that is closed with respect to two binary operations (Notations: $\cup$ and $\cap$) and one unary operation (Notation: -) which, for all $\alpha,\beta,\gamma\epsilon\Gamma$, satisfy the following:

1. $\alpha\cup\beta = \beta\cup\alpha$ $\qquad\qquad$ $\alpha\cap\beta = \beta\cap\alpha$

2. $\alpha\cup(\beta\cup\gamma) = (\alpha\cup\beta)\cup\gamma$ $\qquad$ $\alpha\cap(\beta\cap\alpha) = (\alpha\cap\beta)\cap\gamma$

3. $(\alpha\cup\beta)\cap\beta = \beta$ $\qquad\qquad$ $(\alpha\cap\beta)\cup\beta = \beta$

4. $\alpha\cap(\beta\cup\gamma) = (\alpha\cap\beta)\cup(\alpha\cap\gamma)$ $\qquad$ $\alpha\cup(\beta\cap\gamma) = (\alpha\cup\beta)\cap(\alpha\cup\gamma)$

5. $(\alpha\cup-\alpha)\cap\beta = \beta$ $\qquad\qquad$ $(\alpha\cap-\alpha)\cup\beta = \beta$

Let Γ be a Boolean algebra. A non-empty subset Δ of Γ is a __filter__ (p. 44, bottom) in Γ if, for all $\alpha,\beta\epsilon\Gamma$,

1. $\alpha,\beta\epsilon\Delta$ implies $\alpha\cap\beta\epsilon\Delta$, and

2. $\alpha\epsilon\Delta$ implies $\alpha\cup\beta\epsilon\Delta$.

A filter Δ in Γ is __proper__ in Γ if $\Delta\neq\Gamma$. A proper filter Δ in Γ is an __ultrafilter__ in Γ if, for every proper filter Δ' in Γ, if $\Delta\subset\Delta'$, then $\Delta=\Delta'$ (cf. "maximal" filter, p. 46).

§4.1.2. Basic Facts

Let Γ be a Boolean algebra. The following facts are cited without proof from Rasiowa and Sikorsky (1963), except Fact 10, which is derived.

1. Γ contains a __unit element__ (Notation: $\underset{\sim}{1}$) defined by $\alpha\cup\underset{\sim}{1}=\underset{\sim}{1}$ and $\alpha\cap\underset{\sim}{1}=\alpha$,

   for all $\alpha\epsilon\Gamma$, and a __zero element__ (Notation: $\underset{\sim}{0}$) defined by $\alpha\cap\underset{\sim}{0}=\underset{\sim}{0}$ and

$\alpha \cup \underset{\sim}{0}=\alpha$ for all $\alpha \varepsilon \Gamma$. It follows that $\underset{\sim}{1}=\alpha \cup -\alpha$ and $\underset{\sim}{0}=\alpha \cap -\alpha$, for all $\alpha \varepsilon \Gamma$. (pp. 37 and 68).

2. Every subset Δ of Γ <u>generates</u> a filter in Γ, namely, the intersection of all filters Δ' in Γ such that $\Delta \subset \Delta'$ (p. 45).

3. A filter Δ is proper in Γ iff $\underset{\sim}{0} \notin \Delta$ (p. 46).

4. If Δ is a proper filter in Γ and $\alpha \notin \Delta$, then the filter generated by the set $\Delta \cup \{-\alpha\}$ is a proper filter in Γ (p. 79).

5. Every proper filter is contained in an ultrafilter, i.e., if Δ is a proper filter in Γ, then there exists an ultrafilter Δ' in Γ such that $\Delta \subset \Delta'$ (p. 46).

6. If Δ is an ultrafilter in Γ, then, for all $\alpha \varepsilon \Gamma$, either $\alpha \varepsilon \Delta$ or $-\alpha \varepsilon \Delta$ and not both (pp. 66-67).

7. Let $\alpha => \beta$ denote the <u>complement of</u> α <u>relative to</u> β, defined by $\alpha => \beta = -\alpha \cup \beta$. Then a subset Δ of Γ is a filter in Γ if and only if: $\underset{\sim}{1} \varepsilon \Delta$ and, for all $\alpha, \beta \varepsilon \Gamma$, if $\alpha \varepsilon \Delta$ and $\alpha => \beta \varepsilon \Delta$, then $\beta \varepsilon \Delta$ (pp. 54 and 56).

8. A nonempty subset Δ of Γ is a filter in Γ if and only if, for all $\alpha, \beta \varepsilon \Gamma$, we have $\alpha, \beta \varepsilon \Delta$ iff $\alpha \cap \beta \varepsilon \Delta$ (p. 44).

9. De Morgan's laws: for all $\alpha, \beta \varepsilon \Gamma$, $-(\alpha \cup \beta)=-\alpha \cap -\beta$ and $-(\alpha \cap \beta)=-\alpha \cup -\beta$ (p. 69).

10. If Δ is an ultrafilter in Γ, then, for all $\alpha, \beta \varepsilon \Gamma$, $\alpha \cup \beta \varepsilon \Delta$ iff either $\alpha \varepsilon \Delta$ or $\beta \varepsilon \Delta$.

Fact 10 may be established as follows. If either $\alpha \varepsilon \Delta$ or $\beta \varepsilon \Delta$, then $\alpha \cup \beta \varepsilon \Delta$ by the definition of filter (§3.1.1). Thus, it remains to show that $\alpha \cup \beta \varepsilon \Delta$ implies that either $\alpha \varepsilon \Delta$ or $\beta \varepsilon \Delta$. Suppose $\alpha \notin \Delta$ and $\beta \notin \Delta$. Then, since Δ is an ultrafilter, $-\alpha \varepsilon \Delta$ and $-\beta \varepsilon \Delta$ by Fact 6. It follows by Fact 8 that $-\alpha \cap -\beta \varepsilon \Delta$. Then $-(\alpha \cup \beta) \varepsilon \Delta$ by Fact 9. Hence $\alpha \cup \beta \notin \Delta$ by Fact 6. This completes the proof, by contradiction.

§4.2. Lindenbaum-Tarski Algebras

Let T be any theory; let P and Q be any formulas of L(T). Define:
P∿Q iff T⊢P≡Q.

Proposition 4.2-1. "P∿Q" defines an equivalence relation on the set of formulas of L(T).

Proof. (Details omitted.) By the tautology theorem, all three requirements for an equivalence relation are satisfied; i.e., (i) P∿P, (ii), if P∿Q, then Q∿P, and (iii) if P∿Q and Q∿R, then P∿R. □

Define: $[P]=\{Q|P∿Q\}$, $\Gamma_T=\{[P]|P$ is a formula of L(T)$\}$, and $\Delta_T=\{[P]|T⊢P\}$.

Proposition 4.2-2. Γ_T is a Boolean algebra with respect to the operations ∪, ∩, and − defined by

$$[P]∪[Q] = [P∨Q]$$

$$[P]∩[Q] = [P\&Q]$$

$$-[P] = [¬P].$$

Proof. By the tautology theorem, all the equations 1 through 5 sited in §4.1.1 are satisfied. For example, P∨Q≡Q∨P is a tautology; so T⊢P∨Q≡Q∨P. Then P∨Q∿Q∨P; so [P∨Q]=[Q∨P]. Hence [P]∪[Q]=[P]∪[Q]. This verifies equation 1. □

Proposition 4.2-3. The set Δ_T is a filter in Γ_T.

Proof. This makes use of Boolean algebra Fact 7 (§4.1.2). (a) By the tautology theorem, T⊢P∨¬P; so $\underset{\sim}{1}=[P]∪-[P]ε\Delta_T$. (b) Observe that [P⊃Q]=[P]=>[Q], by the definition of "⊃" and "=>"; it follows by modus ponens that, if [P]εΔ_T and if [P]=>[Q]εΔ_T, then [Q]εΔ_T. □

Proposition 4.2-4. The set Δ_T is a proper filter in Γ_T if and only if T is consistent.

Proof. The set Δ_T is a filter in Γ_T by Proposition 4.2-3. By Boolean algebra Fact 3, Δ_T is proper in Γ_T if and only if $\underset{\sim}{0} \notin \Delta_T$; and $\underset{\sim}{0} = [P] \wedge -[P] = [P \& \neg P] \notin \Delta_T$ iff T is consistent. □

The Boolean algebra Γ_T is commonly referred to as the Lindenbaum-Tarski algebra for T. Rasiowa and Sikorski (1963, pp. 20 ff) uses ideas similar to the above in application to several different kinds of systems. A more recent treatment is Rasiowa (1974).

§4.3. Canonical Embeddings

Throughout this section, let T be any consistent theory; let [P], Γ_T, and Δ_T be as in §4.2; in accordance with Proposition 4.2-4 and Boolean algebra Fact 5, let Δ be any ultrafilter in Γ_T such that $\Delta_T \subseteq \Delta$. Define: $t \sim t'$ iff $[t=t'] \varepsilon \Delta$, where t and t' are any individual terms of L(T); let $[t] = \{t' | t' \sim t\}$

Proposition 4.3-1. "$t \sim t'$" defines an equivalence relation on the set of individual terms of L(T).

Proof. This makes use of the axioms for equality of individuals and the substitution rule. We consider the three criteria of an equivalence relation as follows.

1. Pick any axiom of the form x=x (§3.2.1-2.a). Then T⊢t=t by the substitution rule; so $[t=t] \varepsilon \Delta_T$. Hence $[t=t] \varepsilon \Delta$; so $t \sim t$.

2. The formula $t=t' \supset (t=t \supset (t=t \supset t'=t))$ is a substitution instance of an axiom of the form 2.b (§3.2.1). Hence we have that $([t=t'] \Rightarrow ([t=t] \rightarrow ([t=t] \Rightarrow [t'=t]))) \varepsilon \Delta_T \subseteq \Delta$. We also have that $[t=t] \varepsilon \Delta$, as demonstrated in (1). Thus if $[t=t'] \varepsilon \Delta$, then, by three applications of Boolean algebra Fact 7, we have that $[t'=t] \varepsilon \Delta$. Therefore, if $t \sim t'$, then $t' \sim t$.

3. Similar to (2), using the formula t=t⊃(t'=t"⊃(t=t'⊃t=t")) to show

that, if t∿t' and t'∿t", then t∿t". □

Define: p∿q iff [p≅q]εΔ, where p and q are any linguistic assertions

of L(T); let [p]={q|q∿p}.

Proposition 4.3-2. "p∿q" defines an equivalence relation on the

set of linguistic assertions of L(T).

Proof. Similar to Proposition 4.3-1, using axioms for equality of

linguistic terms (§3.3.1-3), forms a and b. □

Define: χ={[p]|p is a linguistic assertion of L(T)}.

Proposition 4.3-3. The set χ is a de Morgan lattice under the

operations ∪, ∩, and -, defined by

$$[p]\textbf{u}[q] = [p\vee q]$$

$$[p]\cap[q] = [p\wedge q]$$

$$-[p] = [\sim p].$$

Proof. By the axioms listed in §3.3.1-5. □

Since χ is a lattice, the relation ≦ defined by

$$[p]\leqq[q] \text{ iff } [p]\textbf{u}[q]=[q]$$

is a partial ordering of χ (Rasiowa, p. 39).

Proposition 4.3-4. "[p]≦[q]" is a linear ordering of χ.

Proof. It is required to show that, for all [p] and [q], either

[p]≦[q] or [q]≦[p]. Suppose that [p]≰[q]. Then [p]∪[q]≠[q], by defini-

tion of "≦"; so [p∨q≅q]∉Δ. Then, since Δ is an ultrafilter, -[p∨q≅q]εΔ,

by Boolean algebra Fact 6. By the linear-ordering axioms (§3.3.1-6) and

the tautology theorem, T⊢⌐(p∨q≅q)⊃(p∨q≅p). It follows that

-[p∨q≅q]=>[p∨q≅p]εΔ; so an application of Boolean algebra Fact 7 yields

that [p∨q≅p]εΔ. This implies that [p]∪[q]=[p]; whence [q]∪[p]=[p],

since χ is a lattice (axiom form 5.a). Hence [q]≦[p].

Similarly, if $[q] \nleq [p]$, then $[p] \leq [q]$. Thus, necessarily, either $[p] \leq [q]$ or $[q] \leq [p]$. □

Now let $<$ be the strict ordering of χ defined by

$$[p] < [q] \text{ iff } [p] \leq [q] \text{ and } [p] \neq [q].$$

Then, for each $[p]$, exactly one of the following must hold:

1. $-[p] < [p]$, or

2. $-[p] = [p]$, or

3. $-[p] > [p]$.

Let χ_1, χ_2, χ_3, respectively, be the sets of $[p]$ that satisfy 1, 2, and 3. Clearly, $\chi = \chi_1 \cup \chi_2 \cup \chi_3$ (where here $\cup$ is ordinary set union). Define: a mapping $\Phi_\Delta : \chi \to [0,1]$ as follows:

1. $\Phi_\Delta : \chi_1 \to (\frac{1}{2}, 1]$ is any $<$-preserving embedding; i.e., if $[p] < [q]$, then

   $\Phi_\Delta([p]) <_r \Phi_\Delta([q])$, where $<_r$ is the usual strict ordering of the reals,

2. $\Phi_\Delta : \chi_2 \to \{\frac{1}{2}\}$,

3. $\Phi_\Delta : \chi_3 \to [0, \frac{1}{2})$ is defined by $\Phi_\Delta([p]) = 1 - \Phi_\Delta([\sim p])$, where $[\sim p] \varepsilon \chi_1$.

This mapping Φ_Δ will be the _canonical embedding based on_ Δ.

Proposition 4.3-5. The mapping Φ_Δ is well-defined.

Proof. Well-definedness requires that, if $[p] = [q]$, then $\Phi_\Delta([p]) = \Phi_\Delta([q])$. Consider $\Phi_\Delta(\chi_1)$. The set χ_1 has cardinality $\leq \aleph_1$, the first uncountable cardinal, since we have required the cardinality of $L(T)$ to be $\leq \aleph_1$ in §2.1.3. Observe that χ_1 is linearly ordered by $<$. It is easily established that any linearly ordered set having cardinality less than or equal to $\aleph_1$ can be embedded in any subinterval of the reals. Thus, a well-defined mapping of the kind $\Phi(\chi_1)$ exists.

It is obvious that $\Phi_\Delta(\chi_2)$ is well-defined. Consider $\Phi_\Delta(\chi_3)$. If $[p] \varepsilon \chi_3$, then $[p] < -[p]$, by definition of χ_3. But $[p] = --[p]$, because χ is a de Morgan lattice; and $-[p] = [\sim p]$, by definition of $-$. Hence $-[\sim p] < [\sim p]$;

so $[\sim p]\varepsilon\chi_1$. It follows that $\Phi_\Delta(\chi_3)$ is well-defined because $\Phi_\Delta(\chi_1)$ is well-defined. □

Proposition 4.3-6. The mapping Φ_Δ is a $<$-preserving embedding of χ in $[0,1]$.

Proof. Consider each possible way of choosing $[p]$ and $[q]$ from χ_1, χ_2, χ_3. For example, if $[p]$ and $[q]$ are both in χ_3, and $[p]<[q]$, then $[\sim p] = -[p] > -[q] = [\sim q]\varepsilon\chi$; so that $\Phi_\Delta([\sim p])>\Phi_\Delta([\sim q])$, which gives $\Phi_\Delta([p]) = 1-\Phi_\Delta([\sim p]) < 1-\Phi_\Delta([\sim q]) = \Phi_\Delta([q])$. □

Proposition 4.3-7. For all p and q, $\Phi_\Delta([p])=\Phi_\Delta([q])$ iff $[p]=[q]$.

Proof. If $[p]=[q]$, then $\Phi_\Delta([p])=\Phi_\Delta([q])$, by Proposition 4.3-5. Thus, it remains to show that Φ_Δ is one-to-one. Suppose that $[p]\neq[q]$. By Proposition 4.3-4, either $[p]\leq[q]$ or $[q]\leq[p]$. Thus, by the definition of $<$, either $[p]<[q]$ or $[q]<[p]$. But, in either case, $\Phi_\Delta([p])\neq\Phi_\Delta([q])$, by Proposition 4.3-6. Therefore, if $\Phi_\Delta([p])=\Phi_\Delta([q])$, then $[p]=[q]$. □

§4.4. Canonical Interpretations

Let T be a consistent theory; let Γ_T be the Lindenbaum-Tarski algebra for T; let Δ be any ultrafilter in Γ_T, such that $\Delta_T\subseteq\Delta$; let Φ_Δ be the canonical embedding based on Δ, as described in §4.3.

The _canonical interpretation_ for L(T) _based on_ Δ (Notation: I_Δ) may be defined in accordance with §4.2:

1. $U_{I_\Delta}=\{[t]|t$ is an individual term of L(T)$\}$, where $[t]$ is defined in terms of Δ as in §4.3,

2. For each individual constant a of L(T), assign $I_\Delta(a)=[a]$,

3. for each n-ary fuzzy relation symbol α of L(T), define $I_\Delta(\alpha)$ by

$$\mu_{I_\Delta}(\alpha)([t_1],\ldots,[t_n])=\Phi_\Delta([\alpha(t_1',\ldots,t_n')]),$$

where each t_i' is any representative from the equivalence class $[t_i]$,

4. for each n-ary hedge symbol ϕ, define $I_\Delta(\phi)$ by

$$F_{I_\Delta}(\phi)(r_1,\ldots,r_n) = \begin{cases} \phi_\Delta([\phi(p_1,\ldots,p_n)]), \text{ where } p_1,\ldots,p_n \\ \quad \text{are linguistic assertions of } L(T) \text{ such that} \\ \quad \phi_\Delta([p_i])=r_i, \text{ for all } i=1,\ldots,n, \\ \quad \text{assuming such } p_i \text{ exist,} \\ 0 \text{ if, for some } i, \text{ no such } p_i \text{ exists.} \end{cases}$$

Notation: If some individuals in U_{I_Δ} are denoted by $[t]$, $[a]$, and $[x]$, then their respective individual names in $L(T)(I_\Delta)$ will be denoted by $\underset{\sim}{t}$, $\underset{\sim}{a}$ and $\underset{\sim}{x}$.

<u>Proposition 4.4-1</u>. The universe U_{I_Δ} is well-defined.

<u>Proof</u>. The equivalence class $[t]$ is well-defined, by Proposition 4.3-1. □

<u>Proposition 4.4-2</u>. For each individual constant a of $L(T)$, $I_\Delta(a)$ is well-defined.

<u>Proof</u>. The equivalence class $[a]$ is well-defined, by Proposition 4.3-1. □

<u>Proposition 4.4-3</u>. For each fuzzy relation symbol α, $I_\Delta(\alpha)$ is well-defined.

<u>Proof</u>. It is required to show that, if $[t_1]=[t_1'],\ldots,[t_n]=[t_n']$, then $\mu_{I_\Delta}(\alpha)([t_1],\ldots,[t_n])=\mu_{I_\Delta}(\alpha)([t_1'],\ldots,[t_n'])$. If $[t_i]=[t_i']$, then $t_i \sim t_i'$, which means that $[t_i=t_i']\varepsilon\Delta$. Observe that the formula $t_1=t_1' \supset (\cdots \supset (t_n=t_n' \supset \alpha(t_1,\ldots,t_n)\cong\alpha(t_1',\ldots,t_n'))\ldots)$ may be inferred from an axiom of the form 2.c, by means of the substitution rule. It follows that $[t_1=t_1']=>(\cdots=>([t_n=t_n']=>[\alpha(t_1,\ldots,t_n)\cong\alpha(t_1',\ldots,t_n')])\ldots)\varepsilon\Delta$. Then $[\alpha(t_1,\ldots,t_n)\cong\alpha(t_1',\ldots,t_n')]\varepsilon\Delta$ by Boolean algebra Fact 7, so $[\alpha(t_1,\ldots,t_n)]=[\alpha(t_1',\ldots,t_n')]$. Then $\phi_\Delta([\alpha(t_1,\ldots,t_n)])= \phi_\Delta([\alpha(t_1',\ldots,t_n')])$, by Proposition 4.3.7. Hence $\mu_{I_\Delta(\alpha)}([t_1],\ldots,[t_n])=$

$\mu_{I_\Delta(\alpha)}([t_1'],\ldots,[t_n'])$, by definition of $\mu_I(\alpha)$ (§4.4-3). □

Proposition 4.4-4. For each n-ary special operator symbol ϕ, $I_\Delta(\phi)$ is well-defined.

Proof. Suppose that $r_1=r_1',\ldots,r_n=r_n'$, with $r_i,r_i'\varepsilon[0,1]$, and assume that there exist p_i such that $\Phi_i([p_i])=r_i$, for all $i=1,\ldots,n$. (If no such p_i exists for some i, then there is nothing to prove; $F_{I_\Delta}(\phi)(r_1,\ldots,r_n)=0=F_{I_\Delta}(\phi)(r_1',\ldots,r_n')$, by definition of $F_{I_\Delta}(\phi)$.) Suppose that $p_i',\ldots,p_n'$ are such that $\Phi_\Delta([p_i'])=r_i'$ for all $i=1,\ldots,n$. Then $\Phi_\Delta([p_i])=\Phi_\Delta([p_i'])$; so $[p_i]=[p_i']$, by Proposition 4.3-7. Hence $[p_i\cong p_i']\varepsilon\Delta$, for all $i=1,\ldots,n$. Observe that the formula $p_1\cong p_1'\supset$ $(\cdots\supset(p_n\cong p_n'\supset\phi(p_1,\ldots,p_n)\cong\phi(p_1',\ldots,p_n'))\ldots)$ is an axiom of the form 3.f. It follows that $[p_1\cong p_1']=>(\cdots=>([p_n\cong p_n']=>[\phi(p_1,\ldots,p_n)$ $\cong\phi(p_1,\ldots,p_n)])\ldots)$ is in Δ. Then n applications of Boolean algebra Fact 7 yields that $[\phi(p_1,\ldots,p_n)\cong\phi(p_1',\ldots,p_n')]\varepsilon\Delta$. Then $[\phi(p_1,\ldots,p_n)]=$ $[\phi(p_q,\ldots,p_n')]$; so $\Phi_\Delta([\phi(p_1,\ldots,p_n')])=\Phi_\Delta([\phi(p_1',\ldots,p_n')])$, by Proposition 4.3-7. Hence $F_I(\phi)(r_1,\ldots,r_n)=F_I(\phi)(r_1',\ldots,r_n')$. This is what we needed to show. □

Note that if T is not consistent, then no such interpretation I_Δ can exist. For, if T is not consistent, $\Delta_T=\Gamma_T$, so that there is no ultrafilter Δ in Γ_T such that $\Delta_T\subseteq\Delta$ (recall that an ultrafilter must be a proper filter); whence there is no canonical embedding Φ_Δ as required in the definition of I_Δ.

§4.5. Preliminary Results

Proposition 4.5-1. Let T be a consistent theory, and let I_Δ be any canonical interpretation for L(T). For any linguistic assertions p and q of L(T),

1. $\Phi_\Delta([\sim p]) = 1-\Phi_\Delta([p])$,

2. $\Phi_\Delta([p \vee q]) = \max[(\Phi_\Delta([p]), \Phi_\Delta([q]))]$,

3. $\Phi_\Delta([p \wedge q]) = \min[(\Phi_\Delta([p]), \Phi_\Delta([q]))]$,

where Φ_Δ is the canonical embedding associated with I_Δ.

Proof. Let p and q be any linguistic assertions of L(T). Let χ be as in Proposition 4.3-6.

1. To show that $\Phi_\Delta([\sim p])=1-\Phi_\Delta([p])$, consider three cases.

Case a: $[\sim p]\varepsilon\chi_1$. Then $-[\sim p]<[\sim p]$, which means that $[p]<-[p]$. Then $[p]\varepsilon\chi_3$, and $\Phi_\Delta([p])=1-\Phi_\Delta([\sim p])$, by definition of $\Phi_\Delta(\chi_3)$. Solving for $\Phi_\Delta([\sim p])$ gives the desired equation.

Case b: $[\sim p]\varepsilon\chi_2$. Then $-[\sim p]=[\sim p]$, and $\Phi_\Delta(-[\sim p])=\Phi_\Delta([\sim p])=\frac{1}{2}$. But $[p]=-[\sim p]$, so $\Phi_\Delta([p])=\frac{1}{2}$. This yields $\Phi_\Delta([\sim p])=\frac{1}{2}=1-\frac{1}{2}=1-\Phi_\Delta([p])$.

Case c: $[\sim p]\varepsilon\chi_3$. Then $\Phi_\Delta([\sim p])=1-\Phi_\Delta([\sim\sim p])$, by definition of $\Phi_\Delta(\chi_3)$. But $[\sim\sim p]=[p]$; so $\Phi_\Delta([\sim p]=1-\Phi_\Delta([p])$.

2. From the fact that χ is a lattice, we have that, for any [p] and [q],

$$[p]\leqq[p]\cup[q] \text{ and } [q]\leqq[p]\cup[q] \qquad (i)$$

(cf. Rasiowa, p. 39). From the fact that "$\leqq$" is a linear ordering (Proposition 3.3-4), we have that either

$$[p]=[p]\cup[q] \text{ or } [q]=[p]\cup[q]. \qquad (ii)$$

Since Φ_Δ is $<$-preserving and one-to-one (Propositions 4.3-6 and 4.3-7), (i) and (ii) together imply that $\Phi_\Delta([p]\cup[q])=\max([\Phi_\Delta([p]),\Phi_\Delta([q]))$. Since $[p \vee q]=[p]\cup[q]$, by definition of $\cup$, this gives $\Phi_\Delta([p \vee q])=\max[\Phi_\Delta([p]),\Phi_\Delta([q])]$.

3. Similar to (2). $\square$

Proposition 4.5-2. Let T and I_Δ be as in Proposition 4.5-1; let p be any linguistic assertion of L(T); and let $p(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)$ be any I_Δ-instance of p (where the notation $\underset{\sim}{t}_i$ is as in §4.4): Then

$$I_\Delta(p(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n) = \Phi_\Delta([p(t_1,\ldots,t_n)]).$$

<u>Proof</u>: We use mathematical induction on the length of linguistic assertions.

Case 1: p atomic, of the form $\alpha(t_1^*,\ldots,t_n^*)$. We may assume without loss of generality that all the t_i^* are individual variables, so that $p(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)$ is in this case $\alpha(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)$. Then we have the following:

$$I_\Delta(\alpha(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n) = \mu_{I_\Delta}(\alpha)(I(\underset{\sim}{t}_1),\ldots,I(\underset{\sim}{t}_n)) \qquad (3.2.1\text{-}3.a)$$

$$= \mu_{I_\Delta}(\alpha)([t_1],\ldots,[t_n]) \qquad (3.2.1\text{-}2.a)$$

$$= \Phi_\Delta([\alpha(t_1,\ldots,t_n)]). \qquad (4.4\text{-}3)$$

Case 2: p of the form $\sim q$.

$$I_\Delta(\sim q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)) = 1\text{-}I_\Delta(q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n) \qquad (4.2.1\text{-}3.6)$$

$$= 1\text{-}\Phi_\Delta([q(t_1,\ldots,t_n)]) \qquad \text{(Ind. Hyp.)}$$

$$= \Phi_\Delta([\sim q(t_1,\ldots,t_n)]) \qquad \text{(Prop. 4.5-1.1)}$$

Case 3: p of the form $q\text{v}r$.

$$I_\Delta(q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)\text{v}r(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))$$

$$= \max[I_\Delta(q[\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)),I_\Delta(r(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))] \qquad (3.2.1\text{-}3.c)$$

$$= \max(\Phi_\Delta([q(t_1,\ldots,t_n)]),\Phi_\Delta([r(t_1,\ldots,t_n)])) \qquad \text{(Ind. Hyp.)}$$

$$= \Phi_\Delta([q(t_1,\ldots,t_n)\text{v}r(t_1,\ldots,t_n)]) \qquad \text{(Prop. 4.5-1.2)}$$

Case 4: p of the form $q\wedge r$. Similar to Case 3.

Case 5: p of the form $\phi(p_1(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n),\ldots,p_m(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))$, ϕ m-ary.

$$I_\Delta(\phi(p_1(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n),\ldots,p_m(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)))$$

$$= F_{I_\Delta}(\phi)(I_\Delta(p_1(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)),\ldots,I_\Delta(p_m(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))) \qquad (3.2.1\text{-}3.e)$$

$$= F_{I_\Delta}(\phi)(\Phi_\Delta([p_1(t_1,\ldots,t_n)]),\ldots,\Phi_\Delta([p_m(t_1,\ldots,t_n)])) \qquad \text{(Ind. Hyp.)}$$

$$= F_{I_\Delta}(\phi)(r_1,\ldots,r_n), \text{ where } r_i=\Phi_\Delta([p_i(t_1,\ldots,t_n)]) \qquad \text{(Notation)}$$

$$= \Phi_\Delta([\phi(p_1(t_1,\ldots,t_n),\ldots,p_m(t_1,\ldots,t_n))]). \qquad (4.4\text{-}4)$$

This is what we are required to show. □

<u>Proposition 4.5-3</u>. Let T and I_Δ be as in Proposition 4.5-1; let P

be any formula of L(T); let $P(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)$ be any I_Δ-instance of P. Then

$$I_\Delta(P(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))=1 \text{ iff } [P(t_1,\ldots,t_n)]\varepsilon\Delta.$$

<u>Proof</u>. We use mathematical induction on the length of formulas.

Case 1: P atomic of the form t=t'. Then $P(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)$ is of the form

a=b where a,b are either individual constants of L(T) or are names of

individuals in U_{I_Δ}. Without loss of generality, assume that both are

names, so that $P(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)$ is just $\underset{\sim}{t}_1=\underset{\sim}{t}_2$. Then we have

$I_\Delta(\underset{\sim}{t}_1=\underset{\sim}{t}_2)=1$ iff $I_\Delta(\underset{\sim}{t}_1)=I_\Delta(\underset{\sim}{t}_2)$ (3.2.1-3.f)

    iff $[t_1]=[t_2]$ (3.2.2.1-2.a & §3.4)

    iff $t_1 \sim t_2$ (def. [t], §4.3)

    iff $[t_1=t_2]\varepsilon\Delta$.   (def. $\sim$, §3.3)

Case 2: P atomic of the form $p\cong q$.

$I_\Delta(p(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)\cong q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))=1$

  iff $I_\Delta(p(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))=I_\Delta(q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))$ (3.2.1-3.g)

  iff $\Phi_\Delta([p(t_1,\ldots,t_n)])=\Phi_\Delta([q(t_1,\ldots,t_n)])$ (Prop. 3.5-2)

  iff $[p(t_1,\ldots,t_n)]=[q(t_1,\ldots,t_n)]$ (Prop. 4.3-7)

  iff $p(t_1,\ldots,t_n)\sim q(t_1,\ldots,t_n)$ (def. [p], §4.3)

  iff $[p(t_1,\ldots,t_n)=q(t_1,\ldots,t_n)]\varepsilon\Delta$.   (def. $\sim$, §4.3)

Case 3: P of the form $\neg Q$.

$I_\Delta(\neg Q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))=1$

  iff $I_\Delta(Q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))=0$ (2.2.1-3.h)

  iff $[Q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)]\notin\Delta$ (Ind. Hyp.)

  iff $-[Q(t_1,\ldots,t_n)]\varepsilon\Delta$ (B. alg. Fact 6)

  iff $[\neg Q(t_1,\ldots,t_n)]\varepsilon\Delta$.   (Prop. 4.2-2)

Case 4: P of the form Q R.

$I_\Delta(Q(\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n)\vee R((\underset{\sim}{t}_1,\ldots,\underset{\sim}{t}_n))=1$

iff either $I_\Delta(Q(t_1,\ldots,t_n))=1$ or $I_\Delta(R(t_1,\ldots,t_n))=1$ (3.2.1-3.i)

iff either $[Q(t_1,\ldots,t_n)]\varepsilon\Delta$ or $[R(t_1,\ldots,t_n)]\varepsilon\Delta$ (Ind. Hyp.)

iff $[Q(t_1,\ldots,t_n)]\mathsf{V}[R(t_1,\ldots,t_n)]\varepsilon\Delta$ (B. alg. Fact 10)

iff $[Q(t_1,\ldots,t_n)\mathsf{V}R(t_1,\ldots,t_n)]\varepsilon\Delta$.   (Prop. 4.2-2)

This proves Proposition 4.5-3. □

§4.6. Completeness Theorems

Theorem 4.6-1. If T is a consistent theory and Δ is any ultra-filter in Γ_T such that $\Delta_T \subseteq \Delta$, then I_Δ is a model of T.

Proof. Let T and I_Δ be as specified; let P be any formula of L(T) such that $T \vdash P$; let $P(t_1,\ldots,t_n)$ be any I_Δ-instance of P. By the definition of "model" (§3.4.2), it is sufficient to show that $I_\Delta(P(t_1,\ldots,t_n))=1$.

Since $T \vdash P$, we have that $T \vdash P(t_1,\ldots,t_n)$, by the substitution rule. Then $[P(t_1,\ldots,t_n)]\varepsilon\Delta_T \subseteq \Delta$, by definition of Δ_T; so $I_\Delta(P(t_1,\ldots,t_n))=1$, by Proposition 4.5-3. □

Corollary 4.6-1 (Second Form). Every consistent theory has a model.

Proof. If T is consistent, then at least one canonical interpretation exists for L(T) (cf. §4.4). By the theorem, any canonical interpretation for L(T) is a model of T. □

Theorem 4.6-2 (First Form). For any formula P of L(T), if $\Sigma(T)\models P$, then $T \vdash P$.

Proof. Suppose that P is a formula of L(T) such that $T \nvdash P$. Then T is consistent (§3.4.1). Consider the filter Δ_T in Γ_T. Since $T \nvdash P$, $[P] \notin \Delta_T$. Thus, by Boolean algebra Fact 4, the filter generated by the set $\Delta_T \mathsf{V}\{-[P]\}$ is a proper filter in Γ_T. In accordance with Boolean algebra Fact 5, let Δ be any ultrafilter in Γ_T such that $\Delta_T \mathsf{V}\{-[P]\} \subseteq \Delta$.

Let I_Δ be the canonical interpretation for $L(T)$ based on Δ. Consider the I_Δ-instance $P(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_n)$ of P, where $x_1,\ldots,x_n$ are all the distinct individual variables that occur in P, i.e., where P is $P(x_1,\ldots,x_n)$. Then $-[P(x_1,\ldots,x_n)]=-[P]\varepsilon\Delta$, by the choice of Δ; so $[\neg P(x_1,\ldots,x_n)]\varepsilon\Delta$, by definition of $-$. Hence, $I_\Delta(\neg P(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_n))=1$, by Proposition 4.5-3; so $I_\Delta(P(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_n))=0$, by §3.2.1-3.h. Thus, we have shown that $I_\Delta\nvDash P$. But I_Δ is a model of T, by Theorem 4.6-1. Hence we have shown that $\Sigma(T)\nvDash P$. Since P was chosen arbitrarily, it follows that, for any P of $L(T)$, if $T\nvdash P$, then $\Sigma(T)\nvDash P$. Hence, if $\Sigma(T)\vDash P$, then $T\vdash P$. □

CHAPTER 5

EXTENSIONS AND GENERALIZATIONS

The results of Chapters 3 and 4 provide the basis for formalizing a wide variety of fuzzy set theoretic concepts within the framework of a semantically complete class of formal logical systems. Several of these concepts have already been discussed in the foregoing--e.g., linguistic terms, fuzzy Boolean connectives, and semantic equivalence. The present chapter deals with further concepts which may be similarly formalized, either within systems exactly as defined in Chapter 3, or through modifications thereof.

§5.1 discusses the situation of more than one linguistic variable operating simultaneously over a single universe of discourse.

§5.2 describes the modifications necessary for generalization to multiple universes of discourse.

§5.3 discusses fuzzy logic per se, as based on the idea of linguistic truth, and shows how two different versions--one old and one new--may be developed into semantically complete modes of reasoning.

§5.4 explores very briefly the status of possibilistic reasoning within the present context.

§5.5 delves into the formal characterization of alternative modes of fuzzy inference.

§5.6 lays down sufficient conditions for a linguistic theory to be decideable, thereby ensuring its capacity for implementation on a finite state computer.

§5.1. Multiple Linguistic Variables Over A Single Universe

The examples considered thus far have considered the representation of only a single linguistic variable within a formal theory T. However, the definitions laid down in Chapter 3 provide all the necessary apparatus for any finite or infinite number ($\leq \aleph_1$) of linguistic variables, as long as all those variables are intended to range over the same universe of discourse. The task of introducing multiple variables amounts only to that of developing a grammar which is appropriate for generating all of the desired linguistic terms. Once the grammar is established, then the formal theory of those variables may be obtained by including as proper axioms a set of formulas which express the manner in which the variables are interrelated.

For example, suppose that we are discussing a concept of "good automobile," and it is decided that the criteria for "goodness" are that the auto be attractive, dependable, manuverable in traffic, reasonably priced, and fuel-economical. Then "goodness" would be analyzed as some logical combination of the linguistic variables "appearance," "reliabilty," "manuverability," "cost," and "efficiency." Formalization of this concept thus requires a language for six variables together with a grammar which is at least adequate to provide the formal definition of "goodness" in terms of the five criteria.

To illustrate the latter, let $\alpha, \alpha_1, \ldots, \alpha_5$ be unary relation symbols standing respectively for "good," "attractive," etc. If the property of goodness is thought of as just the simple conjunction of the five criteria, then the appropriate defining axiom would be

$$\alpha(x) \cong \alpha_1(x) \wedge \cdots \wedge \alpha_5(x),$$

where x is any individual variable.

The axiomatic theory thus developed would have among its theorems all the formally derivable consequences of this definition. By semantic completeness, moreover, we have assurance that these consequences will be all and only the formulas which one would expect to be derivable in the context of a logic based on the multivalent theory of approximate reasoning.

§5.2. Generalization To Multiple Universes Of Discourse

It is oftentimes necessary to consider several universes of discourse simultaneously. For example, one might wish to discuss a comparison of the heights of trees, the heights of buildings, and the heights of mountains. In this case, even though the various heights may all be measured on the same scale--say, in meters--the meanings of "tall," "short," etc. will differ depending on the kind of object being measured. In order to express a proposition like "tall trees are shorter than tall buildings and tall buildings are shorter than tall mountains," one therefore needs to formally distinguish between the three kinds of objects within the same language, and as well, to allow for comparisons.

A manner in which this can be accomplished is illustrated as follows. Consider a language which has three distinct sets of individual variables--say $x_1^1, x_2^1, \ldots$ for trees, $x_1^2, x_2^2, \ldots$ for buildings, and $x_1^3, x_2^3, \ldots$ for mountains. In addition, include in this language a binary relation β which allows its parameters to be individual variables of any of the three kinds. Then if α stands for "tall," and β is taken as the "shorter than" relation, the aforementioned proposition may be written as

$$\alpha(x^1) \wedge \alpha(x^2) \wedge \alpha(x^3) \rightarrow \beta(x^1, x^2) \wedge \beta(x^2, x^3),$$

in words, "if x^1 (some tree) is tall, x^2 (some building) is tall, and x^3

(some mountain) is tall, then x^1 is shorter than x^2 and x^2 is shorter than x^3." Here $\rightarrow$ is a fuzzy inference as in §5.5.

To accommodate such a language, the earlier notion of semantic interpretation would be modified to have triples of universes (U_I^1, U_I^2, U_I^3) and to have meaning assignments, together with a concept of "valid formula," as dictated by the structure of the language.

The net result of developing this illustration into the general case will provide a new class of formal theories, together with a companion semantics, for which semantic completeness can be established by the same methods as employed in Chapter 4.

§5.3. Linguistic Truth--Fuzzy Logics

The subject of linguistic truth and the concept of a fuzzy set of type 2 was discussed briefly in §2.2.9. The objective of treating "Truth" as a linguistic variable is to develop a formal means of encoding the everyday usage of truth-related discourse. This includes not only the primitives "true" and false," but also the somewhat more complex--and imprecisely intended--expressions like "very true," "more or less true," "not very true," etc. The desire to accommodate linguistic truth into a formal system of reasoning leads to fuzzy logic per se, which differs from the multivalent logic employed in Chapter 3 by having its truth values be fuzzy subsets of [0,1].

Haack (1979) has argued against the need for such a logic, suggesting that sufficient imprecision is already provided by the multivalent approach. This argument is countered, however, by the fact that linguistic truth nearly always appears in vivo as a linguistic variable, so that fuzzy logic is in actuality a more accurate rendering of the every-

day reasoning with truth-related terms. Given that one of the chief anticipated uses of fuzzy linguistics is in machine intelligence-- specifically, improved man-machine interaction via natural language translation--this "naturalness" of fuzzy logic lends it a central position in the work toward this goal.

The present section recounts some of the earlier work that has been done in this area; it discusses how the best of these results may be formalized within a semantically complete class of formal logical systems; and it introduces a somewhat different, but perhaps more readily applicable approach to linguistic truth in general.

§5.3.1. The Germinal Ideas--Basic Truth-Term Connectives

In developing "Truth" as a linguistic variable, Zadeh (1975b, Part II, p. 334) proposes the following interpretations for the linguistic terms "true" and "false":

$$
\mu_{I(true)}(x) = \begin{cases} 0 & \text{for } 0 \leq x \leq a \\ 2(\frac{x-a}{1-a})^2 & \text{for } a \leq x \leq \frac{a+1}{2} \\ 1-2(\frac{x-a}{1-a})^2 & \text{for } \frac{a+1}{2} \leq x \leq 1, \end{cases}
$$

where a is an arbitrarily chosen number in [0,1], and

$$\mu_{I(false)}(x) = \mu_{I(true)}(1-x) \quad \text{for all } x\varepsilon[0,1].$$

These membership functions have graphs as shown in Figure 5-1. Thus depicted, they clearly represent intuitively plausible interpretations of "true" and "false" in that they assume high values in the appropriate regions of [0,1] and each is a mirror image--hence an antonym--of the other.
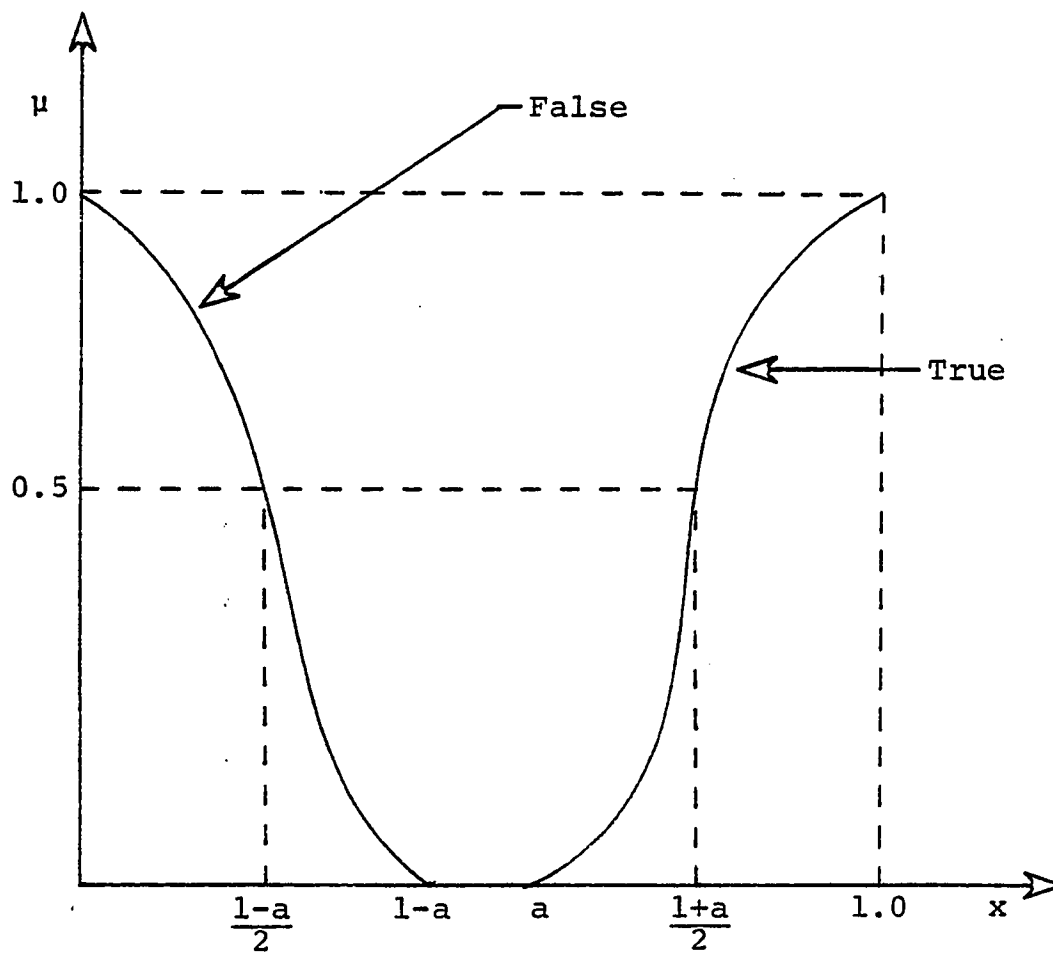
Figure 5-1. Compatibility functions for "True" and "False."

A difficulty arises, however, in that the interpretations of the Boolean connectives given in §2.2.3 for general use with linguistic variables do not serve correctly in the special case of linguistic truth. For example, where common sense requires that "true and false" be reducible to "false," application of the "min" operator to I(true) and I(false) produces a fuzzy subset of [0,1] which is not even approximately equivalent to either I(false) or I(true). Similarly for terms involving "or" and "not."

To rectify this problem, Zadeh therefore equips linguistic truth with its own set of connectives. Let T and T' be terms in the term set for "Truth," and suppose that their meaning assignments are the fuzzy subsets of [0,1] represented by

$$T = \int_i \alpha_i/x_i \text{ and } T' = \int_j \beta_j/y_j.$$

Then the meaning of "not T" is given by

$$-T = \int_i \alpha_i/(1-x_i),$$

the meaning of "T or T'" is given by

$$T \cup T' = \int_{i,j} \sup_{z=\max[x_i,y_j]} \min[\alpha_i,\beta_j]/z,$$

and the meaning of "T and T'" is given by

$$T \cap T' = \int_{i,j} \sup_{z=\min[x_i,y_j]} \min[\alpha_i,\beta_j]/z.$$

These definitions furthermore suggest redefinition of linguistic hedges. For example, "very T" receives the interpretation

$$\text{very-}T = \int_i \alpha_i/x_i^2.$$

This yields many of the basic common sense requirements. In particular, where <--> expresses semantic equivalence, it turns out that

true and false <--> false

true or false <--> true

$$\text{not true} <-> \text{false}$$

$$\text{not false} <-> \text{true.}$$

However, when one considers arbitrary linguistic terms--i.e., arbitrary fuzzy subsets of [0,1]--two further difficulties are seen to persist. First is a problem noted by Haack (1979), that the term set for linguistic truth is not algebraically closed under the above versions of "or" and "and." For example, a compound term like "not very true and not very false" will not in general be reducible to more elementary terms. Now, due to Yager (1979), this problem appears to be resolved by a simple and intuitively appealing form of linguistic approximation.

Second is a somewhat more crucial problem: that of rank ordering the linguistic truth-terms according to a "more true than" relation which is consistent with the everyday usage of "and" and "or." Generally, intuition requires that, for arbitrary truth values T and T', the conjunction, T and T', should be no more true than either T or T', and the disjunction, T or T', should be at least as true as both T' and T.

In algebraic terms, this amounts to requiring that the fuzzy subsets of [0,1], with $\cup$ and $\cap$ as defined above, constitute a lattice; that is, the desired "more true than" relation turns out to be a lattice order. Misumoto and Tanaka (1976) have shown, however, that this algebra of fuzzy subsets of [0,1] is strictly weaker than a lattice; in particular, the absorption laws do not hold. Therefore, it is in fact mathematically impossible to define a rank ordering of the kind desired.

For resolution of this difficulty, two avenues present themselves. First is to again redefine the Boolean connectives, and perhaps also correspondingly redefine the meaning assignments for truth-terms. Second is to retain the above connectives, but limit the base collection

of acceptable meanings for truth terms to some appropriate subcollection of the fuzzy subsets of [0,1]. A solution along the lines of the former is developed in §5.3.3. The section below shows a solution along the latter provided by Mizumoto and Tanaka.

Note: during the research for this dissertation, the possibility of limiting the value set to just fuzzy subsets of [0,1] having monotone membership functions was investigated and was found to be unsatisfactory for the same reason as above: absorption does not apply.

§5.3.2. Fuzzy Numbers

A fuzzy number N in [0,1] is a convex normal fuzzy subset of [0,1], where N is convex if, for all x,y,z in [0,1],

$$x{\leq}y{\leq}z \text{ implies } \mu_N(y){\geq}\min[\mu_N(x),\mu_N(z)],$$

and N is normal if there is at least one x in [0,1] such that

$$\mu_N(x)=1.$$

Mizumota and Tanaka (1976) show that the fuzzy numbers in [0,1], together with Zadeh's connectives as defined in §5.3.1 form a "pseudocomplemented distributive" lattice, i.e., a deMorgan lattice. Hence we have a lattice ordering "$\leq$" with the desired properties: for any two fuzzy numbers N and N', if N$\leq$N', then

$$N{\cup}N' = N'$$

and,

$$N{\cap}N' = N.$$

This ordering is simple to visualize in terms of graphs of the membership function μ_N. The lattice has maximal and minimal elements $\underset{\sim}{1}$ and $\underset{\sim}{0}$ defined by

$$\mu_{\underset{\sim}{1}}(x) = \begin{cases} 0 & \text{if } 0 \leqq x < 1 \\ 1 & \text{if } x = 0 \end{cases}$$

$$\mu_{\underset{\sim}{0}}(x) = \begin{cases} 1 & \text{if } x = 0 \\ 0 & \text{if } 0 < x \leqq 1, \end{cases}$$

and, in general, $N \leqq N'$ if and only if there exists (a unique) $a \varepsilon [0,1]$ such that

$$\mu_N(x) \geqq \mu_{N'}(x) \quad \text{for } 0 \leqq x \leqq a$$

and

$$\mu_N(x) \leqq \mu_{N'}(x) \quad \text{for } a \leqq x \leqq 1.$$

Thus $N \leqq N'$ if and only if the graph of μ_N lies generally to the left of the graph of $\mu_{N'}$ and crosses it at no more than one point. Moreover, the union and intersection of two fuzzy numbers N and N' may be visualized as in Figure 5-2. Note that in this example N and N' are not related by $\leqq$, yet the following relations do hold:

$$N \cap N' \leqq N \leqq N \cup N'$$

$$N \cap N' \leqq N' \leqq N \cup N'.$$

Of especial interest here is that the fuzzy numbers in $[0,1]$ provide a version of linguistic truth which is easily formalizable in a semantically complete class of formal logical systems. There are required only two simple modifications of the definitions developed in Chapter 3: (i) in every linguistic theory, delete from the set of logical axioms all instances of the linear ordering critereon (§3.3.1-6), (ii) in every semantic interpretation, replace the value set $[0,1]$ for linguistic assertions with the set of fuzzy numbers in $[0,1]$, and redefine the connectives $\sim$, $\vee$ and $\wedge$ in terms of the operators of §5.3.1. Note that this automatically transforms the interpretations $I(\alpha)$ of relation symbols α into fuzzy sets of type 2.
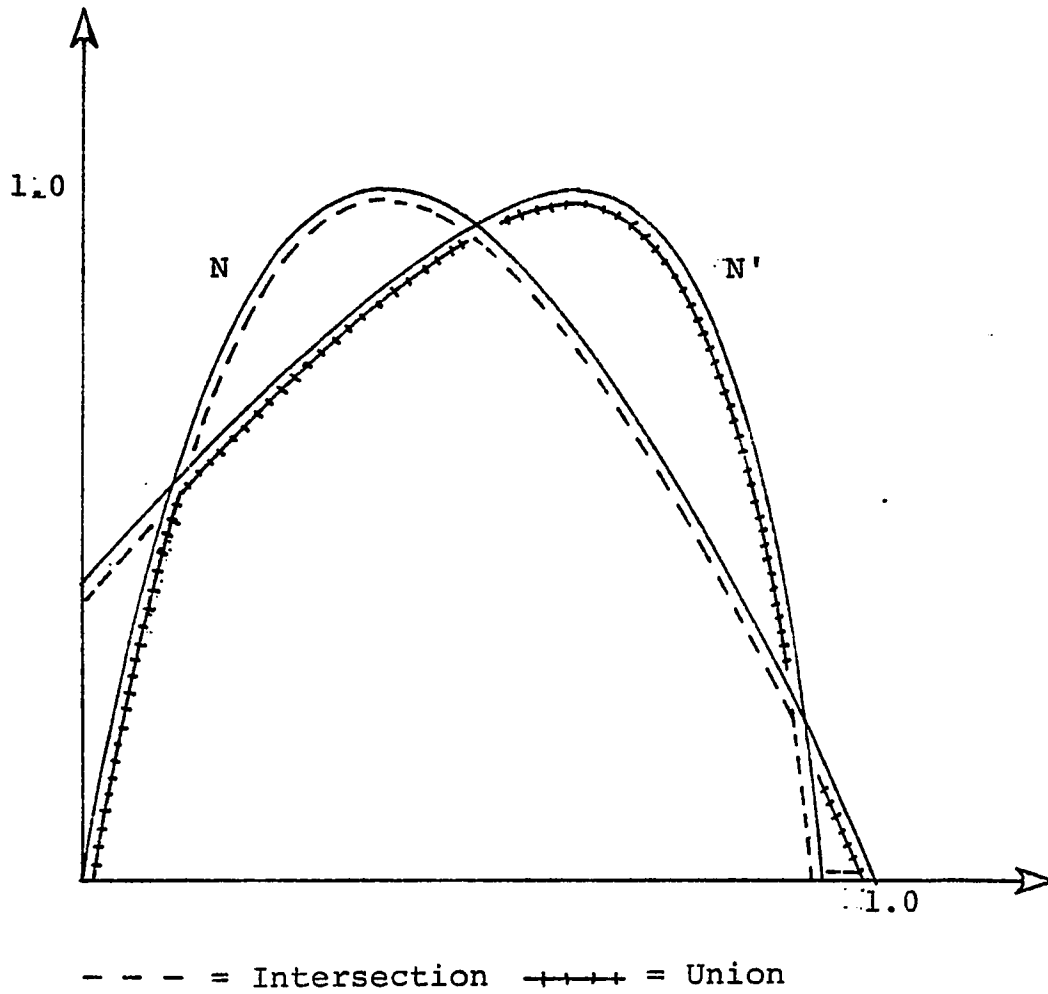
$- - - =$ Intersection $-+-+-+ =$ Union

Figure 5-2. Union and Intersection of fuzzy numbers.

The resultant class of formal systems may be established as semantically complete with respect to the resultant class of semantic interpretations by the methods of Chapter 4. In this case, the crucial step is redefining the cannonical embedding Φ_Δ in such a way that it preserves the lattice ordering of equivalance classes [p] (as opposed to the linear ordering employed in §4.3). The existence of such an embedding is ensured by the fact that the lattice ordering of fuzzy numbers in [0,1] is dense, i.e., for any fuzzy numbers N and N' such that N<N', there is at least one fuzzy number N* such that N<N*<N. This allows us to adapt the definition of Φ_Δ in §4.3 as follows. Let χ_1, χ_2, and χ_3 be as before. Let $\mathcal{N}_1$ be the set of fuzzy numbers N in [0,1] such that -N<N (where "-" is as in §5.3.1), let $\mathcal{N}_2$ be the set such that -N=N, and let $\mathcal{N}_3$ be the set such that N<-N. Then let

1. $\Phi_\Delta : \chi_1 \to \mathcal{N}_1$ be any <-preserving embedding,

2. $\Phi_\Delta : \chi_2 \to \mathcal{N}_2$ be any mapping of χ_2 into $\mathcal{N}_2$,

3. $\Phi_\Delta : \chi_3 \to \mathcal{N}_3$ by $\Phi_\Delta([p])=-\Phi_\Delta([\sim p])$, where $[\sim p]\varepsilon\chi_1$.

It follows that the cannonical interpretation I_Δ, as defined in §3.4 in terms of Φ_Δ, of any consistent theory is well-defined, and that the analogues of all the propositions and theorems of §3.5 and §3.6 hold true. Thus one obtains a semantically complete axiomatization of fuzzy logic.

Because fuzzy numbers in [0,1] form a lattice, the logic captured in this class of systems has many desirable properties. One possible defect persists, however, in that the lattice ordering relation is generally a partial order--i.e., not all truth values will be comparable. Theories with total orders can be defined, nevertheless, through the

techniques described in Example 3.4.3. In this case it must be assumed that there is a 0-ary relation symbol κ_N for every fuzzy number N.

§5.3.3. An Alternate Approach

Since it is customary for logicians to use numbers in [0,1] as truth values, it is only reasonable that when moving to the more general conception represented by linguistic truth one should employ fuzzy subsets of [0,1]. In reflecting on the everyday use of truth-related language, however, it is evident that there is no truely compelling reason why "Truth," even considered as a linguistic variable,. should have [0,1] as its universe of discourse. Indeed, a moments reflection will remind us that the actual universe of discourse for this variable as it appears in natural languages is more exactly the class of all · propositions--to wit, it is actually the linguistic assertions, and not numbers, which are "true," "false," "more or less true," "not very true," etc. It therefore makes sense to look for a formulation of linguistic truth which incorporates this view.

This section presents one such formulation, which is conveniently formalizable in the systems discussed in Chapter 3. In this treatment, linguistic terms from the term set for "Truth" are introduced formally as operators on linguistic assertions--somewhat on the same order as linguistic hedges--with each such term being assumed to either hold or not hold for any given proposition. For example, the claim that a proposition p is "true" will be expressed formally by the linguistic assertion true(p), where "true" here denotes a unary operator symbol; the proposition that "true(p)" holds, will be represented by the formula true(p)$\cong\kappa_1$, where κ_1 is the zero-ary relation symbol with constant value

1; and the proposition that "true(p)" does not hold will be represented by the formula $\text{true}(p) \cong \kappa_0$.

The determination of whether "true(p)" does or does not hold is then dependent on the value in $[0,1]$ which a given interpretation I assigns to p. If it is decided, for example, that a proposition p deserves to be regarded as true if $I(p) \geq 0.6$, then the operator "true" may be formally associated with the interval $[0.6,1]$ by some defining axioms:

$$(\text{true}(p) \cong \kappa_1) \equiv (p \vee \kappa_{0.6} \cong p)$$

$$(\text{true}(p) \cong \kappa_0) \equiv \neg(\text{true}(p) \cong \kappa_1).$$

Additional linguistic truth terms may be defined in a similar fashion. As examples: "false" may be defined by

$$(\text{false}(p) \cong \kappa_1) \equiv (p \wedge \kappa_{0.4} \cong p)$$

$$(\text{false}(p) \cong \kappa_0) \equiv \neg(\text{false}(p) \cong \kappa_1),$$

"more or less true" by

$$(\text{m-l-true}(p) \cong \kappa_1) \equiv (p \vee \kappa_{0.4} \cong p) \& (p \wedge \kappa_{0.6} \cong p)$$

$$(\text{m-l-true}(p) \cong \kappa_0) \equiv \neg(\text{m-l-true}(p) \cong \kappa_1),$$

"very true" by

$$(\text{v-true}(p) \cong \kappa_1) \equiv (p \vee \kappa_{0.8} \cong p)$$

$$(\text{v-true}(p) \cong \kappa_0) \equiv \neg(\text{v-true}(p) \cong \kappa_1)$$

and "absolutely true" by

$$(\text{a-true}(p) \cong \kappa_1) \equiv (p \cong \kappa_1)$$

$$(\text{a-true}(p) \cong \kappa_0) \equiv \neg(\text{a-true}(p) \cong \kappa_1).$$

Choice of the range in $[0,1]$ for each linguistic term is of course arbitrary, but will normally be dictated to a certain extent by the kinds of interrelations that one wishes to set up among the terms being used. With the above definitions of "true" and "false," for example, it

happens that these terms are not autonyms of one another, i.e., for

arbitrary p, both of the formulas

$$\sim true(p) \cong false(p)$$

$$\sim false(p) \cong true(p)$$

are semantically invalid; yet they become valid if "true" is associated

with [0.5,1] and "false" is associated with [0,0.5].

The virtue of this approach to linguistic truth rests on the

following result.

Theorem 5.3.3-1. If the intervals associated with "true" and

"false" are mirror images of one another about the point 0.5, then these

two linguistic terms satisfy all of the classical principles for $\sim$, $\lor$,

and $\land$; that is, the following truth table holds:

p	q	$\sim$p	p$\lor$q	p$\land$q
true	true	false	true	true
true	false	false	true	false
false	true	true	true	false
false	false	true	false	false

Proof. Each entry in the truth table may be verified by straight-

forward appeal to the definitions, together with applications of the

semantic completeness result for formal theories. To illustrate: the

first entry--if p is true, then $\sim$p is false--is established as follows:

$$\neg(true(p) \cong \kappa_1) \equiv (p \lor \kappa_{0.6} \cong p) \qquad \text{(def. of "true")}$$

$$\equiv (\sim p \ \kappa_{0.4} \cong p) \qquad \text{(sem. completeness)}$$

$$\equiv false(\sim p) \cong \kappa_1. \qquad \text{(def. of "false.")}$$

Details of the remaining entries are omitted. □

One thus has in this formulation a "fuzzy logic" on a different

order. Individual linguistic truth terms apply non-fuzzily, i.e., they

either hold for a given proposition p or they do not hold for p; yet

linguistic imprecision is accommodated (i) by allowing a range of [0,1]

to be assigned to each linguistic term, and (ii) by accommodating for

any number of fuzzy linguistic terms, with or without hedges. Moreover,

by Theorem 5.3.3-1, appropriately defining the meanings of "true" and

"false" leaves the full strength of the classical logic completely

embedded within a multivalent fuzzy logic, thereby preserving the common

sense usage of true-false reasoning. Last, this version of linguistic

truth is mathematically simpler than the versions considered in §5.3.1

and §5.3.2, thus making it somewhat more amenable to direct implementa-

tion on a computer.

§5.4. Linguistic Possibility

Zadeh (1978a and 1978b) has shown that the imprecision in everyday

discourse may be regarded as essentially possibilistic in nature. The

basic idea of a possibility distribution was given in §2.2.10. We here con-

sider the status of possibilistic reasoning from the formal point of view.

Consider the fuzzy set SI of "small integers" defined in §2.2.10,

$$SI = 1/0 + 1/1 + 0.8/2 + 0.6/3 + 0.4/4 + 0.2/5.$$

The proposition that x has a possibility of 0.6 of assuming the value 3

may be expressed by the notation

$$Poss\{x=3\} = 0.6.$$

Thus written, possibilistic assertions are clearly metalinguistic, and

are of a kind which is not directly formalizable in the languages

developed in Chapter 3.

Yet because of the correspondence between fuzzy sets and possibility

distributions, the net effect of a possibilistic interpretation is

implicit. It is simply the matter of an intuitive heuristic whether one takes a formal expression like x=a with truth value $I(x=a)=i\varepsilon[0,1]$ as asserting that "x=a is true to the degree i" or as saying "x is possibly a to the degree i." Therefore the necessary apparatus for formalizing many of the concepts from possibility theory is available in the systems of Chapter 3.

This includes in particular the idea of a <u>possibility measure</u> $\Pi(A)$ for a non-fuzzy set A, as induced by a possibility distribution Π_x: if Π_x is given as some fuzzy set B, then $\Pi(A)$ represents the possibility that $x\varepsilon A$, and is defined by

$$\Pi(A) = \sup_{a\varepsilon A} \mu_B(a).$$

By virtue of the two-levels in our formal languages, this definition is effectively captured in a formula of the following form:

$$(\Pi(\alpha(x))\cong\beta(x)) \equiv (\neg(x'=x)\supset\beta(x')\vee\beta(x)\cong\beta(x)).$$

This shows that even fairly sophisticated concepts are formalizable in the languages as given. Many concepts will not be accessible in this manner, however. In particular are the fuzzy quantifiers mentioned in §2.2.11.

§5.5. Fuzzy Inference

As was mentioned in §2.2.8, the fuzzy interpretations of the Boolean connectives are nowdays well-established in the literature, while the question of an appropriate form of the fuzzy "implies" remains open. Gaines (1976) lists several alternative formulations, and other versions appear in papers by Zadeh. We shall here give a sampling of three such connectives and consider their formalizability both within multivalent theories of fuzzy reasoning and within fuzzy logics of the

kind developed in §5.3.2. These three are conveniently chosen as those for which Dubois and Prade (1979b) have already developed the fuzzy-logical analogues, i.e., generalizations for use with fuzzy sets of type 2.

§5.5.1. Dienes-Rescher

This is the $\rightarrow$ connective of classical logic, simply generalized to the case of approximate reasoning. Formally it may be introduced into any theory T--multivalent or fuzzy-logical--by adjoining axioms according to the scheme

$$p \rightarrow q \cong \sim p \lor q.$$

In the multivalent case, this defines an operator => on fuzzy sets of type 1 by

$$\mu_{A=>B}(x,y) = \max[1-\mu_A(x),\mu_B(y)],$$

where A and B are the meanings of p and q. In the fuzzy-logical case, the same scheme defines an operator => on fuzzy sets of type 2 by

$$A=>B = -A \cup B,$$

where - and $\cup$ are as given in §5.3.2.

This mode of inference is generally not satisfactory as a form of approximate inference for the reason that, if $p \rightarrow q$ has a higher truth value than p, then the truth value of q is greater than or equal to the truth value of $p \rightarrow q$, contrary to the intuitive requirement that the conclusion of a logical inference (in this case, modus ponens) should not be more true than any of the hypotheses.

§5.5.2. Lukasiewicz $L_{\aleph_1}$

This mode of inference has been the most popular, partly because it satisfies the above intuitive requirement for modus ponens, and

partly because it has already been studied extensively in the theory of multivalued logics. Gaines (1976) discusses it at length in the context of fuzzy reasoning, and Bellman and Zadeh (1977) have used it as a "base logic" which is "fuzzified" to produce a fuzzy logic over fuzzy sets of type 2.

In the multivalent semantics the Lukasiewicz connective is defined by an operator $\Rightarrow$ on fuzzy sets of type 1, as:

$$\mu_{A \Rightarrow B}(x) = \min[1, 1 - \mu_A(x) + \mu_B(x)],$$

and in the context of fuzzy sets of type 2, it is given by

$$A \Rightarrow B = \underset{\sim}{1} \wedge [-A \oplus B],$$

where $\underset{\sim}{1}$, $\wedge$, and $-$ are as in §5.3.2, and where $\oplus$ is defined by

$$A \oplus B = \int_{x,y} \min[\mu_A(x), \mu_B(y)]/x + y.$$

Because of the use of the arithmetic sum, this mode of inference cannot be formally defined explicitly in terms of the connectives $\sim$, $\vee$, $\wedge$. However, the essence of this $\rightarrow$ connective can be captured in a theory T--either multivalent or fuzzy-logical--by adjoining all axioms for $L_{\aleph_1}$. Following Gaines (1976), this leads to axioms having the forms:

1. $(p \cong \kappa_1 \,\& p \rightarrow q \cong \kappa_1) \supset q \cong \kappa_1$ (modus ponens)

2. $p \rightarrow (q \rightarrow p) \cong \kappa_1$ (paradox)

3. $((p \rightarrow q) \rightarrow p) \rightarrow ((q \rightarrow p) \rightarrow p) \cong \kappa_1$ (disjunction)

4. $(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow (p \rightarrow r)) \cong \kappa_1$ (transitivity)

In addition, one should have axioms which relate $\sim$, $\wedge$, and $\vee$ to $\rightarrow$. The necessary relationships would here be derivable from

5. $\sim p \cong p \rightarrow \kappa_0$

6. $p \vee q \cong (p \rightarrow q) \rightarrow q$.

By using these axioms in a formal theory T, one is ensured that → has

the requisite logical behavior, even though it will in general have

different meanings in different semantic models of T.

Note: An alternative approach would be to introduce a connective +

which satisfies sufficiently many of the rules for addition of real

numbers so that → could be introduced by

$$p→q \cong \kappa_1 \wedge ((\sim p)+q).$$

In fact, it is conceivable that + could be defined in such a way that

its only possible meaning assignment is addition of reals, in which case

p→q is uniquely defined for all semantic interpretations. Development

of the necessary axioms for such a + connective has not been pursued in

this research.

§5.5.3. Zadeh's Generalized Modus Ponens

A relational mode of inference was proposed in Zadeh's first

papers on the subject, and continues to be discussed (c.f. Zadeh 1975,

Part III for a review). The assertion "if p then q" is considered as a

special case of "if p then q, else r," which receives as its meaning the

union of cartesian products

$$(A \times B) \cup (-A \times C),$$

where it is assumed that B and C are fuzzy subsets of the same universe V.

The requisite special case is gotten by taking C=V, so that

$$A \Rightarrow B = (A \times B) \cup (-A \times V)$$

This leads to a generalized modus ponens through its combination

with a compositional rule of inference, which is defined as follows.

Let A be a fuzzy subset of U, and let R be a fuzzy subset of U×V, i.e.,

a binary fuzzy relation. Then a fuzzy set B may be <u>inferred</u> from the fuzzy sets A and R as the composition

$$B = AoR.$$

Since the meaning of "p implies q" given above is a binary fuzzy relation of U and V, the rule of modus ponens "from p and p→q infer q" obtains as its meaning

$$B = Ao(A=>B).$$

This mode of inference is formalizable in the multivalent systems of Chapter 3, since cartesian products of fuzzy sets and compositions of fuzzy relations are formally definable in the given two leveled languages. (For compositions, the role of the "sup" operator is captured as for Π(A) in §5.4.)

Dubois and Prade (1976) give a fuzzy logical version of this mode of inference as well. Here it is left as a conjecture for future investigations that this mode is similarly representable in a fuzzy logic as described in §5.3. The task is to formalize the generalizations of "max," "min," and "sup" to fuzzy sets of type 2.

§5.6. Decidability

Given that a substantial portion of the theory of approximate reasoning has been axiomatized in a semantically complete class of formal logical systems, the question next arises concerning which of these systems are representable on a finite state computer. In mathematical terms this amounts to determining decidability, where a theory T is <u>decidable</u> if there is a mechanical procedure (a <u>decision procedure</u>) by which one can effectively determine in finitely many steps whether a formula P of L(T) is a theorem of T. (For a more detailed discussion of

decision procedures, cf. §8.1 and §8.3.) Here decidability will be established for a subclass of the theories defined in Chapter 3.

By semantic completeness, a formula P is a theorem of a theory T if and only if it is valid in every semantic model of T. Therefore, a decision procedure for T would amount to a procedure for determining the semantic validity of formulas of L(T). By the definition of semantic validity, for such a procedure to exist it is sufficient that T satisfy:

1. T has only finitely many semantic models,

2. every model of T has a finite universe of discourse, and

3. every model of T effectively uses only finitely many of the truth values in [0,1].

In certain cases it may be possible to modify a given theory so that these requirements are satisfied. We here discuss a procedure for doing this, by treating the three conditions in reverse order.

Condition 3 may be ensured for a theory T by the techniques discussed in Example 3.4.3. Let the language L(T) contain a finite set of 0-ary relation symbols κ_i, where the set of indices $i_1,\ldots,i_n$ is symmetric about the point 0.5. Adjoin as proper axioms all formulas of the form

$$p \cong \kappa_{i_1} \lor p \cong \kappa_{i_2} \lor \ldots \lor p \cong \kappa_{i_n}.$$

Then in every semantic model of T, the linguistic assertions in a valid formula P can assume only the values $i_1,\ldots,i_n$.

Condition 2 can be ensured in an analogous manner. Let L(T) contain a finite collection of individual constants $a_1,\ldots,a_m$, and adjoin as a proper axiom a formula of the form

$$x = a_1 \lor x = a_2 \lor \ldots \lor x = a_m.$$

Then individual variables in a valid formula can range over only finitely many individuals.

Condition 1 is ensured by requiring that the theory T have a property of categoricity, which may be defined as follows. Two models I and I' of T are _isomorphic_ if there is a one-to-one correspondence Ψ between U_I and $U_{I'}$ such that

$$I(\alpha(\underset{\sim}{i}_1,\ldots,\underset{\sim}{i}_n))=I(\alpha(\Psi(\underset{\sim}{i}_1),\ldots,\Psi(\underset{\sim}{i}_n))$$

for all relation symbols α in L(T) and all n-tuples of individuals $i_1,\ldots,i_n \varepsilon U_I$. It can be shown that two models I and I' of T are isomorphic if and only if exactly the same formulas of L(T) are valid in both I and I'. Then, a theory T is _categorical_ if all of its semantic models are isomorphic. Because of the manner in which Conditions 2 and 3 are ensured, categoricity can be ensured by explicitly defining the meanings of every relation symbol within the formal theory T: for every k-ary relation symbol α, and every k-tuple $a_{i_1},\ldots,a_{i_k}$ of the individual constants specified for Condition 2, adjoin a proper axiom of the form

$$\alpha(a_{i_1},\ldots,a_{i_k}) = \kappa_i,$$

where κ_i is one of the 0-ary relation symbols specified for Condition 1.

When invoking this procedure, however, it is important to verify that this does not lead to a theory which is inconsistent. In particular, the procedure is not applicable to any theory whose axioms already require that either the set of individuals or the set of truth values be infinite. For in this case, the above methods for satisfying conditions 2 and 3 will yield a theory that has no semantic models and which is therefore inconsistent by the second form of the completeness theorem (Theorem 4.6-1). An illustration of this is the theory CM developed in

§6.3, which requires that its set of truth values obey the axioms for an algebraic field. Condition 3 cannot be satisfied in this case, since it is known that the field axioms are not satisfied by any finite subset of the real numbers. On the other hand, the procedure succeeds for the theory described in Example 3.4.3; and many other, considerably more complex, theories of one or more linguistic variables can be rendered decidable in this fashion without difficulty.

Note also that undecidability does not of itself forbid computer applications. It rather points out potential sources of error where assumptions of finitude are involved. This situation is also illustrated by the theory CM of §6.3.

CHAPTER 6

AN APPLICATION--SPIN$^+$

Uncertainty about the future is in part a consequence of the imprecision in our knowledge of the present, together with our manner of reasoning with imprecise ideas. This chapter outlines an approach to structural modeling of complex systems which makes this present-future relationship formally explicit, via fuzzy sets and the theory of approximate reasoning. The main result is a direct mathematical connection between the imprecision in a given structural model and the degree of uncertainty that this imprecision contributes to the model's forecasted results.

§6.1 states the problem context, taking into consideration the aim of forecasting methodology, the use of structural models, and the question of "certainty accounting."

§6.2 takes McLean's (1976) modeling package SPIN, and enhances it to a package SPIN$^+$ by adding some algorithms for computing certainty values for forecasted events. SPIN$^+$ is illustrated with sample calculations for a five-variable "Citizen's Model" of a transit system.

§6.3 lays down all the details needed for representing SPIN$^+$ and the Citizen's Model in a multivalent linguistic theory.

§6.4 discusses the possibility and potential advantages of improving this approach through an implementation of fuzzy logic as is described in §5.3.

§6.1. Forecasting Under Uncertainty

The task besetting contemporary forecasters stems from man's being caught between the two sides of a perennial dilemma. First, we must plan. This clearly is necessary merely for human survival. Even the simplest of cultures must undertake a small amount of planning in order to meet the minimum requirements of food and shelter. And in the more sophisticated cultures, such as ours, planning is a common aspect of everyday life. In order to survive we must have well-defined goals and a reasonable idea of how they can be achieved.

At the same time, however, in order that our plans have a practical viability, they must also be couched within a forecast of possible future events. A plan which is crafted in a vacuum, that is, in complete disregard for the context within which it is to be executed, is not really a plan at all. Any project can be affected by myriad factors of different kinds, e.g., physical, biological, human, economic, environmental, and so on. Hence no plan can be crafted without a reasonable prognosis of what lies ahead.

But herein resides the second half of the dilemma. We cannot accurately predict. No one knows for sure what will happen in a given situation (except perhaps a few self-proclaimed soothsayers and clairvoyants, whose claims still remain dubious). Therefore, all our predictions leave an element to chance. Cognizance of chance overshadows our plans with uncertainty.

Thus arises the problem of forecasting: to improve our ability to see into the future so as to reduce uncertainty. The basic tool of the forecaster is the concept of a model, by which one abstracts form a given situation those aspects deemed relevant to the plan or goals one

has in mind. Such models may be roughly grouped into the categories of mathematical and nonmathematical, while the former may in turn be subdivided according to mechanistic and probabalistic. Mechanistic mathematical models normally implement the mathematics of dynamical systems, such as linear algebra or differential equations, and treat the future as being predetermined by the characteristics of the present. Probabalistic models implement the laws of averages, i.e., statistics, and investigate the future in terms tendencies, or trends, of events. Nonmathematical models, on the other hand, forego rigorous computations and focus on the nonmeasurable "ill-defined" aspects of a system, e.g., problems of management or individual human behavior. These distinctions are not at all crisp, however, as is illustrated in this dissertation, and particularly, in this chapter. There are mathematical techniques which can be applied to certain kinds of "ill-definedness."

The best forecasts combine an assortment of modeling techniques, so as to bring into consideration an optimal collection of relevant factors. Yet even given the most carefully wrought prognosis, a degree of uncertainty must remain. First, there will often be aspects of the system in question for which no known modeling procedure applies. This inability to model leaves uncertainty regarding those aspects unaffected, i.e., it is at the maximum. Second, inasmuch as models are only abstractions of carefully chosen factors, they characteristically blur a large quantity of detail. Hence, models are always imprecise, even where they apply, and this imprecision automatically propogates an element of uncertainty into the models predictions. Last, experience teaches that we in any case must always expect the unexpected. Here we note a fact

of the planner-decisionmaker's life, that any number of human factors
and unexpected events may affect what actually occurs.

For these reasons, forecasters have of late adopted the practice
of accompanying their forecasts with an accounting, as best as is feasible,
of the regions where uncertainty remains. This assists the planner by
pointing out those areas for which he should incorporate flexibility.
In the face of uncertainty, of course, this is the planner's only realis-
tic alternative: to design courses of action which can adapt.

The work in this chapter deals with a somewhat more specialized
question within this problem of uncertainty accounting, that of develop-
ing an actual measure of the uncertainty that obtains. In this, the
current approach is a priori limited to a consideration of only those
forecasting methods for which such a measure can be defined, namely, the
mathematical models. Specifically, we consider the "structural models,"
of which a wide variety are currently in vogue.

An extensive study of selected structural modeling techniques,
explicitly for their use in forecasting, was recently undertaken by a
team at Portland State University under the auspices of the National
Science Foundation (Linstone, et. al. 1978 and 1979). An outgrowth of
that project is a more recent paper which gives a thorough analysis of
the nature of structural models, a comparison of the essential features
of a selected collection of structural modeling tools, and a discussion
of the kinds of information about a system that can be obtained in this
way (Lendaris 1980).

Structural models differ from other mathematical models in that
they directly represent systems in terms of their structure and inner
dynamics, while relegating quantitative (i.e., numerical) considerations

to a level which is mostly nonsignificant. In general, a structural model is defined as a collection of variables (or parameters) together with their interactions (Lendaris, p. 808). A chief advantage of structural modeling is that the system being modeled is visualized geometrically as a digraph which shows the impacts of each parameter on the other (as will be seen later in Figure 6-1). Impacts may be either positive or negative, linear or nonlinear, have short or long delay times, etc., and they may be measurable in real units (e.g., dollars or BTU's) or in nominal units on a qualitative scale (such as from "high" to "low"). The digraph in turn can be transformed into a matrix, which is analyzed on a computer.

Use of the model in generating forecasts involves a "what if..." analysis wherein one tests the future consequences over nominal or real time units of hypothetical changes in the present or future values of the parameters. Any one forecasting study may include a variety of different sets of hypotheses and results, and as well, the results of several different structural models of the same system. These increase our understanding of the system by illustrating the implications of conceptualizing the system in different ways.

As both Linstone and Lendaris point out, an important aspect of this technique is the insight into the system which is gained by the modeler during the process of building the model itself--i.e., creating the digraph, assigning weights, delay times, etc. Given that the model output is to be published as part of a forecasting study, however, it is mainly the results which bear meaning for the planner. For this reason, the planner is best advised that the results should not be taken too literally, i.e., as predictions. As was stressed above, such forecasts

always inherit an element of imprecision from the imprecision of the model.

Hence arises the question of how much imprecision obtains; or more exactly, given a measure of model imprecision, to what extent does this affect the confidence that one may place in the forecasted results. The sections to follow provide an approach to this question via the theory of approximate reasoning. Here the modeling package SPIN of McLean, et. al. (1976) is enhanced by allowing the modeler to adjoin judgments as to the "certainty of impacts" along with the usual weights and delays. These certainty values are then manipulated in parallel with the ordinary SPIN calculations, according to a property chosen from of fuzzy inference. The result is a unique "certainty of forecast" computation along the individual system parameters for each time unit that the parameters are projected into the future. A sample run of the modified package shows that a reasonable link between model imprecision and uncertainty of forecast is established. If the assigned certainty of impacts in the digraph are high, then certainty of projected parameters degrades slowly with respect to time, whereas certainty degrades rapidly if the assigned certainties are low.

In using this approach, however, two points regarding the present use of the word "uncertainty" should be noted. First, the "certainty measures" which the modeler contributes to the digraph are strictly subjective measures of the modeler's confidence that the impacts will occur and will persist over time. Hence they are measures of "imprecision" as such, but only given that the basic structure and dynamics of the model is correct. They do not account for imprecison due to "hidden variables" or other system aspects which the model does not represent.

Second, the computed "certainty of forecasts" is strictly connected only with the model imprecision and should not be taken in the broader sense discussed above. Uncertainty due to an inability to model or an awareness of the unexpected must be dealt with by other means.

Last it should be mentioned that, whereas SPIN and its relatives among the class of structural models are strictly mechanistic in their representation of reality, the present proposal to introduce approximate reasoning is a move to combine this mechanism with a form of stochasm. This of course is not stochasm in the traditional sense of being based on probability theory (for the reasons cited in §2.2.10). But it is stochastic inasmuch as it deals with uncertainty and is based on an idea which is akin to subjective probability. Here judgments as to model imprecision are obtained in the form of subjective choices of degrees of certainty that a given impact relation will persist over time.

In focusing specifically on the impact relations, moreover, the present approach differs further from conventional stochastic methods, which deal only with uncertainty in the values of the system parameters. Yet it at the same time still leaves unaddressed the common limitation of all structural models, this being that they do not account for uncertainty regarding the model's topology. There is no provision for the possibility that a given impact relation may cease to be effective, or that new impact relations may appear. Even given these limitations, however, it is clear that the models developed here are equally applicable, and in the same situations, as their nonstochastic predecessors.

§6.2. McLean's SPIN And The Enhancement SPIN$^+$

It will be convenient to preface the technical details of SPIN$^+$ with an outline of the basic idea. Suppose that a digraph has been constructed as in Figure 6-1, and assume that two variables A and B are related by

"an increase of ΔA in the value of A produces an

increase of $\Delta B = i\% \times \Delta A$ in the value of B after a

delay of d time units."

This assertion may recast as a logical implication

$$P_A \rightarrow P_B,$$

where

P_A is "A increases by ΔA," and

P_B is "B increases by $\Delta B = i\% \times \Delta A$

after a delay d."

Then a <u>degree</u> <u>of</u> <u>certainty</u> $c(P_A \rightarrow P_B)$ measured on a scale from 0 to 1 can be assigned, with a high (low) degree indicating a high (low) confidence that the inference $P_A \rightarrow P_B$ is true. Given an explicit definition of $\rightarrow$ as a function on $[0,1]$, this leads to a capability for computing an in-ferred degree of certainty $c(P_B)$ from the pair $c(P_A)$, $c(P_A \rightarrow P_B)$. In turn this leads to a capability for computing degrees of certainty for the consequent P_{A_n} of a sequence of inferences $P_{A_1} \rightarrow P_{A_2}, \ldots, P_{A_{n-1}} \rightarrow P_{A_n}$ from $c(P_{A_1})$ and the $c(P_{A_{i-1}} \rightarrow P_{A_i})$, $i=2,\ldots,n$.

Certainty accounting is accomplished automatically by the computa-tions, with the specific characteristics of the accounting being dependent on the choice of the implication operator $\rightarrow$. For the present purposes, the most appropriate choice appears to be the Lukasiewicz operator discussed in §5.5.2. Given the definition
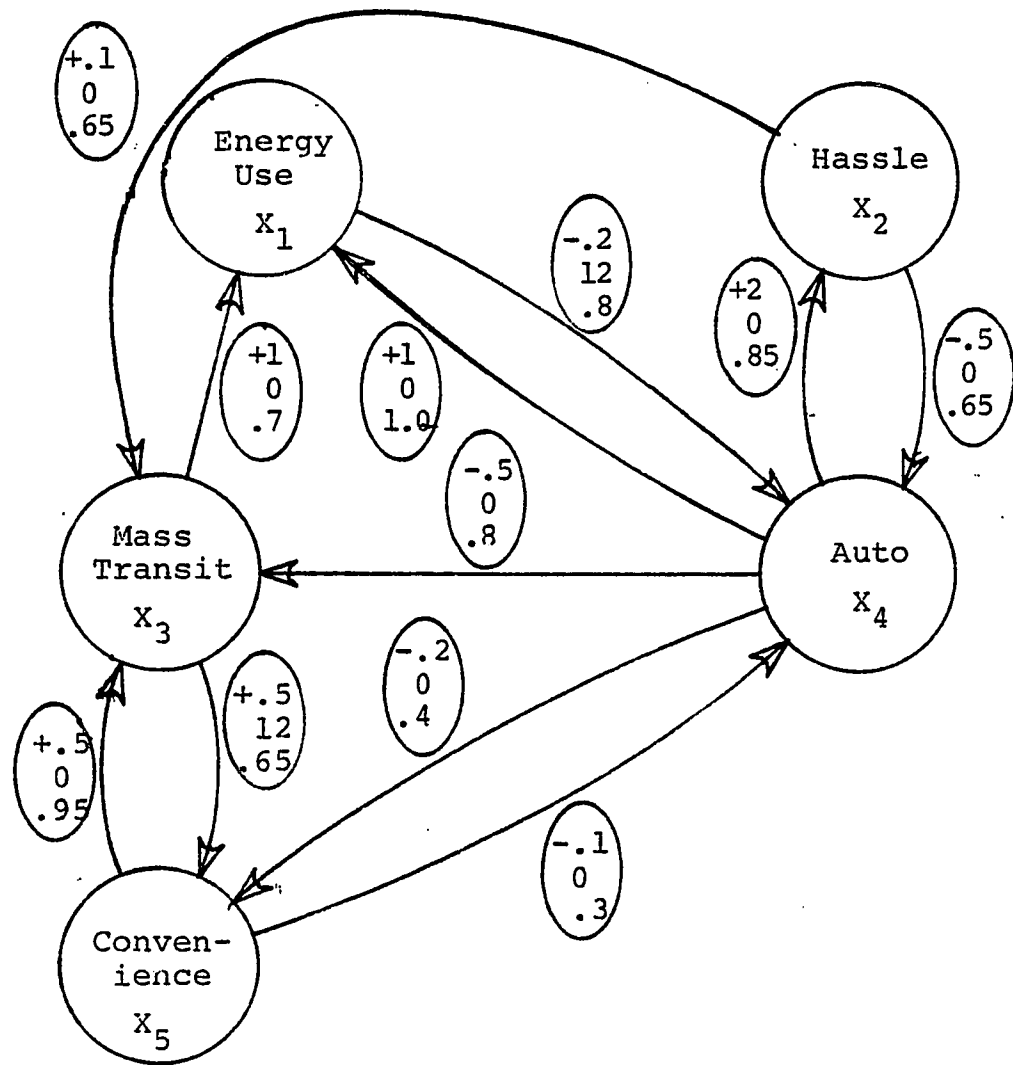
Variables

X_1 = Use of Energy

X_2 = Hassle of Driving Car
(Parking,concestion,etc.)

X_3 = Use of Mass Transit

X_4 = Use of Automobile

X_5 = Convenience of Mass Transit

Impacts

1st no. is a_{ij} = Weight

2nd no. is d_{ij} = Delay

3rd no. is c_{ij} = Certainty

Figure 6-1. The Citizen's Model for SPIN[+].

$$c(P_A \rightarrow P_B) = \min[1, 1-c(P_A)+c(P_B)] \tag{1}$$

we can solve for $c(P_B)$ obtaining

$$c(P_B) \geqq \max[0, c(P_A)+c(P_A \rightarrow P_B)-1], \tag{2}$$

which forms a basis for the accounting procedure. For example, suppose

that $c(P_A)=1.0$, $c(P_A \rightarrow P_B)=0.9$, and $c(P_B \rightarrow P_C)=0.9$. Then, by (2), $c(P_B) \geqq$

1.0+0.9-1=0.9, from which it follows in turn that $c(P_C) \geqq 0.9+0.9-1=0.8$.

Note that, while the certainty of the two inferences are equal, we have

that $c(P_A) > c(P_B) > c(P_C)$, i.e., the certainties of inferred events steadily

degrade. For computational simplicity, equality is assumed in (2),

which has the effect of making each inferred value be the highest

certainty estimate that is ensured by the previously computed values.

The core of McLean's modeling package may be described as follows

(Linstone, et. al., 1978, p. D-9). Let $X_1, \ldots, X_n$ be the parameters in a

digraph representation of a given system; let w_{ij} and d_{ij} be the weight

and delay of the impact of X_i on X_j; let $X_i(0)$ be the assumed value of

X_i at time t=0; and assume that one or more of the X_i are "pulsed" by

letting $X_i(1)=X_i(0)+$pulse. Then future values of the X_i are computed by

the formula

$$X_j(t+1) = X_j(t)+\sum_{i=1}^{n} w_{ij} \cdot (X_i(t)-X_i(t')),$$

where $t'=\max[0, t-1-d_{ij}]$. In words, the next value of X_j is the current

value of X_j plus the net change (which may be positive or negative)

brought about by the cumulative impacts of the most recent changes in

all of the X_i.

Calculation of a degree of certainty for $X_j(t+1)$ may now be de-

veloped in parallel. Let $\Delta X_{ij}(t)$ denote $X_i(t)-X_i(t')$, the change in X_i

over the time of the delay d_{ij}; and let $\Delta_i X_j(t+1)$ denote $w_{ij} \cdot \Delta X_{ij}(t)$,

the change in X_j due to the change $\Delta X_{ij}(t)$ in X_i. Then let c_{ij} be the certainty of the impact of X_i on X_j; let $c(X_i(0))$ be the certainty of the initial values of the X_i; and assume that pulses have a certainty of 1.0. Then the certainty of the change $\Delta_i X_j$ in X_j due to $\Delta X_{ij}(t)$ may be calculated in accordance with the Lukasiewicz $\rightarrow$ by

$$c(\Delta_i X_j(t+1))=\max[0,c(\Delta X_{ij}(t))+c_{ij}-1].$$

To obtain next a certainty for the net impact of all the changes in the X_i—i.e., the quantity $\Delta X_j(t+1)=\sum_{i=1}^{n} \Delta_i X_j(t+1)$—compute a weighted average of certainties by

$$c(\Delta X_j(t+1)) = \frac{\sum_{i=1}^{n} c(\Delta_i X_j(t+1)) \cdot \left| \Delta_i X_j(t+1) \right|}{\sum_{i=1}^{n} \left| \Delta_i X_j(t+1) \right|}$$

The rationale for this calculation is that the certainty of a sum of two changes in X_j is partly dependent on the sizes of those changes, thus ruling out a straightforward application of the simple logical conjunction (i.e, saying that the certainty of the sum is the "and" of the certainties of the summands, where "and" is interpreted as the arithmetic min). For example, if a change of 10 units has certainty 0.9, and a change of -2 units has certainty 0.2, then the weighted averaging technique yields a net change of 8 units with net certainty 0.8, which is much more reasonable than a net certainty of only 0.2. To distinguish this weighted average interpretation for the logical "and," we here use the term balanced conjunction, the formal details of which will be given in §6.3.

The same rationale may finally be applied again to obtain

$$c(X_j(t+1)) = \frac{c(X_j(t)) \cdot |X_j(t)| + c(\Delta X_j(t+1)) \cdot |\Delta X_j(t+1)|}{|X_j(t)| + |\Delta X_j(t+1)|},$$

which completes the certainty calculates for the SPIN computed value $X_j(t+1)$.

The acronym SPIN[+] refers to the modification of SPIN which is obtained by including all of these computations. To test the kind of output that one can obtain using SPIN[+], a small FORTRAN program was written for a model using five variables, with a forecast covering 30 time units. The data used was as shown in Figure 6-1. All variables were assumed to have an initial value of 50 units, to be bounded by $0 \leq X_i \leq 100$, and to have initial certainty $c(X_i(0))=1.0$. A pulse of 1 unit was applied to each of X_3 (use of mass transit) and X_4 (use of autos).

The output at selected time intervals is shown in Figure 6-2. The five variables fluctuate in accordance with the original version of SPIN, while certainties degrade, more rapidly for some variables than for others. Most importantly, further runs of the same program showed that adjusting the certainties of the impact relations in the model has the effect described in §6.1. Certainty of forecasted events degrades slowly if the certainty in the model is high, and it degrades rapidly if the certainty of impacts is low. Thus is established a firm mathematical link between the amount of confidence that may be placed in the implications of a given model and the amount of imprecision that the model is acknowledged to bear as a description of the corresponding system in the real world.

§6.3. Formalization Of SPIN[+] In A Multivalent Linguistic Theory

The motivation for formalizing a given model of a real-world system is discussed at length in Part III. In brief, where non-classical

Time	Energy Use		Hassle		Mass Transit		Auto		Conven- ience	
t	X_1	$c(X_1)$	X_2	$c(X_2)$	X_3	$c(X_3)$	X_4	$c(X_4)$	X_5	$c(X_5)$
1	50	1.00	50	1.00	50	1.00	50	1.00	50	1.00
5	50	.99	49	.99	52	.99	49	.99	51	.99
10	52	.98	46	.96	54	.98	50	.96	57	.95
15	57	.96	45	.92	61	.96	48	.92	73	.87
20	67	.92	36	.86	78	.91	38	.84	100	.77
25	61	.89	13	.72	83	.88	30	.77	100	.66
30	58	.85	3	.59	84	.85	27	.68	100	.60

X_i = parameter value calculated by SPIN

$c(x_i)$ = certainty of X_i calculated by SPIN[+]

Figure 6-2. Sample output for SPIN[+].

logics are involved, representation of the model in a semantically complete formal logical system is called for in order to obtain an explicit characterization of the underlying logic, which can then be analyzed for its consistency or coherence. In the present case of approximate reasoning, which is classical at the outer level, consistency is equivalent with coherence. Hence, if the formalization were discovered to be inconsistent, this would mean that <u>any</u> result is derivable from the given digraph, so that all model output would be devoid of credibility.

The following establishes that the logic of the SPIN$^+$ Citizen's Model is indeed consistent. At the same time it pinpoints a possible source of inconsistency in the computer implementation, namely, errors of rounding due to decimal approximation. Since such errors will in practice be negligible, however, one is hereby assured that the model will always produce logically meaningful results.

Formalization of SPIN$^+$ and the Citizen's Model requires developing a formal theory with linguistic and axiomatic capabilities adequate to express all of the calculations for an arbitrary run of the SPIN$^+$ package. This involves four main steps:

1) Starting with the minimal multivalent theory, form the theory FL of algebraic fields,

2) From FL, form the theory OF of ordered fields, thus in two steps laying down the needed properties of real-number arithmetic,

3) Develop from OF the theory SP$^+$, which includes explicit definitions of (i) the Lukasiewicz $\rightarrow$, (ii) a notion of "balanced conjunction" as mentioned in §6.2, and (iii) general calculations of forecasted parameters and certainty measures,

4) From SP^+ build the theory CM which describes the specific characteristics of the Citizen's Model--i.e., the weights, delays, and certainties of impacts, together with initial certainties and values of the model parameters.

In order to develop the theory FL it is necessary first to extend the basic definition of a linguistic theory, to allow for the formal representation of functions defined on the individuals of a semantic interpretation. The necessary modifications are:

1) proper symbols now include n-ary function symbols for any finite n, with general denotations f,g,h, etc.,

2) individual terms include strings of the form $f(t_1,...,t_n)$, where f is n-ary and the t_i are individual terms,

3) semantic interpretations I for a language L now assign a function $I(f):U_I^n \rightarrow U_i$ to each n-ary function symbol f in L.

To ensure semantic completeness, the following addition can be made to the definition of the canonical interpretation I_Δ given in §4.4: for any n-ary function symbol f in L(T), assign the function $I_\Delta(f)$ defined by

$$I_\Delta(f)([t_1],...,[t_n])=[t] \text{ iff } [f(t_1,...,t_n)=t]\varepsilon\Delta.$$

This function is well-defined, and leads to the desired completeness result.

Now let F be the minimal linguistic theory. To form the theory FL, adjoin to L(F) three individual constants, denoted by 0,1, and -1, two binary function symbols, denoted by + and ·, and one unary function symbol, denoted by -1. The terms +(t,t'), ·(t,t'), and -1(t) will be written t+t', t·t', and t^{-1}. Next, where x,y, and z are arbitrary individual variables, adjoin the following formulas as proper axioms:

FL1. (x+y)+z=x+(y+z)

FL2. x+0=x

FL3. x+(-1·x)=0

FL4. x+y=y+x

FL5. (x·y)·z=x·(y·z)

FL6. x·1=x

FL7. x·y=y·x

FL8. x·(y+z)=x·y+x·z

FL9. y=x^{-1} ≡ x·y=1

FL10. ⌐(0=1).

This axiomatization follows Shoenfield (1967, p. 70) with the exception that the concept of multiplicative inverse is here introduced in terms of a unary function, -1, rather than by the existential axiom

$$\neg(x=0) \supset \exists y(x \cdot y = 1).$$

It can be shown that the two approaches are equivalent, i.e., in either case multiplicative inverses are properly defined. To simplify notations, assume that FL furthermore contains two binary function symbols denoted by - and / having the defining axioms:

FL11. x-y=x+(-1·y)

FL12. x/y=x·y^{-1}.

Readability may be enhanced by writing terms of the form (x-y)-z as x-y-z.

To define the theory OF of ordered field, adjoin a binary fuzzy relation symbol, denoted by <, and make it non-fuzzy by the axiom

OF1. x<y≅$\kappa_{0.0}$ ∨ x<y≅$\kappa_{1.0}$,

where $\kappa_{0.0}$ and $\kappa_{1.0}$ are the zero-ary relation symbols defined in §3.1.1. For notational convenience, formulas of the form t<t'≅$\kappa_{1.0}$ will be

written as t<t'. Then the axioms specific to ordered fields may be written exactly as in Shoenfield (1967, p. 87):

OF2. ⌐(x<x)

OF3. x<y⊃(y<z⊃x<z)

OF4. x<y∨x=y∨y<x

OF5. x<y⊃x+z<y+z

OF6. 0<x⊃(0<y⊃0<x·y).

The theory OF will furthermore be assumed to contain a binary relation symbol $\leq$ (with notational convention similar as for <), a unary function symbol | |, and two binary function symbols max[ ] and min[ ], defined by:

OF7. x$\leq$y ≡ x<y∨x=y

OF8. (|x|=x ≡ 0$\leq$x)&(|x|=-1·x ≡ x<0)

OF9. (max[x,y]=x ≡ y$\leq$x)&(max[x,y]=y ≡ x$\leq$y)

OF10. (min[x,y]=x ≡ x$\leq$y)&(min[x,y]=y ≡ y$\leq$x).

We are now in a position to consider SP^+. Adjoin to OF: (i) for each rε[0,1], the zero-ary relation symbol κ_r defined in §3.1.1, (ii) for each rε[0,1], an individual constant denoted by a_r, (iii) the defining axiom SP^+1. $a_{0.00}=0$ & $a_{1.00}=1$,

and (iv) for each pair r,r'ε[0,1] such that r<r', the axiom

SP^+2. $a_r<a_{r'}$.

Axioms SP^+1 and SP^+2, together with the axioms for FL and OF, ensure that the constants a_r serve as formal representatives of the truth values in [0,1].

In terms of the a_r and κ_r, two operators, denoted K and V, may be introduced by adjoining, for each rε[0,1], the axiom

SP^+3. $K(a_r)\cong\kappa_r$

together with all formulas of the form

SP^+4. $V(p)=a_r \equiv p\cong\kappa_r$,

where p is a linguistic assertion. Although these operators are not

formally accommodated by the present definition of linguistic theory,

their use here causes no difficulty. Semantic completeness is unaffected

since the behaviors of K and V are explicitly defined in terms of the a_r

and κ_r.

Given these operators, the Lukasiewicz $\rightarrow$ has the explicit definition

SP^+5. $p\rightarrow q \cong K(\min[1,1-V(p)+V(q)])$.

Assurance that there is a unique certainty value for q, given the certain-

ties of p and $p\rightarrow q$, can be gained by adjoining

SP^+6. $q \cong K(\max[0,V(p)+V(p\rightarrow q)-1])$.

In order to discuss certainties of model parameters within SP^+ it

is necessary to have, for an arbitrary SPIN calculation of X_i as having

a value b, a means of expressing the degree of certainty that $SPIN^+$

assigns to the proposition that $X_i=b$. To this end, we may introduce a

binary <u>certainty</u> <u>of</u> <u>equality</u> fuzzy relation symbol, denoted by $\hat{=}$ (and

writing $t\hat{=}t'$ for $\hat{=}(t,t')$). A formula of the form $x\hat{=}a_r \cong \kappa_{0.50}$, for

example, expresses the proposition that the variable x assumes the value

r with a certainty of 0.50. It is useful to relate certainty of equality

to ordinary equality by

SP^+7. $x\hat{=}y\cong\kappa_{1.0}\supset x=y$.

The aforementioned <u>balanced</u> <u>conjunction</u>, denoted by , may now be

defined for a finite sequence of certainty of equality relations as follows:

SP^+8. $x_1\hat{=}y_1\tilde{\wedge}\cdots\tilde{\wedge}x_n\hat{=}y_n \cong K([V(x_1\hat{=}y_1)\cdot|y_1|+\cdots+V(x_n\hat{=}y_n)\cdot|y_n|]$

$\qquad\qquad\qquad \cdot[|y_1|+\cdots+|y_n|]^{-1})$.

The manner in which this is used for dealing with cumulative impacts will be seen below.

The next step in the definition of SP^+ is to introduce a formal representation of the time parameter. At this stage SP^+ becomes a theory of the kind described in §5.2, having the universes of discourse for its semantic interpretations being pairs of universes, one of which in this case is understood as a set of times. To this end adjoin: (i) a new set of individual variables $\tau_1, \tau_2, \ldots$, with general denotations τ, τ', etc., (ii) for each $s=0,1,2,\ldots,M$ (some finite M) an individual constant denoted by b_s, (iii) for each individual constant, function symbol, and relation symbol of OF (which contains FL), a corresponding symbol bearing the subscript τ, (iv) the corresponding analogues of all the proper axioms of OF, but excluding FL9, and (v) the axioms

$SP^+9.$ $b_0 = 0_\tau$

$SP^+10.$ $b_{s+1} = b_s +_\tau 1_\tau$ (for each $s=1,\ldots,M-1$).

These axioms ensure that the constants b_s behave formally as a set of times which obey the necessary laws of real arithmetic. The subscript τ will be omitted when the intended meaning is clear.

Now, for each $i=1,\ldots,N$ (some finite N), adjoin a unary function symbol, denoted by X_i, and for each pair $i,j=1,\ldots,N$, adjoin two unary function symbols denoted by DX_{ij} and D_iX_j. The expressions $X_i(\tau)$, $DX_{ij}(\tau)$, and $D_iX_j(\tau)$ will represent the SPIN computed values at time τ of, respectively, the parameter X_i, the change in X_i since the time $\tau'=\max[\tau-1-\text{delay}_{ij}]$, and the change in X_j due to the most recent change DX_{ij}.

The general notion of a weight of impact is introduced by adjoining, for each pair $i,j=1,\ldots,N$, an individual constant denoted by a_{ij}. To represent delays of impacts, similarly adjoin a (time) constant b_{ij} for each pair i,j. Then the SPIN computations are described by adjoining, for each pair i,j, the axioms

SP^+11. $DX_{ij}(\tau)=X_i(\tau)-X_i(\max[0,\tau-1-b_{ij}])$

SP^+12. $D_iX_j(\tau+1)=a_{ij}\cdot DX_{ij}(\tau)$

SP^+13. $DX_j(\tau+1)=D_1X_j(\tau+1)+\cdots+D_NX_j(\tau+1)$

SP^+14. $X_j(\tau+1)=X_j(\tau)+DX_j(\tau+1)$.

Proceeding to the $SPIN^+$ calculations, it will be convenient to have the following abbreviations:

$\delta_{ij}(\tau)$ for $DX_{ij}(\tau)\triangleq X_i(\max[0,\tau-1-b_{ij}])$

$\delta'_{ij}(\tau+1)$ for $D_iX_j(\tau+1)\triangleq a_{ij}\cdot DX_{ij}(\tau)$

$\delta_i(\tau+1)$ for $DX_j(\tau+1)\triangleq D_1X_j(\tau+1)+\cdots+D_NX_j(\tau+1)$

$\gamma_j(\tau+1)$ for $X_j(\tau+1)\triangleq X_j(\tau)+DX_j(\tau+1)$.

Next let $c_{ij}\varepsilon[0,1]$ be the certainty of the impact of X_i on X_j. Then the certainty computations of $SPIN^+$ are defined by adjoining, for each pair i,j, the axioms

SP^+15. $\delta_{ij}(\tau)\rightarrow\delta_{ij}'(\tau+1)\cong\kappa_{c_{ij}}$

SP^+16. $\delta_j(\tau+1)\cong\delta_{1j}'(\tau+1)\tilde{\wedge}\cdots\tilde{\wedge}\delta'_{Nj}(\tau+1)$

SP^+17. $\gamma_j(\tau+1)\cong\gamma_j(\tau)\tilde{\wedge}\delta_j(\tau+1)$.

The use of these axioms in SP^+ has the effect that, given a specific model (such as the Citizen's Model formalized below), the certainty of the value of X_j at time t, $X_j(t)$, will be the truth value of the linguistic assertion $\gamma_j(b_t)$.

To formalize the Citizen's Model in a theory CM: let M=30, the number months used in §6.2; let N=5, the number of model parameters; for each pair $i,j=1,\ldots,5$, let $c_{ij}\varepsilon[0,1]$ be the certainties of impacts, with $c_{ij}=0$ if no impact was given; for each pair i,j, adjoin the axiom

CM1. $a_{ij}=a_{w_{ij}}$,

where $w_{ij}\varepsilon[0,1]$ is the weight of impact, with $w_{ij}=0$ if no impact was given; and, for each pair i,j, adjoin the axiom

CM2. $b_{ij}=b_{d_{ij}}$,

where d_{ij} is the delay of impact, with $d_{ij}=0$ if no impact was given.

For convenience, the range $[0,100]$ of the model parameters will be compressed to $[0,1]$. This allows the constants a_r to also be used as specific values of the $X_i(\tau)$. Thus the parameters are initialized by adjoining, for each $i=1,\ldots,5$, the axiom

CM3. $X_i(0)\hat{=}a_{0.50}\cong\kappa_{1.0}$.

There remains only the specification of pulses. Following the example of §6.2, this may be accomplished by adjoining the five axioms

CM4. $X_1(b_1)\hat{=}X_1(b_0)+a_{0.00}\cong\kappa_{1.0}$

CM5. $X_2(b_1)\hat{=}X_2(b_0)+a_{0.00}\cong\kappa_{1.0}$

CM6. $X_3(b_1)\hat{=}X_3(b_0)+a_{0.01}\cong\kappa_{1.0}$

CM7. $X_4(b_1)\hat{=}X_4(b_0)+a_{0.01}\cong\kappa_{1.0}$

CM8. $X_5(b_1)\hat{=}X_5(b_0)+a_{0.00}\cong\kappa_{1.0}$

These also provide initial values to the SPIN computations, by axiom SP^+7.

It is easily verified that the Citizen's Model of §6.2 constitutes a semantic model of the theory CM. The relation symbol $<$ can be interpreted

as the usual ordering of the reals, and the function symbols +, ·, and -1 can be interpreted as the usual addition, multiplication, and inversion of reals. Thus, by Proposition 3.4.2-2, CM is consistent.

CM is not categorical, however, since it is known that the theory OF is not categorical. Thus CM has semantic models which are not iso-morphic with the Citizen's Model. On the other hand, since all of the functions and relations of CM, other than t, ·, -1, and <, are explicitly defined in terms of the individual constants a_r and b_s, nonisomorphic models will always agree at least on the interval [0,1] and for the times from 0 to 30. Since CM clearly encodes all of the computations of the computer run described in §6.2, it follows by semantic completeness that a formula of L(CM) which expresses a property of the X_i--within the given ranges of values, certainties, and times--will be a theorem of CM if and only if it is generated by the SPIN$^+$ computer run of the Citizen's Model with the given initial values. This of course excepts minor discrepancies due to CM's being infinitary, while the computer deals with decimal approximations. Since such discrepancies will in practice be negligible, we may in any case conclude that CM is a correct characteriza-tion of the logic of the Citizen's Model, and hence that the model output will in general always be logically reliable.

§6.4. Future Directions

This chapter shows how a reasonable rendering of certainty account-ing can be built into the SPIN modeling package through some fairly simple modifications, and it may be expected that similar modifications can be developed for most of the structural modeling packages in use today. The formalized logic of approximate reasoning is generally

applicable to any "A impacts B" relationship, or for that matter, any transitive relation such as "A is preferred to B," which occurs in applications of ISM (see Linstone, et. al., 1978).

At the same time, this work is properly regarded as only an initial step toward a more sophisticated approach which would use fuzzy logic per se as described in §5.3, and as recently extended to the possibilistic logic described in §5.4. This would invoke "Certainty" as a linguistic variable having linguistic terms such as "certain," "uncertain," "very certain," "absolutely certain," "more or less certain," and so on, and having its universe of discourse be the interval [0,1]. In other words "Certainty" may be developed exactly as Zadeh's linguistic "Truth," which has its truth values being fuzzy subsets of [0,1].

This logic has now been axiomatized (§5.3), but the entirety of the mathematics for full implementation as an enhancement to SPIN has yet to be worked out. Partial steps in this direction are available from the work of Dubois and Prade (1979), which provides the needed generalization of the Lukasiewicz →. The main task remaining is a corresponding generalization of the balanced conjunction defined in §6.3.

The advantages of using fuzzy logic, rather than the multivalent logic, are two-fold:

1) user-machine interaction can be carried out exclusively with the qualitative linguistic terms in the term set for certainty--i.e., both for input certainty values and as output--thus approaching a natural language conversation,

2) in addition to the qualitative values given by the linguistic term set, one also has a measure of the precision of those values.

The latter may conceivably add a further dimension to certainty accounting, since it is known (Dubois and Prade 1979) that in sequences of inferences in fuzzy logic, the consequents tend both to become less true (i.e., less certain) and to become less precise.

PART II

LOGICS OF FORM AND SELF-REFERENCE

CHAPTER 7

THE LAWS OF FORM

The aim in this chapter is to show that the "primary algebra" of Spencer-Brown's <u>Laws of Form</u> (1969) is essentially isomorphic with the axiomatized classical propositional calculus.

In §6.1 the classical propositional calculus is presented as a specific logic (cf. §2.1.3) consisting of a formal language L(PC) together with a semantics $\Sigma(PC)$. The formal system PC is then obtained by adopting the axioms and inference rules of Hilbert and Ackerman (1928). This axiomatization is known to be a consistent and semantically complete characterization of the classical logic.

In §6.2 we define an extension PC* of PC, and then expand $\Sigma(PC)$ to an appropriate semantics $\Sigma(PC^*)$ for L(PC*). PC* is a consistent and semantically complete inessential extension of PC.

In §6.3 we rewrite Spencer-Brown's primary algebra as a formal system PA having a semantics $\Sigma(PA)$. Brown's work established that PA is consistent and semantically complete with respect to $\Sigma(PA)$.

In §6.4 the latter result is used to show that PA is isomorphic with PC* in the general sense that exactly the same English language assertions can be expressed in both L(PA) and L(PC*) and, of those assertions, exactly the same ones can be formally derived as theorems of both PA and PC*. This in turn leads to a proof that "equality" in PA is isomorphic with "logical equivalence" in PC*.

§6.5 reviews and offers new proofs of some earlier results of Spencer-Brown (1969) and Orchard (1975) concerning the same questions as addressed here.

As motivation for the development in this chapter, an outline of the isomorphism theorem is provided in Figure 7-1. The main steps are as follows. First, a translation τ is defined. τ is a one-to-one mapping of expressions of L(PA) into propositions of L(PC*). Next, Σ(PA) is shown to be identical with Σ(PC*) in the sense that an equation E=⌐ is valid in Σ(PA) if and only if the proposition τ(E) is tautology of Σ(PC*). One then has the following simple argument:

PA⊢E=⌐ iff E=⌐ is valid in Σ(PA)

(completeness of PA and Proposition 1.1.6-1)

iff τ(E) is a tautology of Σ(PC*)

(by the result mentioned above)

iff PC*⊢τ(E)

(completeness of PC* and Proposition 1.1.6-1).

The task in the first three sections is mainly to describe the systems PC* and PA with sufficient precision so that the translation τ can be provided a rigorously explicit definition.

§7.1. The System PC

As symbols for a language L(PC) select a countably infinite set from the collection which was assumed to exist in §2.1.1. Let two of these symbols be denoted by ⌐ and V, to serve as logical connectives-- respectively, negation and disjunction. Let two further symbols be denoted by (and), to serve formally as parentheses. Assume that the
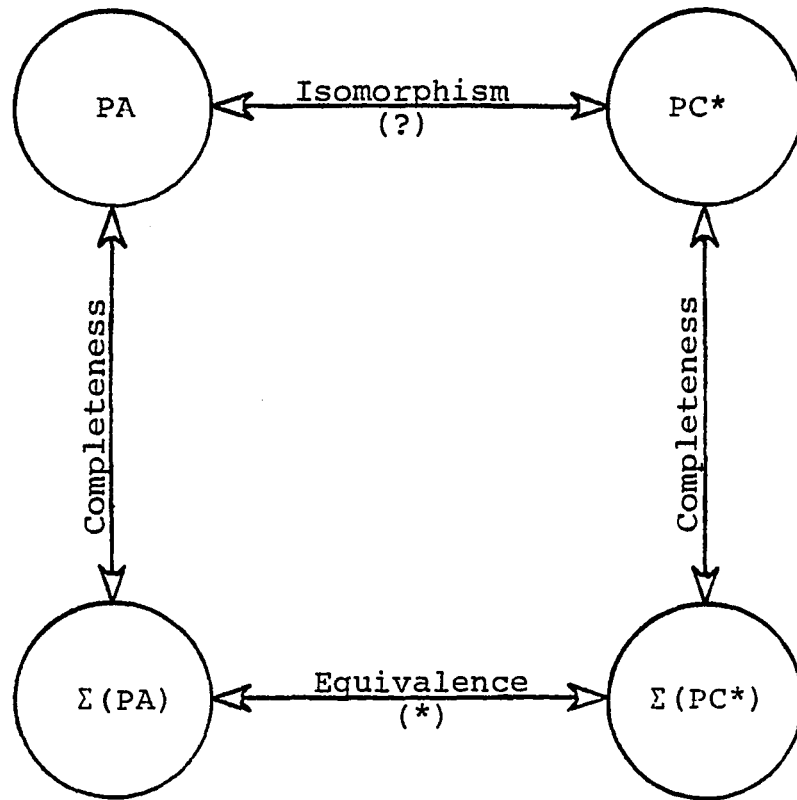
Figure 7-1.   Isomorphism of PC* and PA.

other, infinitely many, symbols of L(PC) are denoted by $p_1, p_2, \ldots$, and

serve as propositional variables.

The propositions of L(PC) are defined inductively as follows:

1. The propositional variables are propositions of L(PC),

2. If a string P is a proposition of L(PC), then the string of the
   form ¬P is a proposition of L(PC),

3. If two strings P and Q are propositions of L(PC), then the string of
   the form (PVQ) is a proposition of L(PC).

It is convenient to introduce other logical connectives--&
(conjunction), ⊃(logical implication), and ≡ (logical equivalence)--for
purposes of abbreviating complex formulas, as follows: P&Q abbreviates
¬(¬PV¬Q); P⊃Q abbreviates ¬PVQ; P≡Q abbreviates (P⊃Q)&(Q⊃P). The use of
parentheses is necessary for distinguishing between such formulas as
¬P&Q and ¬(P&Q). On the other hand, for writing out formulas of L(PC),
parentheses may be dropped when the intended grouping is clear.

Now define a semantics Σ(PC) for L(PC) as follows. Let the truth
values be the numbers 1 and 0--with 1 standing for "true" and 0 standing
for "false." Let the truth valuations (or truth-value assignments) be
the mappings V of formulas of L(PC) into {1,0}, such that:

1. For each propositional variable p_i, $V(p_i)\varepsilon\{1,0\}$,

2. For each proposition of the form ¬P, $V(¬P) = 1-V(P)$,

3. For each formula of the form PVQ, $V(PVQ) = \max[V(P),V(Q)]$.

It is easy to see that Σ(PC) encodes the usual truth-tables for ¬ and
V. Furthermore, each V in Σ(PC) acts on propositions abbreviated by
means of &, ⊃, and ≡, in the expected ways; for example, V(P&Q)=1 if and
only if both V(P)=1 and V(Q)=1. A proposition P of L(PC) is a tautology
of Σ(PC) if V(P)=1 for all truth-valuations V of Σ(PC).

By <u>classical</u> <u>propositional</u> <u>calculus</u> is meant the logic which is represented by the pair L(PC),Σ(PC). In this dissertation, the term <u>propositional</u> <u>calculus</u> shall sometimes also refer to the <u>axiomatized</u> propositional calculus--i.e., a formal logical system--denoted by PC. The following axiomatization was developed by Hilbert and Ackerman (1928, cf. Rescher 1969). As <u>axioms</u>, PC has the following propositions of L(PC):

1. $(p_1 \lor p_1) \supset p_1$

2. $p_1 \supset (p_1 \lor p_2)$

3. $(p_1 \lor p_2) \supset (p_2 \lor p_1)$

4. $(p_1 \supset p_2) \supset ((p_3 \lor p_1) \supset (p_3 \lor p_2))$.

As <u>inference</u> <u>rules</u> PC has the following mappings of L(PC) into itself:

1. <u>Modus</u> <u>Ponens</u>: the binary rule consisting of all triples of the form $(P, P \supset Q, Q)$ where P and Q are propositions of L(PC),

2. <u>(Uniform)</u> <u>Substitution</u>: the unary rule consisting of all pairs of the form $(P, P(Q/p_i))$, where Q is a proposition of L(PC), p_i is a propositional variable, and the notation $P(Q/p_i)$ is as defined in §2.1.1.

The <u>theorems</u> of PC are generated from the axioms by means of the inference rules, in the manner described in §2.1.4.3.

Hilbert and Ackerman's work establishes that PC is semantically <u>complete</u> with respect to Σ(PC). A <u>contradiction</u> of L(PC) is a proposition of the form $P \& \neg P$. To see that PC is <u>consistent</u> observe that no contradiction of L(PC) can be a tautology of Σ(PC). Hence no contradiction can be a theorem of PC, by Proposition 2.1.6-1.

§7.2. The Inessential Extension PC*

Let $\underline{F}$ denote a symbol which is not in the symbol set for L(PC). An extension PC* of PC can be obtained as follows:

1. Adjoin the new __proposition symbol__ $\underline{F}$ to the set of symbols of L(PC), as a new __proposition__, thereby generating a larger set of propositions,

2. Adjoin the proposition $\underline{F} \equiv p_1 \& \neg p_1$ to the set of axioms of PC, thereby generating a larger set of theorems.

The semantics $\Sigma(PC)$ can be expanded to a __semantics__ $\Sigma(PC^*)$ for L(PC*) by extending the domain of each valuation V in $\Sigma(PC)$ by the additional requirement that $V(\underline{F})=0$. Thus the symbol $\underline{F}$ may be considered semantically as standing for "falsehood," and syntactically as being a "name" for the contradiction $p_1 \& \neg p_1$, which name has been introduced by means of a "defining axiom."

__Contradiction__ for L(PC*) is defined similarly as for L(PC); and the same simple technique discussed in §7.1 may be used to show that PC* is __consistent__.

__Proposition 7.2.-1__. The system PC* is complete with respect to $\Sigma(PC^*)$.

__Proof__. Let P be a proposition of L(PC*) which is a tautology of $\Sigma(PC^*)$. It is desired to show that PC*⊢P. Since $V(\underline{F})=V(p_1 \& \neg p_1)=0$ for all truth valuations V of $\Sigma(PC^*)$, the proposition $P(p_1 \& \neg p_1/\underline{F})$--which is obtained from P by replacing each occurrence of $\underline{F}$ with an occurrence of $p_1 \& \neg p_1$--must be a tautology of $\Sigma(PC^*)$. But $P(p_1 \& \neg p_1/\underline{F})$ is a proposition of L(PC), whence it follows that it must be a tautology of $\Sigma(PC)$. Then PC⊢$P(p_1 \& \neg p_1/\underline{F})$ by the completeness of PC with respect to $\Sigma(PC)$. Since the theorems of PC are a subset of the theorems of PC*, this means that PC*⊢$P(p_1 \& \neg p_1/\underline{F})$. From this it may be established, by induction on the

length of propositions of L(PC*), that PC*⊢P. An example of a proof by induction on the length of strings appears in Theorem 7.4-1. □

Proposition 7.2-2. The system PC* is an inessential extension of PC.

Proof. Let P be a proposition of L(PC). It is desired to show that, if PC*⊢P, then PC⊢P (cf. §2.1.5.3). For a proof by contradiction, assume that PC⊬P. Then, by the completeness of PC with respect to Σ(PC), P is not a tautology of Σ(PC). It follows that P is not a tautology of Σ(PC*). Hence PC*⊬P, by Proposition 1.1.6-1. □

§7.3. The System PA

Spencer-Brown's <u>Laws of form</u> (1969) develops an equation calculus in which logical reasoning is connected to a few simple rules governing the manipulation of some elementary conceptual "forms." The work begins with the concept of a "distinction" in a uniform conceptual space. One side of the distinction is termed the "marked state" and is indicated by a "token," ⌐; the other side of the distinction is termed the "unmarked state" and is indicated by the absence of a token, i.e., a blank space. A "primary arithmetic" is devised, which has two basic axioms:

$$\neg\,\neg\ =\ \neg \qquad \text{(Law of Calling)}$$
$$\overline{\neg}\ =\ \qquad \text{(Law of Crossing)}.$$

The Law of Calling may be interpreted by the statement: To indicate (or "call") the marked state twice is the same as to indicate it once. The Law of Crossing may be interpreted by: To cross the boundary of the distinction twice is the same as to not cross at all.

Expressions in the primary arithmetic can be defined formally by: (i) the token and the blank space are expressions, (ii) if E is an

expression, then $\overline{E}$ is an expression, and (iii) if E and F are expressions, then the concatenation EF is an expression. Spencer-Brown's Theorem 3 (1969, p. 14) establishes that the above two laws, together with a few intuitively plausible "canons," enable one to reduce any expression of the arithmetic to either the token or the blank space. Thus every expression of the arithmetic indicates exactly one of the "marked" or "unmarked" states.

When reducing expressions of the arithmetic by means of the two laws, it is evident that the token has two different interpretations, depending on whether it "covers" a blank space or some other (non-blank) expression--to wit, the expression $\overline{}$ is always interpreted as the "marked state," while an expression of the form $\overline{E}$ is interpreted as the state that one arrives at by crossing from the state that is indicated by E. These two interpretations may be united, however, by interpreting as the state one arrives at by crossing from the unmarked state--namely the marked state. In other words, the "crossing" interpretation serves the purpose for all expressions.

Given the primary arithmetic, Spencer-Brown next generalized by introducing expression variables into the language, and by formalizing "equality" of expressions so that one obtains equations like $\overline{\overline{e}\,\overline{f}}$=g and $\overline{e}$= . This leads to a "primary algebra" having two "initials" and two "rules" (see the following). An equation of the algebra is termed "valid" if it is true in the primary arithmetic for all possible (uniform) substitutions of arithmetical expressions into that equation in place of the variables. Spencer-Brown's principal result is that the primary algebra is complete under this notion of validity. Because of the above-mentioned Theorem 3, it is evident that this in effect is

completeness with respect to a two valued semantics, having as its values the marked and the unmarked states.

The primary algebra will here be represented as a formal logical system in the sense of §2.1.4. To this end, several mathematical objects and properties left unwritten or simply assumed in the primary algebra must here be made formally explicit. These are discussed in what follows.

As symbols for a language L(PA), select a countably infinite set of symbols which are different from those used for L(PC*). Let two of these symbols be denoted by [and], to serve formally as square brackets; let two further symbols be denoted by (and), to serve as parentheses; let a fifth symbol be denoted by =, to serve as an equality symbol; and let a sixth symbol be denoted by ε, serving to formalize the notion of blank space. The infinitely many remaining symbols will be denoted by $e_1, e_2, \ldots$, and will be referred to as expression variables. The language L(PA) has expressions, defined inductively by:

1. Expression variables are expressions of L(PA),

2. The symbol ε is an expression of L(PA),

3. If a string E is an expression of L(PA), then the string of the form [E] is an expression of L(PA),

4. If two strings E and F are expressions of L(PA), then the string of the form (EF) is an expression of L(PA).

In writing out expressions, parentheses may be dropped if the symbol groupings are clear or irrelevant; but square brackets may not be dropped. Brackets are here being used as logical connectives, and therefore are always essential to the syntactical form of an expression.

An equation of L(PA) is a string of the form E=F, where E and F are expressions of L(PA).

The formal logical system PA has as <u>axioms</u> the following five equations of L(PA):

1. $[[e_1]e_1]=\varepsilon$

2. $[[e_1e_3][e_2e_3]]=[[e_1][e_2]]e_3$

3. $e_1\varepsilon=e_1$

4. $e_1e_2=e_2e_1$

5. $e_1(e_2e_3)=(e_1e_2)e_3.$

The <u>inference rules</u> of PA are the following mappings of L(PA) into itself:

1. (<u>Not-Necessarily-Uniform</u>) <u>Substitution</u>: all pairs of the form $(E=F,G=G^*)$, where $E=F$ is an equation of L(PA), G is an expression of L(PA), and G^* is an expression of L(PA) that is obtained from G by substituting an occurrence of F for each of one or more (but not necessarily all) occurrences of E in G,

2. (<u>Uniform</u>) <u>Replacement</u>: all pairs of the form $(E=F,(E=F)(G/e_i))$, where G is an expression of L(PA), e_i is an expression variable, and $(E=F)(G/e_i)$ is as defined in §2.1.1,

3. <u>Symmetry of Equality</u>: all pairs of the form $(E=F,F=E)$, where E and F are expressions of L(PA),

4. <u>Transitivity of Equality</u>: all triples of the form $(E=F,F=G,E=G)$, where E, F, and G are expressions of L(PA).

The system LOF has a set of <u>theorems</u> as defined in §2.1.4.3.

The manner in which LOF represents Spencer-Brown's primary algebra may now be summarized as follows:

1) The blank space is here replaced by a specific symbol, denoted by the Greek letter ε,

2) In order to have expressions be strings--instead of two dimensional arrays made up of tokens and blank spaces--the token ⌐, is here replaced by the pair of square brackets, [and],

3) On account of item 1, the token ⌐, thought of as "covering" a blank space, is now written as [ε],

4) In the present notation, the above axioms 1 and 2 of PA are Spencer-Brown's "initials" of the primary algebra, entitled "position" and "transposition"; and the above inference rules 1 and 2 are Spencer-Brown's rules of "substitution" and "replacement,"

5) Axiom 3 encodes the use of the blank space, which in Laws of form (p. 15) is expressed by the "Rule of Dominance" and, briefly stated, says that "marked" dominates "unmarked",

6) Commutativity and associativity of concatenation are here formalized, by means of axioms 4 and 5,

7) Parentheses are introduced for expressing the associativity property of concatenation, in axiom 5,

8) Symmetry and transitivity of equality are here formalized, by inference rules 3 and 4.

From these notes it can be seen that exactly the same equations--under transliteration in accordance with the above items 1 and 2--are expressible and derivable in both the primary algebra and the formal system PA.

The semantics Σ(PA) for L(PA) has as values the letters $\underline{m}$ and $\underline{u}$, standing respectively for the marked and unmarked states. The valuations of Σ(PA) are the mappings v of expressions of L(PA) into $\{\underline{m},\underline{u}\}$ such that:

1. For each elementary variable e_i, $v(e_i)\varepsilon\{\underline{m},\underline{u}\}$,

2. $v(\varepsilon)=\underline{u}$,

3. For any expression of the form [E], v([E])=$\underline{m}$ if v(E)=$\underline{u}$; and v([E])=$\underline{u}$ if v(E)=$\underline{m}$,

4. For any expression of the form EF, v(EF)=$\underline{m}$ if either v(E)=$\underline{m}$ or v(F)=$\underline{m}$; otherwise v(EF)=$\underline{u}$.

Item 4 evidently is in keeping with the aforementioned "rule of dominance." An equation E=F of L(PA) is <u>valid</u> in Σ(PA) if v(E)=v(F) for all valuations v of Σ(PA).

An <u>inconsistency</u> of L(PA) is an equation of the form E=[E]. It can be shown that the system PA is <u>consistent</u>, by Proposition 1.1.6-1 and the same technique as discussed for PC in §7.1.

<u>Proposition 7.3-1</u>. The system PA is complete with respect to Σ(PA).

<u>Proof</u>: Spencer-Brown's completeness proof applies, since the foregoing remarks show that Σ(PA) is essentially the same semantics as used in <u>Laws of form</u>. More exactly, if an equation E=F is valid in Σ(PA), then it is also valid in Spencer-Brown's semantics, since every arithmetical expression is uniquely reducible to either ⌐ or the blank space (<u>Laws of form</u>, Theorem 3); and conversely for the same reason. □

§7.4. The Systems PA And PC* Are Isomorphic

A <u>translation</u> τ of expressions of L(PA) into propositions of L(PC*) can now be defined as follows:

1. For each index i, τ(e_i)=p_i, where e_i is the i-th expression variable of L(PA), and p_i is the i-th propositional variable of L(PC*),

2. τ(ε)=$\underline{F}$,

3. For each expression of the form [E], τ([E])= ¬τ(E),

4. For each expression of the form EF, τ(EF)= τ(E)∨τ(F).

It can be shown, by induction on the length of expressions of L(PA), that the mapping τ is well-defined. Furthermore, by induction on the length of propositions of L(PC*), τ is a one-to-one correspondence between expressions of L(PA) and propositions of L(PC*). Consequently, the inverse translation, τ^{-1}, is also well-defined.

The following theorem establishes an _isomorphism_ of PA and PC* in the precise sense that, for any expression E of L(PA), PA⊢E=[ε] if and only if PC⊢τ(E). This notion of isomorphism may be interpreted as saying that essentially the same assertions of ordinary English are expressible in both L(PA) and L(PC*), and, of those so expressible, exactly the same assertions are formally derivable in both PA and PC*. This interpretation is justified by reading τ(E) as usual for formulas of PC--i.e., simply as asserting "τ(E)"--and by reading the expression [ε] as standing for "true"--so that "E=[ε]" may be interpreted as asserting "the proposition 'E' is true".

Theorem 7.4-1. The systems PA and PC* are isomorphic.

Proof. By the discussion at the beginning of this chapter, it is sufficient to show that Σ(PA) can be identified with Σ(PC*) in the sense that for any expression E of L(LOF),

$$\Sigma(PA)\models E=[\varepsilon] \text{ iff } \Sigma(PC^*)\models\tau(E).$$

By the definition of Σ(PA), each valuation v is uniquely determined by its action on the expression variables e_i. Similarly, by the definition of Σ(PC*), each truth-value assignment V is uniquely determined by its action on the propositional variables p_i. Consequently, there is a one-to-one correspondence between the two sets of valuation mappings, given by:

$$v(e_i)=\underline{m} \text{ iff } V(p_i)=1.$$

Since $v([\varepsilon])=\underline{m}$, for all v in $\Sigma(PA)$, it now follows by the definitions of "valid in $\Sigma(PA)$" and "tautology of $\Sigma(PC^*)$" that it is sufficient to show: for every expression E of $L(PA)$, and every valuation v of $\Sigma(PA)$, if V is the truth-value assignment of $\Sigma(PC^*)$ corresponding to v, then

$$v(E)=\underline{m} \text{ iff } V(\tau(E))=1.$$

This will be accomplished by induction on the length of expressions E of $L(PA)$. Let E be any expression of $L(PA)$; and let n be the length of E.

Case 1: n=1. By definition of the expressions of $L(PA)$, there are two possibilities:

Case 1.a: E is an expression variable, e_i. Then $\tau(e_i)=p_i$, by the definition of τ; so $v(E)=\underline{m}$ iff $V(\tau(E))=1$, by virtue of the assumed correspondence between v and V.

Case 1.b: E is the symbol ε. This case does not arise, since here there is no v such that $v(E)=\underline{m}$.

Case 2: n>1. Here also there are two possibilities:

Case 2.a: E is an expression of the form [F]. Then F has length less than n, and the induction hypothesis applies, giving

$$v(F)=\underline{m} \text{ iff } V(\tau(F))=1.$$

It follows that

$$v(F)=\underline{u} \text{ iff } V(\tau(F))=0. \tag{1}$$

Thus,

$$v(E)=\underline{m} \text{ iff } v(F)=\underline{u} \qquad \text{(def. of } \Sigma(PA))$$
$$\text{iff } V(\tau(F))=0 \qquad \text{(by (1))}$$
$$\text{iff } V(\neg\tau(F))=1 \qquad \text{(def. of } \Sigma(PC^*))$$
$$\text{iff } V(\tau(E))=1 \qquad \text{(def. of } \tau)$$

Case 2.b: E is an expression of the form FG. Then each of F and G have length less than n, and the induction hypothesis provides

$$v(F)=\underline{m} \text{ iff } V(\tau(F))=1 \tag{2}$$

$$V(G)=\underline{m} \text{ iff } V(\tau(G))=1. \tag{3}$$

Thus,

$v(E)=\underline{m}$ iff either $v(F)=\underline{m}$ or $v(G)=\underline{m}$ (def. of Σ(PA)

     iff either $V(\tau(F))=1$ or $V(\tau(G))=1$ (by (2) and (3))

     iff $V(\tau(F)\vee\tau(G))=1$ (def. of Σ(PC*))

     iff $V(\tau(FG))=1$.   (def. of τ)

This completes the proof, by induction on n. □

The following theorem shows that equality in PA is isomorphic with logical equivalence in PC*.

Theorem 7.4-2. For all equations E=F of L(PA),

$$PA\vdash E=F \text{ iff } PC*\vdash\tau(E)\equiv\tau(F).$$

Proof. Let G be the expression $\tau^{-1}(\tau(E)\equiv\tau(F))$ of L(PA). By Theorem 7.4-1 and the fact that τ^{-1} is well-defined,

$$PC*\vdash\tau(E)\equiv\tau(F) \text{ iff } PA\vdash G=[E]. \tag{1}$$

By Propositions 7.2-1 and 1.1.6-1,

$$PC*\vdash\tau(E)\equiv\tau(F) \text{ iff } \Sigma(PC*)\vDash\tau(E)\equiv\tau(F). \tag{2}$$

By Propositions 7.3-1 and 1.1.6-1,

$$PA\vdash G=[E] \text{ iff } \Sigma(PA)\vDash G=[E]. \tag{3}$$

It follows by (1), (2), and (3) that it is sufficient to show

$$\Sigma(PA)\vDash G=[E] \text{ iff } \Sigma(PA)\vDash E=F. \tag{4}$$

Item (4) can be established as follows. Observe that G can be written explicitly as [[[E]F][[F]E]], by the definition of τ and the definition of the connective $\equiv$.

For the "only if" part of (4), suppose that G=[E] is valid in
$\Sigma(PA)$. Then, for all v, v(G)=$\underline{m}$. It is easy to verify that, if v(E)$\neq$v(F),
for any v, then v(G)$\neq\underline{m}$. Consequently, for all v, v(E)=v(F)--i.e., the
equation E=F is valid in $\Sigma(PA)$.

For the "if" part of (4), suppose that E=F is valid in $\Sigma(PA)$.
Then, for all valuations v, v(E)=v(F). Hence, for all v, either
v(E)=v(F)=$\underline{m}$ or v(E)=v(F)=$\underline{u}$. Here one can verify that, in either case,
v(G)=$\underline{m}$. Thus, for all v, v(G)=v([E])--i.e., the equation G=[E] is valid
in $\Sigma(PA)$. $\square$

§7.5. Review of Spencer-Brown's and Orchard's Results
Concerning Propositional Calculus

Let τ^* be the restriction of the inverse translation τ^{-1} to just
the propositions of L(PC). Then τ^* turns out to be Spencer-Brown's
"interpretation" of propositions of the propositional calculus into
expressions of the primary algebra. In the present notation, his result
concerning propositional calculus may be expressed as follows.

Proposition 7.5-1. For P any proposition of PC, PA$\vdash\tau^*$(P)=[E] iff
P is a tautology of $\Sigma(PC)$.

Proof. By Theorem 7.3-1, PA$\vdash\tau^*$(P)=[E] iff PC$^*\vdash\tau(\tau^*$(P)). But
$\tau(\tau^*$(P)) is just P. Hence

$$PA\vdash\tau^*(P)=[E] \text{ iff } PC^*\vdash P. \tag{1}$$

Since PC* is an inessential extension of PC (Proposition 7.2-2) and P is
a proposition of L(PC),

$$PC^*\vdash P \text{ iff } PC\vdash P. \tag{2}$$

By the completeness of PC with respect $\Sigma(PC)$,

$$PC\vdash P \text{ iff } P \text{ is a tautology of } \Sigma(PC). \tag{3}$$

Lines (1),(2), and (3) together imply the desired result. $\square$

A proposition P of L(PC) is an <u>antitautology</u> of Σ(PC) if V(P)=0

for all V in Σ(PC). A proposition P of L(PC) is a <u>contingency</u> of Σ(PC)

if V(P) may be either 1 or 0, depending on V. Orchard (1975) establishes

the following.

<u>Proposition 7.5-2</u>. For any proposition P of PC, (i) PA⊢τ*(P)=ε

iff P is an antitautology of Σ(PC), and (ii) neither PA⊢τ*(P)=[ε] nor

PA⊢τ*(P)=ε iff P is a contingency of Σ(PC).

<u>Proof</u>: (i) Similar to the proof of Proposition 7.5-1. (ii) It is

obvious that P is either a tautology, an antitautology, or a contingency

of Σ(PC). Thus (ii) follows by (i) and Proposition 7.5-1. □

These results established conclusively that the propositional

calculus can be embedded in the primary algebra; but they also raised

the question of why, conversely, the primary algebra cannot be embedded

in propositional calculus. This led to the isomorphism theorem of §7.4.

It was observed that the translation τ* could not be inverted to provide

a translation of expressions of the primary algebra into propositions of

the propositional calculus, simply because the language of the proposi-

tional calculus contains no symbol corresponding to Spencer-Brown's

blank space. This suggested adding a new symbol, together with a defin-

ing axiom, as was done in §7.2 to obtain PC* from PC. The fact that PC*

is an inessential extension of PC justifies the terminology of the

conclusion: the primary algebra is <u>essentially isomorphic</u> with classical

propositional calculus.

CHAPTER 8

THE CALCULUS FOR SELF-REFERENCE

The primary aim of this chapter is to establish that F. J. Varela's calculus for self-reference (Varela 1975) provides a semantically complete axiomatization of S. C. Kleene's three-valued logic of partial recursion (Kleene 1938 and 1952).

In §8.1 the calculus is represented as a formal system CSR having a semantics Σ(CSR). Varela's work establishes that CSR is complete with respect to Σ(CSR).

In §8.2, first, the three-valued truth-table system studied by Kleene is represented as a formal logic, consisting of a language L(K), together with a semantics Σ(K). This formalizes a notion of "strong equivalence" of logical propositions, denoted by $\cong$. Next is defined a translation τ of expressions of L(CSR) into propositions of L(K), which is then used to define a formal system K. The axioms and inference rules of K are just the images under τ of the axioms and inference rules of CSR. Thus, it is immediate that CSR is isomorphic with K.

§8.3 establishes that K is complete with respect to the semantics Σ(K). This proof is similar to the proof of Theorem 7.4-1.

In §8.4 it is shown that PC, PC* and PA are Aristotelean systems, while CSR and K are non-Aristotelean systems. Thus we establish that extensions of K having the form $K(P \cong \neg P)$, although inconsistent, are nevertheless coherent.

§8.5 studies the completeness and coherence of some special extensions of K. This in part corrects an apparent oversight in Varela's paper. Where that work uses E=☐ as an "abbreviation" for E=Ē⌐, it happens that these two equations are in fact formally derivable from one another in CSR. This section also establishes necessary and sufficient conditions for a simple extension to be coherent.

As motivation for the developments in this chapter, an outline of the abovementioned completeness result for K is provided in Figure 8-1. Here the primary task is to show that $\Sigma(CSR)$ is identical with $\Sigma(K)$ in the sense that an equation E=F of L(CSR) is valid in $\Sigma(CSR)$ if and only if the strong equivalence $\tau(E) \cong \tau(F)$ is valid in $\Sigma(K)$. With this result, one may then argue as follows:

$$K \vdash P \cong Q \quad \text{iff} \quad CSR \vdash \tau^{-1}(P) = \tau^{-1}(Q)$$

$$\text{(isomorphism of CSR and K)}$$

$$\text{iff} \quad \tau^{-1}(P) = \tau^{-1}(Q) \text{ is valid in } \Sigma(CSR)$$

$$\text{(completeness of CSR and Proposition 1.1.6-1)}$$

$$\text{iff} \quad P \cong Q \text{ is valid in } \Sigma(K)$$

$$\text{(by the result mentioned above)}.$$

No such axiomatization of Kleene's logic has heretofore been known.

§8.1. The System CSR

The development in Varela's "A calculus for self-reference" parallels the discussion in Laws of form. First, the notion of a marked and an unmarked state is augmented by the notion of an "autonomous state." This state may be conceptualized as the entire uniform space in which Spencer-Brown makes his "distinction," and it is indicated formally by a new sign, ☐.
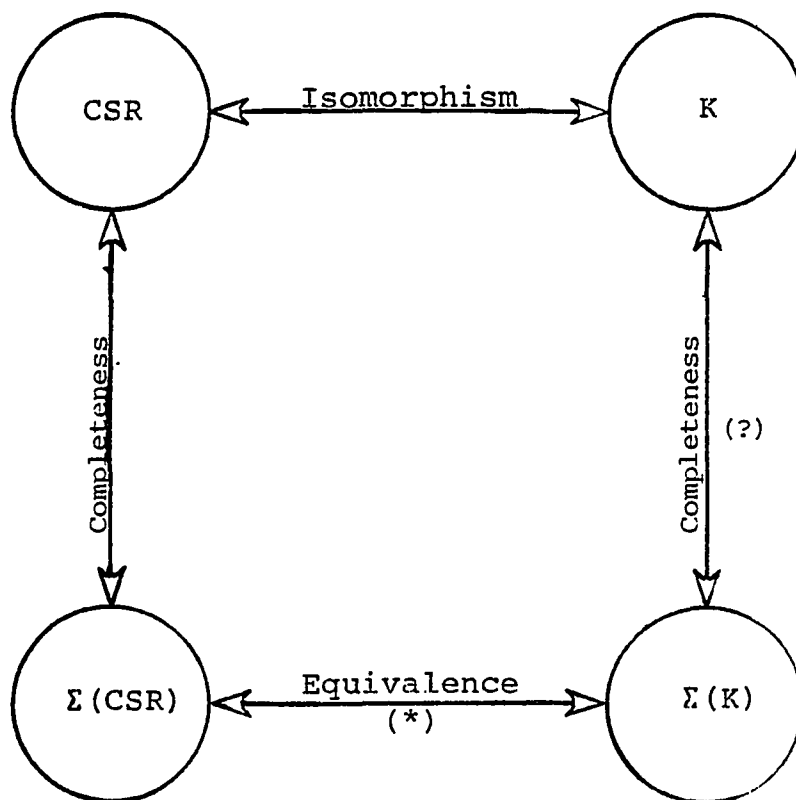
Figure 8-1. Semantic completeness of K.

This leads to an "extended arithmetic" having the following axioms:

$\boxed{e} = \daleth$, where e is either   , blank, or   (Dominance)

$\boxed{\daleth} = $ (Crossing)

$\boxed{\sqsupset} = \sqsupset$ (Reflexion)

$\sqsupset \sqsupset = \sqsupset$ (Autonomy)

which axioms may be interpreted similarly as Spencer-Brown's two "laws." For example, the third axiom (Reflexion) may be interpreted as saying: to cross from the autonomous state is to remain in the autonomous state. Furthermore, the rule of dominance is hereby modified to read: marked dominates autonomous, and autonomous dominates unmarked.

Based on these axioms, Varela established an analogue of Spencer-Brown's Theorem 3: that any expression of the extended arithmetic can be reduced to exactly one of $\daleth$, blank, or $\square$. The extended arithmetic is then generalized to form an "extended algebra," which includes a formalized equals sign, expression variables, and three axioms (represented below as axioms 1, 2, and 3). Last, equations of the extended algebra are termed "valid" in the analogous sense as equations of the primary algebra, and it is established in full detail that the extended algebra is semantically complete with respect to this notion of validity.

It follows that remarks similar to those given in §7.3 will establish that the following system CSR and semantics $\Sigma(\text{CSR})$ provides a correct formalization of the extended algebra. We may here note that the Greek letter α is now used as a replacement for the sign $\square$.

Let the formal <u>language</u> L(CSR) be the language that is obtained from L(PA) by adjoining a new symbol, denoted by α, to the symbols of L(PA) (cf. §2.1.5.3). Let the <u>semantics</u> $\Sigma(\text{CSR})$ have as <u>values</u> the

letters $\underline{m}$, $\underline{u}$, and $\underline{a}$, and as <u>valuations</u> have the mappings v of expressions of L(CSR) into $\{\underline{m},\underline{u},\underline{a}\}$ such that:

1. For each expression variable e_i, $v(e_i)\varepsilon\{\underline{m},\underline{u},\underline{a}\}$,

2. $v(\varepsilon)=\underline{u}$ and $v(\alpha)=\underline{a}$,

3. For any expression of the form [E], $v([E])=\underline{m},\underline{u}$, or $\underline{a}$, according as $v(E)=\underline{u},\underline{m}$, or $\underline{a}$,

4. For any expression of the form EF, if $v(E)\neq v(F)$, then $v(EF)$ equals whichever one of $v(E)$ and $v(F)$ is most "dominant", where $\underline{m}$ dominates $\underline{a}$, and $\underline{a}$ dominates $\underline{u}$; otherwise $v(EF)$ is the value that v assigns to both E and F.

An equation E=F of L(CSR) is <u>valid</u> in Σ(CSR) if $v(E)=v(F)$ for all valuations v of Σ(CSR).

The formal logical system CSR has as <u>axioms</u> all equations of L(CSR) having the following forms:

1. $[[E]G]E = E$

2. $[[EG][FG]] = [[E][F]]G$

3. $[E\alpha]E = E$

4. $E\varepsilon = E$

5. $EF = FE$

6. $E(FG) = (EF)G$.

As <u>inference rules</u> CSR has the following mappings of L(CSR) into itself, all defined similarly as for PA:

1. <u>(Not-Necessarily-Uniform) Substitution</u>

2. <u>Symmetry of Equality</u>

3. <u>Transitivity of Equality</u>

These axioms and inference rules generate a set of <u>theorems</u> of CSR as in §2.1.4.3. A rule of (uniform) replacement is not used in CSR since the

axioms are here defined in accordance with some axiom schemes--i.e., the above "forms" 1 through 6--which ensures a priori that all substitution instances of theorems are also theorems.

Proposition 8.1-1. The system CSR is complete with respect to Σ(CSR).

Proof. By Varela's completeness theorem and essentially the same argument as in Proposition 7.3-1. □

An inconsistency of L(CSR) shall here be defined as an equation of the form E=[E], exactly as for PA. It should be noted that this definition differs from that of Varela, who adheres more closely to Spencer-Brown in having an inconsistency be any one of the specific equations $[\varepsilon]=\alpha$, $\alpha=\varepsilon$, and $[\varepsilon]=\varepsilon$. The present definition is a much more natural one, however, especially in view of the following two propositions.

Proposition 8.1-2. The system CSR is inconsistent.

Proof. The inconsistency $\alpha=[\alpha]$ is valid in Σ(CSR). Hence CSR$\vdash\alpha=[\alpha]$ by Proposition 8.1-1. □

Proposition 8.1-3. The system CSR is coherent.

Proof. Consider any one of Varela's inconsistencies mentioned above, say $\alpha=\varepsilon$. This equation is not valid in Σ(CSR). Hence CSR$\nvdash\alpha=\varepsilon$, by Proposition 2.1.6-1. □

Thus CSR, although inconsistent, is nevertheless coherent--and is therefore a nontrivial non-Aristotelean system in the sense described in §2.1.5.4.

§8.2. Kleene's Logic And The System K

Kleene's three-valued system of truth-tables incorporates an "undefined" truth-value and a notion of "strong equivalence" between

propositions. This system is here presented as a formal logic consist-
ing of a language L(K) and a semantics Σ(K), wherein the undefined
truth-value is represented by the number $\frac{1}{2}$ and the notion of strong
equivalence is denoted, as in Kleene's work, by $\cong$. An addition to
Kleene's system is the special proposition symbol $\underline{A}$, introduced formally
below.

Let $\underline{A}$ and $\cong$ denote two symbols which are not included in the
symbol set for L(PC*). The <u>language</u> L(K) may be obtained from L(PC*) as
follows:

1. Adjoin the new symbol $\underline{A}$ as a <u>proposition</u> symbol, thereby generating
    a larger set of propositions,

2. Let the well-formed strings of L(K) be the <u>equivalences</u> of the form
    P$\cong$Q where P and Q are propositions of L(K).

The same abbreviations with &, $\supset$, and $\equiv$, apply here as for L(PC)
and L(PC*).

The <u>semantics</u> Σ(K) for L(K) may be obtained from the semantics
Σ(PC*) by:

1. Include the number $\frac{1}{2}$ in the set of <u>truth-values</u>, thereby forming the
    set $\{0,\frac{1}{2},1\}$,

2. Extend the <u>truth-valuations</u> of Σ(PC*) to the propositions of K
    according to: for all v, v($\underline{A}$)=$\frac{1}{2}$.

An equivalance E$\cong$F of L(K) is <u>valid</u> in Σ(K) if v(E)=v(F) for all
valuations v in Σ(K). This concept of validity exactly captures the
sense in which Kleene (1952) asserts that a strong equivalence of propo-
sitions is "true."

Given the language L(K), one may now define a formal system K, as
follows. Let the Greek letter τ here denote the <u>translation</u> of L(CSR)

into L(K) that is obtained by extending the translation of §7.4 according to:

$$\tau(\alpha) = \underline{A}.$$

It can be verified that this τ is one-to-one, so that the inverse mapping τ^{-1} is well-defined. Then the <u>axioms</u> and <u>inference</u> <u>rules</u> of a formal logical system K are obtained simply by applying this translation to the axioms and inference rules of CSR. This leads to an explicit description of the axioms of K as being all equivalences of L(K) having the following forms:

1. $\neg(\neg P \lor Q) \lor P \cong P$

   alternatively, $(P \supset Q) \supset P \cong P$,

2. $\neg(\neg(P \lor R) \lor \neg(Q \lor R)) \cong \neg(\neg P \lor \neg Q) \lor R$

   alternatively, $(P \lor R) \& (Q \lor R) \cong (P \& Q) \lor R$,

3. $\neg(P \lor \underline{A}) \lor P \cong P \lor \underline{A}$

   alternatively, $(P \lor \underline{A}) \supset P \cong P \lor \underline{A}$,

4. $P \lor \underline{F} \cong P$

5. $P \lor Q \cong Q \lor P$

6. $P \lor (Q \lor R) \cong (P \lor Q) \lor R$.

The inference rules of CSR may similarly be transcribed, leading to a set of <u>theorems</u> for K.

For the systems CSR and K, it is natural to say that they are <u>isomorphic</u> if the translation τ has the property that, for all equations E=F of L(CSR), CSR⊢E=F if and only if K⊢τ(E)≅τ(F). In this terminology, we have:

<u>Theorem 8.2-1</u>. The systems CSR and K are isomorphic.

<u>Proof</u>. Given the definition of K in terms of the translation τ, one can show: (i) proofs in CSR translate into proofs in K, by induction

on the length of proofs in CSR, and (ii) proofs in K translate into proofs in CSR, by induction on the length of proofs in K. □

The following results establish that K is non-Aristotelean.

<u>Proposition 8.2-2</u>. The system K is inconsistent.

<u>Proof</u>. Similarly as for CSR, Proposition 8.1-2. □

<u>Proposition 8.2-3</u>. The system K is coherent.

<u>Proof</u>. Similarly as for CSR, Proposition 8.1-3. □

§8.3. Semantic Completeness of K

This result is patterned after the isomorphism theorem of §8.4.

<u>Theorem 8.3-1</u>. The system K is complete with respect to $\Sigma(K)$.

<u>Proof</u>. Let $P \cong Q$ be any equivalence of $L(K)$. Let $E=F$ be the equation $\tau^{-1}(P)=\tau^{-1}(Q)$ of $L(CSR)$, where τ is the translation defined in §8.2. By the argument given at the beginning of this chapter, it is sufficient to show that $\Sigma(CSR)$ and $\Sigma(K)$ can be identified in the sense that

$$\Sigma(CSR) \vDash E=F \quad \text{iff} \quad \Sigma(K) \vDash P \cong Q.$$

Observe that there is a one-to-one correspondence between the valuations v in $\Sigma(CSR)$ and the truth-value assignments V of $\Sigma(K)$, given by

$$v(e_i)=\underline{m} \quad \text{iff} \quad V(p_i)=1,$$

and

$$v(e_i)=\underline{u} \quad \text{iff} \quad V(p_i)=0.$$

Here it is implicit that $v(e_i)=\underline{a}$ iff $V(p_i)=\frac{1}{2}$. It follows by this correspondence, and the definitions of "valid" for $\Sigma(CSR)$ and $\Sigma(K)$, that it is sufficient to establish: for all expressions E of $L(CSR)$, and all valuations v of $\Sigma(CSR)$,

$$v(E)=\underline{m} \quad \text{iff} \quad V(\tau(E))=1, \tag{1}$$

and

$$v(E)=\underline{u} \text{ iff } V(\tau(E))=0, \tag{2}$$

where V is the truth-value assignment of $\Sigma(K)$ corresponding to v. This may be established as follows. Let E be an expression of L(CSR), and let n be the length of E.

Case 1: n=1. There are three possibilities:

Case 1.a: E is an expression variable, e_i. In this case, both (1) and (2) hold by virtue of the stated correspondence between v and V.

Case 1.b: E is the symbol ε. On the one hand, $v(E)=\underline{u}$, for all v, by definition of $\Sigma(CSR)$; on the other hand, $\tau(E)=\underline{F}$, so that $V(\tau(E))=0$, for all V, by definition of $\Sigma(K)$. Thus (2) is established for all v and V. Assertion (1) holds for all v and V by default; there is no v such that $v(\varepsilon)=\underline{m}$, and there is no V such that $V(\underline{F})=1$.

Case 1.c: E is the symbol α. In this case, both (1) and (2) hold by default.

Case 2: n>1. Here there are two possibilities:

Case 2.a: E is an expression of the form [F]. Since the expression F has shorter length than E, the induction hypothesis provides that (1) and (2) hold for F. Thus,

$$v(E)=\underline{m} \text{ iff } v(F)=\underline{u} \qquad \text{(def. of } \Sigma(CSR))$$
$$\text{iff } V(\tau(F))=0 \qquad \text{(ind. hyp. (2))}$$
$$\text{iff } V(\neg\tau(F))=1 \qquad \text{(def. of } \Sigma(K))$$
$$\text{iff } V(\tau(E))=1 \qquad \text{(def. of } \tau)$$

which establishes (1) for E. A similar argument establishes (2).

Case 2.b: E is an expression of the form FG. In this case the induction hypothesis provides that (1) and (2) hold for both F and G. Thus,

$$v(E)=\underline{m} \text{ iff either } v(F)=\underline{m} \text{ or } v(G)=\underline{m} \qquad \text{(def. of } \Sigma(CSR))$$

$$\text{iff either } V(\tau(F))=1 \text{ or } V(\tau(G))=1 \qquad \text{(ind. hyp. (1))}$$

$$\text{iff } V(\tau(F)\lor\tau(G))=1 \qquad \text{(def. of } \Sigma(K))$$

$$\text{iff } V(\tau(E))=1 \qquad \text{(def. of } \tau)$$

which establishes (1); and

$$v(E)=\underline{u} \text{ iff both } v(F)=\underline{u} \text{ and } v(G)=\underline{u} \qquad \text{(def. of } \Sigma(CSR))$$

$$\text{iff both } V(\tau(F))=0 \text{ and } V(\tau(G))=0 \qquad \text{(ind. hyp. (2))}$$

$$\text{iff } V(\tau(F)\lor\tau(G))=0 \qquad \text{(def. of } \Sigma(K))$$

$$\text{iff } V(\tau(E))=0 \qquad \text{(def. of } \tau)$$

which establishes (2). This completes the proof, by induction on n. □

§8.4. Aristotelean Vs. Non-Aristotelean Classification Of PC, PC*, PA, CSR, And K

Classification of CSR and K has already been accomplished in the preceding sections. These results may be summarized as follows.

Proposition 8.4-1. The systems CSR and K are non-Aristotelean systems.

Proof. This fact about CSR is a consequence of Propositions 8.1-2 and 8.1-3. The proposition about K is established by Propositions 8.2-2 and 8.2-3. □

The corresponding facts about PC, PC*, and PA can be demonstrated as follows.

Proposition 8.4-2. The systems PC, PC* and PA are Aristotelean systems.

Proof. For PC: Let PC' be any inconsistent simple extension of PC. It is required to show that PC' is incoherent. Let P&¬P be an inconsistency of L(PC) such that PC'⊢P&¬P. Let Q be any proposition of L(PC'). Since PC' is a simple extension of PC, Q is also a proposition

of L(PC). Observe that the proposition (P&¬P)⊃Q is a tautology of

Σ(PC). Then PC⊢(P&¬P)⊃Q, by the completeness of PC with respect to

Σ(PC). Hence, PC'⊢(P&¬P)⊃Q. By modus ponens, it follows that PC'⊢Q.

Since Q was chosen as an arbitrary proposition of L(PC'), this shows

that PC' is incoherent.

The proof for PC* is identical to the above. Then the fact that

PA is Aristotelean may be established as a consequence of isomorphism,

Theorem 7.4-1. □

These results show that the present characterization of formal

logical systems as Aristotelean vs. non-Aristotelean is both natural and

intuitively appealing. Its strength lies in its being sufficiently

general to apply for <u>all</u> systems as defined in §2.1.

§8.5. Completeness And Coherence Of Simple Extensions

Let Γ be a set of equivalences of L(K). By the <u>semantics</u> Σ(K[Γ])

for the simple extension K[Γ] is meant the semantics that is obtained

from Σ(K) by limiting the set of valuation mappings to be the set of

only those V such that, for every equivalence P≅Q in Γ, we have

V(P)=V(Q).

<u>Proposition 8.5-1</u>. For any simple extension K' of K, K' is

complete with respect to Σ(K').

<u>Proof</u>. Let CSR' denote the inverse image of K' under the transla-

tion τ of §8.2; and let a semantics Σ(CSR') be defined analagously as

the semantics Σ(K'). By going back to Varela's original proof (Varela

1975), which was used in §8.1 to affirm that K is complete with respect

to Σ(K), and everywhere replacing the word "true" with the phrase "valid

in Σ(CSR')," one obtains a proof that CSR' is complete with respect to

$\Sigma(CSR')$. Then the methods of §8.3 apply to show that K' is complete with respect to $\Sigma(K')$. □

The above proof makes tacit use of the fact that Varela's completeness theorem can be generalized to the statement: If the inference "from E=E' infer F=F'" is valid in the extended arithmetic, then that inference is formally derivable within the extended algebra. Proposition 8.5-1 leads to a formal connection between paradox and autonomy in K as follows.

<u>Proposition 8.5-2</u>. For all formulas P of L(K),

$$K[P \cong P] \vdash P \cong \underline{A}$$

and

$$K[P \cong \underline{A}] \vdash P \cong P.$$

<u>Proof</u>. Since, for all V in $\Sigma(K)$, $V(P)=V(\neg P)$ if and only if $V(P)=V(\underline{A})$, both assertions follow from Proposition 8.5-1 by the definition of $\Sigma(K')$. In the first case, K' is $K[P \cong \neg P]$; in the second, K' is $K[P \cong \underline{A}]$. □

The next two propositions establish some necessary and sufficient conditions for a simple extension to be coherent.

<u>Proposition 8.5-3</u>. For all simple extensions K' of K, K' is a coherent system if and only if the set of valuation mappings for $\Sigma(K')$ is nonempty.

<u>Proof</u>. Let K' be any simple extension of K'; and let Φ denote the set of valuation mappings for $\Sigma(K')$. Suppose that K' is coherent. Then, by the definition of coherence, there exists an equivalence $P \cong Q$ in L(K') such that $K' \nvdash P \cong Q$. Hence, by Proposition 8.5-1, the equivalence $P \cong Q$ is not valid in $\Sigma(K')$. But then, by the definition of "valid," there must exist V in $\Sigma(K')$ for which $V(P) \neq V(Q)$. Thus, necessarily, Φ is nonempty.

Suppose next that Φ is nonempty. Then some valuation V in Σ(K) is a valuation in Σ(K'). Observe that, for all V in Σ(K), V($\underline{F}$)≠V($\underline{A}$). It follows that there exist V in Σ(K') such that V($\underline{F}$)≠V($\underline{A}$). But then, $\underline{F}$≅$\underline{A}$ is not valid in Σ(K'), so Proposition 2.1.6-1 shows that K'⊬$\underline{F}$≅$\underline{A}$. Accordingly, K' is coherent. □

Let the notation V* denote the valuation V of Σ(K) defined by: V*(p_i)=½, for all i=1,2,... .

Proposition 8.5-4. For all formulas P of L(K), the simple extension K(P≅P) is a coherent extension if and only if V*(P)=½.

Proof. It can be shown, by induction on the length of formulas P of L(K), that V*(P)=½ iff there exists at least one valuation V of Σ(K) such that V(P)=V(⌐P). We can make use of this result in the following way. Let K' denote a simple extension K[P≅⌐P] of K.

Suppose that K' is coherent. Then, by the Proposition 7.5-3, the set Φ of valuations of Σ(K') is nonempty. Say that V is in Φ. By the definition of the semantics Σ(K'), we have that the equivalence P≅⌐P is valid in Σ(K'). It follows that V(P)=V(⌐P). Then the above-stated result implies that V*(P)=½.

Suppose that V*(P)=½. Then, by the same result there exist V in Σ(K) such that V(P)=V(⌐P). By the definition of Σ(K'), this means that the set of valuations for Σ(K') is nonempty. Hence, K' is coherent, by Proposition 8.5-3. □

An immediate consequence of Proposition 8.5-4 is that the definition of "inconsistency" used by Varela is, more exactly, a form of incoherence--i.e., the three extensions CSR[α=ε]. CSR[α=[ε]], and CSR[ε=[ε]] are all incoherent extensions of CSR. This also illustrates by example that coherence, and not consistency, is the relevant critereon for determining the meaningfulness of non-Aristotelean systems.

CHAPTER 9

UNDECIDABILITY AND UNKNOWABILITY

The aim of this last chapter of Part II is two-fold: (i) to establish a formal connection between Varela's concept of logical autonomy and the notion of a totally undecidable partial recursive set, (ii) to define a general concept of "mechanical unknowability" and to show how various instances of this idea may be represented in a non-Aristotelean, yet coherent, frame. The former is accomplished in §9.1, the latter in §9.2.

§9.1.1 lays down basic terminology from recursion theory. Then §9.1.2 defines a system PRS of partial recursive sets, which in principle is a simple recasting of the system K. Then the desired connection between autonomy in CSR and undecidability in PRS is then established in §9.1.3.

§9.2.1 begins the study of mechanical unknowability with a reorientation on the concepts from recursion theory, taking Turing machines as the basic givens. §9.2.2 extends this into a definition of mechanical unknowability vs. mechanical conceivability, based on the popular brain-computer analogy. §9.2.3 looks more carefully at the general form of paradox. §9.2.4 then proceeds to some illustrations regarding (i) "perfect" systems, (ii) time, (iii) the concept of "goodness," and (iv) transcendent realms of being.

§9.1. Formal Autonomy Vis-A-Vis Recursive Undecidability

§9.1.1. Partial Recursive Sets

The following is an admixture from Kleene (1952) and Rogers (1967). Let N denote the natural numbers. A _partial_ _function_ ψ is a mapping of some subset Dom(ψ) N into N. Dom(ψ) is the _domain_ _of_ _definition_ of ψ, and, for each _argument_ xεN, the _value_ ψ(x) is _defined_ or _undefined_ according as xεDom(ψ) or x$\notin$Dom(ψ). A partial function is a _total_ _function_ if Dom(ψ)=N.

The concept of _recursion_ may be given in terms of a mathematically well-defined class $\mathfrak{M}$ of _Turing_ _machines_. For an explicit definition, see Rogers (1967). Loosely, a Turing machine is an algorithm consisting of a set of instructions describing precise mechanical steps by which, given any number in N, one can proceed with a well-defined computation. Let M be a machine in $\mathfrak{M}$ and let M(x) represent the application of M to the argument x. If the computation procedure terminates, producing a value y, one writes M(x)=y and says that the computation _converges_; otherwise, the computation _diverges_ (loops, provides no information, etc.). A partial function ψ is _algorithmic_ if there is a Turing machine M_{ψ} such that:

1. If ψ(x) is defined, then M_{ψ}(x)=ψ(x),

2. If ψ(x) is undefined, then M_{ψ}(x) diverges.

Each M describes a unique ψ; on the other hand, it is possible that a given ψ have more than one choice for M. A _recursive_ _function_ is an algorithmic total function. A _partial_ _recursive_ _function_ is an algo-rithmic partial function. Hence, by definition, recursive functions are partial recursive.

A <u>characteristic</u> <u>function</u> is a partial function κ such that either $\kappa(x)=1$ or $\kappa(x)=0$, for all $x \varepsilon \text{Dom}(\kappa)$. For each characteristic function κ, there is a <u>partial</u> set p_κ defined by

$$x \varepsilon p_\kappa \text{ if } \kappa(x)=1$$

$$x \not\varepsilon p_\kappa \text{ if } \kappa(x)=0$$

undecided if $\kappa(x)$ is undefined.

If κ is total, then p_κ is a total set (hence a "set" in the classical sense). For a given partial set p, let κ_p denote the characteristic function for p. A partial set p_κ is a <u>recursive</u> <u>set</u> if κ is a recursive function, and is a <u>partial</u> <u>recursive</u> <u>set</u> if κ is a partial recursive function.

The interrelation of partial sets, sets, partial recursive sets, and recursive sets is illustrated in Figure 9-1. We may see that the various regions of this figure are nonempty as follows. A simple example of a recursive set is the <u>empty</u> <u>set</u> ϕ defined by $\kappa(x)=0$, for all $x \varepsilon N$. The set ϕ is recursive because there is a Turing machine M such that $M(x)=0$ for all $x \varepsilon N$. A partial recursive set which is not a recursive set is the <u>empty</u> <u>partial</u> <u>set</u> v defined by having $\kappa_v(x)$ be undefined for all $x \varepsilon N$. v is recursive because κ_v can be represented by a machine which always "loops" or is otherwise nonterminating. That there are sets which are not recursive sets and that there are partial sets which are not partial recursive sets is an immediate consequence of there being only countably infinitely many Turing machines, while the sets and partial sets are uncountably many. A specific example of a (total) set which is not recursive can be given as an instance of the well-known "halting problem." Let $\psi_0, \psi_1, \psi_2, \ldots$ be a standard (Gödel) numbering of the unary partial recursive functions. Define q by
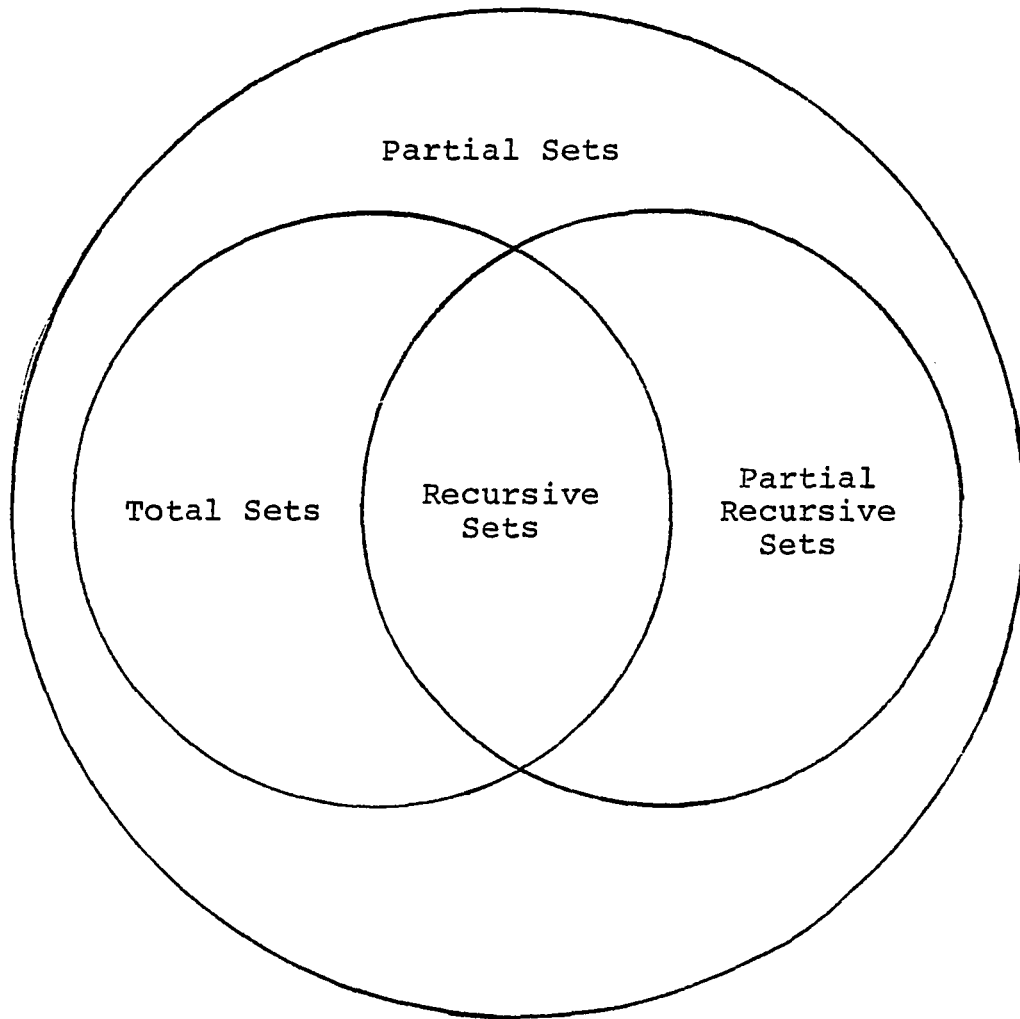
Figure 9-1. Interrelation of different "sets."

$$x\varepsilon q \text{ iff } x\varepsilon Dom(\psi_x).$$

Then, for all xεN, either xεq or x$\notin$q; so κ_q must be total. However, Roger's Corollary 1-VII shows that κ_q is not ψ_i for any i=1,2,.... Hence, κ_q is not recursive. A specific example of a partial set which is not a partial recursive set can be obtained from q by the ad hoc specification that $\kappa_q(0)$ is undefined.

§9.1.2. The System PRS

Kleene's original definition of L(K) treated propositional variables as n-ary relation symbols, and each was thought of as representing an "n-ary partial recursive relation" on N. For convenience, we here deal only with unary relations, i.e., partial recursive sets.

The system PRS is obtained from K by modifying L(K) as follows: let $\underset{\sim}{x}$ denote a new symbol, called an __individual variable__; replace p_i with $p_i(\underset{\sim}{x})$, for all i=1,2,....; replace $\underline{A}$ and $\underline{F}$ with $\underline{A}(\underset{\sim}{x})$ and $\underline{F}(\underset{\sim}{x})$. Then propositions P become denoted by $P(\underset{\sim}{x})$.

To motivate the definition of Σ(PRS), Kleene's results may be summarized as follows. A partial set p is alternatively denoted by p(x), taking x as ranging over N, in which case the notation p(n) represents the assertion that nεp. Then some __truth-values__ may be defined by:

1. p(n) is $\underline{t}$ if $\kappa_p(n)$=1,

2. p(n) is $\underline{f}$ if $\kappa_p(n)$=0,

3. p(n) is $\underline{u}$ if $\kappa_p(n)$ is undefined.

This leads to truth tables for $\neg$, $\vee$, &, $\supset$, $\equiv$, and $\cong$, with the following properties: (i) if p(x) and q(x) are partial recursive sets, then $\neg$p(x) and p(x)$\vee$q(x) defined by

$$\kappa_{\,p}(n) = 1-\kappa_p(n)$$

$$\kappa_{p\,q}(n) = \max[\kappa_p(n),\kappa_q(n)]$$

are partial recursive sets, (ii) if &, ⊃, and ≡ are defined in terms of ⌐ and V as in §7.1, then they similarly yield partial recursive sets, (iii) the strong equivalence p(x)≅q(x) asserts that p(x) and q(x) are identical, i.e., that $\kappa_p(n)=\kappa_q(n)$ for all n∊N, (iv) the connective ≅ is not explicitly definable in terms of ⌐ and V, and (v) ⌐, V, and & correspond to the customary partial-recursive set complementation, union, and intersection.

Now, the semantics Σ(PRS) has as <u>values</u> the numbers 1, 0, and ½-- corresponding to <u>t</u>, <u>f</u>, and <u>u</u>--and as <u>valuations</u> has all ordered pairs (Φ,â), where â is an <u>assignment</u> such that

1. $p_i^{\hat{a}}$ is a partial recursive set,

2. $\underline{F}^{\hat{a}}$ is the empty set φ,

3. $\underline{A}^{\hat{a}}$ is the empty partial set v,

4. $x^{\hat{a}}$ is a number in N,

and where Φ is defined by

5. If P is $\underline{F}$, $\underline{A}$, or p_i for some i=1,2,..., then

$$\Phi(P(\underset{\sim}{x}^{\hat{a}})) = \begin{cases} 1 \text{ if } \kappa_p\hat{a}\ (\underset{\sim}{x}^{\hat{a}})=1 \\ 0 \text{ if } \kappa_p\hat{a}\ (\underset{\sim}{x}^{\hat{a}})=0 \\ \tfrac{1}{2} \text{ if } \kappa_p\hat{a}\ (\underset{\sim}{x}^{\hat{a}}) \text{ is undefined} \end{cases}$$

6. If P is of the form ⌐Q, then

$$\Phi(P(\underset{\sim}{x}^{\hat{a}})) = 1-\Phi(Q(\underset{\sim}{x}^{\hat{a}})),$$

7. If P is of the form QVR, then

$$\Phi(P(\underset{\sim}{x}^{\hat{a}})) = \max[\Phi(Q(\underset{\sim}{x}^{\hat{a}})),\Phi(R(\underset{\sim}{x}^{\hat{a}}))]$$

An equivalence P(x)≅Q(x) is <u>valid</u> in Σ(PRS) if $\Phi(P(\underset{\sim}{x}^{\hat{a}}))=\Phi(Q(\underset{\sim}{x}^{\hat{a}}))$ for every assignment â. It follows that P(x)≅Q(x) is valid in Σ(PRS) if and

only if P≅Q is valid in Σ(K). Then Theorem 8.3-1 implies that PRS is complete with respect to Σ(PRS).

Evidently, PRS is essentially identical with K, and Σ(PRS) is only a slightly more sophisticated version of Σ(K).

§9.1.3. Undecidability

A partial set is <u>undecidable</u> if it is not a recursive set. Thus, a partial set can be undecidable in either of two ways: (i) it is non-total, or (ii) it is total but not recursive. An example of (ii) is the set q of §9.1.1. In terms of Turing machines, this means that there is no $M\epsilon\,\mathcal{M}$ which effectively distinguishes between the members and non-members of q. The modified q, on the other hand, is an example of (i); membership for 0 in q is "undecided" a priori.

This terminology motivates referring to the empty partial set $\vee$ as the <u>totally</u> <u>undecidable</u> <u>partial</u> <u>recursive</u> <u>set</u>, since (i) $\vee$ is in fact partial recursive, and (ii) $\vee$ strictly speaking is not "empty," rather its membership relation is merely "undecided" for all $x\epsilon N$. The afore-mentioned correspondence between $\vee$ and Varela's concept of formal autonomy can be exhibited as follows.

The isomorphism τ of CSR and K can be taken as an isomorphism of CSR and PRS. Then $\tau(\alpha)=\underline{A}(\underline{x})$. Recall that $v(\alpha)=\underline{a}$, for all valuations v in Σ(CSR). From §9.1.2, $\underline{A}^{\hat{a}}=\vee$ for all assignments $\hat{a}$ of Σ(PRS). Since CSR is complete with respect to Σ(CSR), and PRS is complete with respect to Σ(PRS), it follows that any assertion about the autonomous state $\underline{a}$ that is both formalizable and derivable in CSR is simultaneously an assertion about the totally undecidable partial recursive set $\vee$ that is

formalizable and derivable in PRS; and conversely. Thus, $\underline{a}$ (autonomy) and $\vee$ (total undecidability) are formally identical.

§9.2. Mechanical Unknowability

A saying among logicians is that in recursion theory one perceives the world "through a glass darkly." The sense of this is that the concept of recursive undecidability can be regarded as a concept of mechanical unknowability; to wit, from the standpoint of the class of partial recursive functions, our knowledge about the subsets of N is limited to the information that is provided us by the collection $\mathcal{M}$ of Turing machines. This section makes this idea explicit in a model of the observer-observed relationship based on the popular brain-computer analogy, together with Church's Thesis about the relation of real-world machines to recursive functions.

To this end, the concept of recursion is here recast in terms of just $\mathcal{M}$, rather than with the concept of a "partial set," as follows.

§9.2.1. Recursion Revisited

An alternative formulation of recursion is as follows. Let N denote the natural numbers; let $\mathcal{M}$ be the class of Turing machines operating on N. If $M\varepsilon\mathcal{M}$ converges for all $x\varepsilon N$, then M is a total machine; otherwise, M is a partial machine. A characteristic machine M will be such that, for each $x\varepsilon N$, either $M(x)=1$, $M(x)=0$, or $M(x)$ diverges. For each characteristic machine M, let a partial recursive set σ_M be defined by: (i) $x\varepsilon\sigma_M$ if $M(x)=1$, (ii) $x\notin\sigma_M$ if $M(x)=0$, and (iii) undecided if $M(x)$ diverges. σ_M is said to be a recursive set if M is total. These collections are identical with the corresponding collections defined in

§9.1.1, except that here "recursive" and "partial recursive" are mutually exclusive.

Church's Thesis asserts that the intuitive concept of "effectively computable" is exactly captured by the class of partial recursive functions. More exactly, if a (partial or total) function defined on N satisfies one's reasonable understanding of what it means for a function to be algorithmic, then its values will be computable by a Turing machine. (The converse, that recursive implies effectively computable, is obvious from Turing's definition of the class $\mathcal{M}$.) The support for Church's Thesis is largely empirical: (i) every intuitively acceptable mathematical formulation of "effective computability" has been shown to be equivalent to the Turing characterization, and (ii) no function on N has been discovered which satisfies the intuition as being "effectively computable" and is not representable in $\mathcal{M}$.

Let σ be a subset of N. A decision method for σ is a method by which, given any $x \varepsilon N$, one can determine in a finite number of steps whether $x \varepsilon \sigma$ or $x \notin \sigma$. Evidently, this would be a method of effective computability for the characteristic function κ_σ. The decision problem for σ is: find such a decision method or prove that no such method exists. By Church's Thesis, a decision method exists for σ if and only if σ is a recursive set, so the decision problem amounts to finding a Turing machine for κ_σ or showing that there isn't one. One says that the decision problem for σ is recursively solvable if σ is recursive, and recursively unsolvable if not. It follows that the decision problem for σ is recursively solvable if and only if σ is decidable in the sense of §9.1.3.

The import of recursive unsolvability is that, given any "candidate" decision method for an undecidable σ, there will be numbers $x\epsilon N$ for which that method produces no result. Yet this does not exclude the possibility that a decision problem be _approximately_ _recursively_ _solvable_ or that different decision methods might be better than others in providing _recursive_ _approximations_ of a complete determination for σ. By Church's Thesis, such partial approximations will be represented by partial characteristic machines. For each such M in $\mathfrak{M}$, let $\hat{\sigma}_M$ be the subset of N defined by $x\epsilon\hat{\sigma}_M$ if and only if $M(x)=1$. Then a _maximal_ _approximation_ of σ will be any maximal $\hat{\sigma}_M{}^{\subset}\sigma$. It follows that σ is decidable if and only if $\hat{\sigma}_M=\sigma$. In any case, maximal approximations are recursive sets.

A _totally_ _undecidable_ _set_ in this context will be any set v such that every maximal approximation $\hat{v}_M=\phi$. Such a set might be one whose members are prescribed completely at random, so that for each $x\epsilon N$, there is only a 50-50 chance that $x\epsilon v$. In contrast with the v defined in §8.1.1, as a unique partial set, there here may be any number of such totally undecidable subsets of N. From the standpoint of recursive solvability via the class $\mathfrak{M}$, however, all such sets are equivalent. It follows that the semantics $\Sigma(PRS)$ can alternatively be defined in terms of equivalence classes of partial characteristic machines, rather than partial recursive sets.

§9.2.2. Universes: Conceivable Vs. Knowable

The popular brain-computer analogy asserts that the human brain may be regarded merely as an extremely sophisticated computing machine. Given the present state of the art for computers, however, it is perhaps

more reasonable to refer to an "intellect-computer" analogy, since machines have not yet been designed which replicate the emotive, intuitive, and similar aspects of the human personality (even though it might be speculated that they can be). Given an intellect-computer analogy, the foregoing analysis of recursive solvability leads to a provocative--albeit simple--model of the limits of "intellectual knowability."

One easily conceptualizes a countably infinite set of natural numbers N and its set of subsets $\underline{P}(N)$, and it is reasonable to assume that this manner of conceptualization is mechanical. Indeed, a computer program could be developed which describes all the axioms and inference rules of classical set theory, and proceeds to derive proofs of theorems. Thus, if the human intellect is a machine, the ordered pair $(N,\underline{P}(N))$ constitutes a <u>mechanically conceivable universe</u> U_c.

Yet, as was seen in §9.1.1, not all of the subsets of N--in fact only countably infinitely many of the uncountable collection $\underline{P}(N)$--are recursively decidable. Let the decidable sets in $\underline{P}(N)$ be denoted by $\underline{D}(N)$. Then it makes sense to take the ordered pair $(N,\underline{D}(N))$ as a corresponding <u>mechanically knowable universe</u> U_k. $\underline{D}(N)$ is that segment of $\underline{P}(N)$ which is completely discernible by the Turing machines $\mathcal{M}$, where $\mathcal{M}$ by Church's Thesis, includes <u>all</u> machines that can possibly exist.

The interconnection between U_c and U_k can be elaborated. Let σ be any subset of N; let $\hat{\sigma}_M$ be a maximal recursive approximation of σ; and let δ_M be the set of $x \varepsilon N$ such that M(X) diverges. Then δ_M consists of all potential elements of σ which are left undecided by M. Consider U the collection of all subsets of N having the form $\delta \cup \hat{\sigma}$ where $\delta \subset \delta_M$. Clearly $\sigma \varepsilon U$; and σ is the unique member of U if and only if σ is decidable. Suppose that σ is recursively undecidable and let c be the

cardinality of δ_M. It may be noted that the cardinality of U is 2^c, where 2^c is the uncountable infinity $\aleph_1$ in case c is infinite. Most importantly, from the standpoint of U_c, the members of U are distinct members of $\underline{P}(N)$, while from the standpoint of U_k, the members of U are indistinguishable from the set σ. These considerations suggest the following interpretation for U_c and U_k.

A being which has a strictly mechanical intellect is able to conceptualize an uncountably infinite universe U_c of which only a countable part U_k is really (intellectually) knowable. The intellectually unknowable part, however, is approximately knowable--more or less so depending on how closely he obtains a knowable approximation. The same being is furthermore able to determine a closeness of fit for his approximation by looking at the number of elements of N which his approximation leaves undecided. And he can sometimes, through a concept of recursive unsolvability, be cognizant of whether a complete approximation is possible.

Thus, this being resides in a world where he cannot "know" everything that he can "conceptualize"; he can approximate a complete knowledge of some unknowns, but he is aware that absolute knowledge of everything is impossible. Indeed, he can be aware of "totally unknowable" concepts as represented by intellectual loops or otherwise paradoxical circular reasoning.

In this manner, paradox represents the absolute limits of intellectual knowability. By means of circularity, intellect formulates concepts of things which are completely beyond its ability to comprehend.

The interpretation of mechanical unknowability as an idea of intellectual unknowability can be extended to an idea of empirical

untestability. For, all manners of scientific observation, measurement, and analysis here become strictly mechanical as products of the human intellect. To elaborate, suppose that the physical universe consists of countably infinitely many objects. Then these objects can be enumerated by the members of N, and $\underline{P}(N)$ represents the collection of all conceivable properties of objects--i.e., x has the property represented by σ if and only if $x\varepsilon\sigma$. Thus, one has a physical interpretation of U_c, while U_k correspondingly becomes interpreted as giving the properties for which one can determine by strictly scientific analysis whether an object does or does not have that property. Thus $\underline{D}(N)$ represents those properties which alone are empirically testable. Partially knowable concepts accordingly become partially testable properties, and totally unknowable concepts become totally untestable properties. Thus, a being having a strictly mechanical intellect and living in a countably infinite universe can conceptualize properties of objects which he should not hope to establish by strictly empirical methods. While this in itself is not surprising--indeed it is rather common knowledge--it is remarkable that recursion theory provides such a simple model which suggests why this is true.

§9.2.3. Paradox

As a preface to specific examples of mechanical knowability, it will be useful to look more carefully at the general form of paradoxical assertions. In the classical notations, a proposition P is paradoxical if it enters into an assertion of the form

$$P \equiv \neg P, \tag{1}$$

which reads "P is true if and only if not-P is true." This is to be distinguished from an ordinary contradiction having the form

$$P \ \& \ \neg P, \tag{2}$$

which reads "both P and not-P are true." The latter asserts the simultaneous truth of P and $\neg P$, while the former asserts that the truth of each is predicated on the truth of the other.

However, both (1) and (2) are antitautologies of $\Sigma(PC)$, and hence are formally equivalent with one another--i.e., each implies the other. Thus paradoxes cannot be adjoined to a classical Aristotelean system such as PC without rendering the system trivially incoherent. Similar remarks apply to the adjunction of a paradoxical assertion of the form

$$P = \overline{P\rceil}$$

to the primary algebra. This shows that in Aristotelean systems, paradoxes are always formally meaningless.

Paradoxes can become meaningful, however, in non-Aristotelean systems. In the calculus for self-reference, paradoxical assertions have the form

$$P = \overline{P\rceil} \tag{3}$$

and, when translated into the system K, have the form

$$P \cong \neg P. \tag{4}$$

Each of (3) and (4) can be validated in their respective semantics--the former by any v such that $v(P)=\underline{a}$, and the latter by any V such that $V(P)=\frac{1}{2}$. Thus (§8.5), they can be adjoined to their respective systems without ruining coherence.

This shows generally the value of Kleene-Varela type systems for the discussion of paradoxical ideas. Moreover, it suggests a generalized definition of paradox per se as being any assertion to the effect that a

proposition is equivalent with its negation. The familiar semantic paradoxes, such as the paradox of the liar (Martin 1970) and Russell's set of all sets which are not members of themselves (Russell 1902 and 1908) are special cases of paradox which involve an internal linguistic self-reference. The general paradoxical form does not require such self-reference, but when it is regarded in the context of the calculus for self-reference, it is revealed as involving a self-reference on another level; to wit, the logical status (i.e., autonomy) of a paradoxical proposition is formally defined in term of itself through "re-entry" into the "form of indication" ⌐ (Varela, 1975). The logical autonomy of paradoxical propositions P is furthermore made explicit in the systems K, PRS, and the calculus for self-reference by Proposition 8.5-2.

§9.2.4. Sample Interpretations

The foregoing shows how the model of a universe, the knowledge of which is limited by the capabilities of a collection of machines, can be formalized in PRS. Of course, for more sophisticated interpretations of "mechanical unknowability," a much richer linguistic structure than L(PRS) is required. To this end, PRS might be expanded, in the same manner that PC is normally expanded to obtain first-order and second-order systems. Indeed a work by Zadeh shows that the prototype self-referential paradox, the paradox of the liar, is representable in the logic of possibilistic reasoning (Zadeh 1979), which suggests that the required richer languages may be just those, or developments of those, for the linguistic theories defined in Chapter 3.

The interpretations described below are intended merely as illustrations of how such expanded systems might be employed. Moreover, they focus only on "total unknowability" and therefore do not explore the full range of potential application to "approximately knowable" ideas.

§9.2.4.1. Perfect Systems

Weinberg's "perfect systems law" asserts: "true systems properties cannot be investigated" (Weinberg 1975, p. 160). The standard method of investigating a real-world system is to observe it while it undergoes transformations. But "true systems properties"--such as "perfect wholeness" or "absolute autonomy"--are invariants intrinsic to the system and are therefore lost if the system is changed. Consequently, a "perfect system"--i.e., a system which has only true system properties--"could not be studied if found."

The concept of a perfect system may be formalized in a Kleene-Varela system by developing a proposition $P(\underset{\sim}{x})$ which asserts that $\underset{\sim}{x}$ is perfect--e.g., P might be built up of propositions P_w, P_a, etc., which assert $\underset{\sim}{x}$'s wholeness, autonomy, etc.--and by adjoining the equivalence $P(\underset{\sim}{x}) \cong A(\underset{\sim}{x})$ as an axiom. Then P becomes interpreted as a totally unknowable property of real-world objects.

In this context, the semantics may receive an alternative interpretation, based on a logic of empirical methods, which is defined by having a proposition be

1. "true" if it has been verified by experiment,

2. "false" if it has been refuted by experiment,

3. "unknown" if it has been neither verified nor refuted.

Then $P(\underset{\sim}{x}) \cong A(\underset{\sim}{x})$ asserts that for all $\underset{\sim}{x}$ the truth-value of $P(\underset{\sim}{x})$ is always unknown.

§9.2.4.2. McTaggart's Paradox

Gale (1968) "uses McTaggart's argument for the unreality of time as a starting-point for an examination of the logic of and interconnectedness between tensed and tenseless discourse" and defends "the common belief in the objectivity of temporal becoming" (from the book jacket). McTaggart's paradox involves two conceptions of time: an A-series based on the notions of "past," "present," and "future," and a B-series based on the relation of "before and after." A "positive thesis" shows that the A-series is reducible to the B-series, while a "negative thesis" shows that the B-series leads to an irreconciable infinite regress of paradoxes.

Gale's treatment begins with the observation that all attempts to define time inadvertently lead to paradox, and it proceeds by the method of linguistic analysis to show that people naturally accept a belief in the objectivity of temporal becoming even though they cannot define it or rationally prove it to be real. "What time makes it possible for us to say is exactly what cannot be said about time" (p. 234).

This implies that the natural use of temporal language involves paradoxical, or otherwise circular, reasoning. Thus a Kleene-Varela system which includes McTaggart's paradox as an axiom might serve for formalizing temporal discourse as it actually appears in observable uses of human language. Time is a clear example of something that can be conceptualized but cannot be understood by ordinary "rational" means.

§9.2.4.3. Moore's Paradox

Another is the idea of "goodness." Hartman (1967) develops a "logic of value" which uses Moore's paradox as a prototype test case. This paradox is quoted from Moore (p. 18) as consisting of "the two different propositions that are both true of goodness, namely: (i) that it does depend only on the intrinsic nature of what possesses it, and (ii) that, though this is so, it is as yet not itself an intrinsic property." Hartman "solves" this paradox by unraveling a confusion of levels, showing that "the negative proposition--concerning what goodness is not--refers to the thing itself; and the positive proposition--concerning what goodness is--refers to the concept of the thing."

Hartman's work then disbands the self-referential negative proposition and elaborates the positive one in terms of properties of concepts rather than of things. While this is certainly a reasonable approach to the problem of values, it leaves unaddressed a fact that remains embodied in Moore's paradox: that people nevertheless do speak and think as if goodness is an inherent property of things. This raises the question of why it is so natural for people to think of value in this way. Evidently, one has here another example of circularity in natural human reasoning, and the property of goodness becomes another candidate for formulation in a logically autonomous proposition.

§9.2.4.4. Transcendent Realities

God, Truth, Justice, Beauty, Perfection, Freedom of the Will, and so on all represent realities which are completely beyond the human ken. Indeed, it is almost an embarrassment to philosophers and theologians that the existence of such entities has no logical or empirical verification.

In point of fact, the only position of complete intellectual integrity on such matters is Skepticism: the existence of transcendent realities is uncertain. In theology, this is Agnosticism.

Yet it is equally true that intuition, beliefs, experiences of faith, and the higher ideals, frequently take precedence over the rational intellect as prescriptors of human behavior. While much effort has been directed toward justifying this phenomenon, little has been done to explain how it occurs.

The Kleene-Varela logic can be offered as a model of the kind of intellect which makes such occurrences possible. In this model, the intellect remains strictly mechanical, but goes beyond the exclusive "true of false." Clearly, if the intellect was exclusively two-valued, there would be no question of transcendental realities. It is only by virtue of man's ability to entertain logical uncertainty that he has the opportunity to extend beyond the purely "rational" and act intuitively, to formulate beliefs, and to experience faith--i.e., to be "human" in the higher sense of the word.

PART III

THE USE OF NON-CLASSICAL LOGICS

CHAPTER 10

THEORETICAL FOUNDATIONS

The primary aim of this chapter is to set the stage for the methodo-
logical issues taken up in Chapter 11. We lay down a few items of
terminology, discuss the sense in which the proposed use of nonclassical
logics obtains a theoretical rationale, and generally motivate the
ensuing considerations.

§10.1 constitutes a brief survey of the development of logical
ideas insofar as this illustrates a particular view regarding the
processes by which mathematical studies evolve.

§10.2 takes up the matter of "selecting" one's mode of reasoning,
and considers the manner in which this relates to an inherent "bimodality"
of mathematical thinking.

§10.3 discusses the use of nonclassical logics for purposes of
modeling qualitative aspects of real-world systems.

§10.4 is a general introduction to the subject of formal systems
and their semantics, which is intended both as a supplement to §2.1 and
as a prelude to Chapter 11.

§10.1. An Evolutionary Overview

The prospect of using nonstandard modes of reasoning in the design
of system models stems largely from there being a wide variety of known
nonclassical logics. In order to obtain a better idea of how such
logics may be employed, it will be useful to first consider the manner

in which they have evolved. We therefore begin here with these more general considerations, before going into the particulars concerned with "selecting" alternative modes of reasoning. In process, this brings to fore some ideas which will play a role in later sections.

The development of logical ideas may be viewed as involving two distinct but interrelated processes of human intellection. First is the one by which mathematics most directly serves the purposes of science, and which shall herein be referred to as observation formulation. Specifically, this designates a process which springs from studies of concrete physical, biological, social, etc. systems under the motivation to provide these systems with precise analytical descriptions. As such, observation formulation sometimes leads to implementation of existing bodies of mathematical knowledge, whereas it at other times leads to the invention of completely new kinds of mathematics. To illustrate: the differential calculus was created for purposes of formulating the concept of a "rate of change" as it appears in the laws of mechanics; and statistics were developed to capture the notion of "tendencies" or "trends." Observation formulation thus construed clearly underlies the development of all mathematical models of real-world situations.

The second process to be considered is a somewhat more subtle one, pertaining to the intrinsic nature of mathematics proper. It may be characterized as a process by which the mathematician makes explicit certain aspects of his own role in relation to his work--i.e., his manners of thinking about specific mathematical ideas. Two examples here are the invention of the "variable quantity" and the definition of a "set of elements." Ordinarily, the process by which such concepts as these come into being is referred to as one of "abstraction." A truer

assessment, however, is that these concepts are _tools_ by which certain acts of abstraction may be carried out. The concepts themselves arise, rather, only when the mathematician strives to comprehend and then make explicit his own manners of thinking, for purposes of organizing and clarifying ideas or for simplifying the mental exercises that his work requires him to perform. Accordingly, this second process shall be referred to as _introspection articulation_, which emphasizes furthermore that it customarily leads to new items of mathematical language.

This classification of mental activities is of course not inclusive of all aspects of mathematical endeavor. It is concerned rather only with the creation of new ideas and therefore omits other activities, such as theorem proving, by which mathematical theories grow to fruition. Nor is the classification crisp, since most mathematical investigations invoke both processes. On the one hand, the desire to formulate a model of a given real-world situation may lead to introspection in search of new modeling techniques. On the other hand, the results of mathematical introspection may inspire the development of modeling approaches not previously conceived. In physics, for example, an illustration of the former would be functional analysis, which generalizes the theory of differential equations; and an instance of the latter would be quantum logic (Reichenbach 1944), which applies formal systems to the treatment of anomalies in the physical theory.

In any case, the present distinction is not to be confused with the conventional classification of mathematics as being "pure" or "applied." Among contemporary mathematicians, the latter distinction is almost universally regarded as an unfortunate choice of words, since the underlying intention of all mathematics is that it is either directly or

indirectly meant to be applied. This pertains even to philosophical studies, which are undertaken to provide the mathematics of science with a solid foundation. Yet even taking this into account, it may be argued that various forms of mathematics involve more of one of the above processes than the other. Mathematical logic, in particular, is fundamentally a product of introspection articulation and, only after reaching a sufficient maturity, became available for direct applications.

As motivation for the style of applications proposed in this dissertation, and to illustrate more fully the notion of introspection articulation, it will be useful to recount some of the milestones in the development of logical ideas. Except where indicated otherwise, the following references are from Kleene (1952, pp. 60-65).

The discovery of the axiomatic deductive method is attributed to Pythagoras (c. 700 B.C.) who used it to derive many of the fundamental theorems of elementary geometry. After several centuries of continued development, the Pythagorean geometry received its optimum crystalization in Euclid's "Elements" (c. 300 B.C.), which work for many centuries thereafter remained unchallenged as representing the ideal of mathematical elegance and clarity of thought.

Institution of the formal treatment of logic per se, as a method of reasoning with sentences in terms of their grammatical forms, is due to Aristotle (also c. 300 B.C.). Out of Aristotle's work come many of the basic principles of classical logic, notably: Excluded Middle (for any proposition A, either A is true or not-A is true); Non-Contradiction (for no proposition A do we have that both A is true and not-A is true); and Modus Ponens (if A is true, and it is true that A implies B, then B is true). Also due to Aristotle is the idea of "syllogism," which constituted

a major step toward the theory of logical "quantification" as used in modern first- and second-order logics.

The invention of the variable quantity, and in general, the development of the algebraic symbolism, is attributed to Vieta (1591) and others working around the same time. Kleene indicates this period as a significant turning point in the history of mathematical ideas: "The discovery of simple symbolic notations which lend themselves to manipulation by formal rules has been one of the ways by which modern mathematics has advanced its power" (p. 61). Prior to Vieta, mathematicians suffered the tediousness of describing their deductions in words.

Significant early developments in the contemporary logic are Leibniz (1866), de Morgan (1847 and 1864), Boole (1847 and 1854), Pierce (1867 and 1880), and Schröder (1877 and 1890-1905). These works collectively led to formalizations of various portions of mathematics. The theory of arithmetic was first axiomatized by Dedekind (1888), and then developed deductively from Dedekind's axioms by Peano (1889). Frege (1893 and 1903) wrote down the first definition of formal "proof," by means of diagrams showing interconnections between assertions, and used this to study the "logical foundations" of arithmetic. Whitehead and Russell (1910-1913) reduced much of the prior work to a concise symbolism, thereby providing what is today the standard prototype of a formal logical system. The sense in which logical systems are "formal," however, may be attributed to Hilbert, who was first to emphasize the total abstraction of symbols from meanings, so that proofs became simply mechanical derivations carried out in accordance with precise rules for operating on symbols and symbolic expressions.

Not included in Kleene's list are important advances by Tarski and Gödel. To Tarski we are indebted for articulating the concept of mathematical "truth," which now forms the basis for the study of "semantic interpretations" of formal systems (model theory). Tarski's definition (cf. Tarski 1949), "The assertion 'Snow is white' is _true_ if and only if snow is white," is an excellent example of how a simple statement of the obvious can have profound consequences in the development of mathematical thought.

Gödel's contribution (1931) made use of all the foregoing ideas, together with a development of Skolem's idea of "recursive function" (Skolem 1923). This work showed that the system of Whitehead and Russell's _Principia_ _Mathematica_ (1910-1913), which derives Peano arithmetic within a formal logical system, can be interpreted within its own language by means of a straightforward enumeration of that system's symbolic components. Attendant with this came the now famous result (the incompleteness theorem) that that system contains "formally undecideable" propositions--i.e., propositions which can neither be affirmed nor denied on the basis of the system's axioms. This result in turn led to a general study of undecideability and the discovery of many other formally "unsolvable" problems of contemporary mathematics. A collection of these is documented by Odifreddi (1981).

In viewing this development as a phenomenon of articulated introspection, it is easy to see how mathematics arrived at the genre of systems that are studied in this dissertation. As soon as the concept of a formal logical system was fully articulated, and became a subject on its own for mathematical investigation, a phenomenon analagous to the development of the non-Euclidean geometries began to occur--namely, the

invention of non-Aristotelean or otherwise nonclassical logics. The
first of these new logics were studied for purely philosophical reasons,
by Brower's school of Intuitionism (cf. Heyting 1956). However, once
those studies were begun, it soon became evident that the basic idea of
a formal logic could be extended for purposes of describing many further
aspects of human reasoning. Thus ensued the plethora of multi-valued
logics, modal logics, temporal logics, inductive logics, autogenetic
logics, and so on, that we have today. Very recently these logics have
begun to be applied, and the aim of the next following section is to
consider the foundational justifications for this development.

Before proceeding to that issue, however, it should be said that
merely regarding a particular logic--be it classical or nonclassical--as
articulated introspection is not tantamount to upholding that logic as
being in any sense superior to another as a model of natural human
reasoning. A case in point is Boole's "laws of thought," which has
hence been regarded as mistitled. What is intended by "introspection
articulation" is rather only to emphasize that different logics have
arisen in the effort to model different aspects of reasoning. Whether
any one logic gives a better fit to the natural reasoning is a separate
matter.

The prevailing opinion regarding formal logics is therefore upheld,
that these serve mainly as guides to thinking, and that no one logic
should be sacrosanct. This opinion shall here be overlayed, however,
with a less conventional view regarding, if not the sacredness, then at
least the inherent fundamentality of bimodal reasoning, insofar as
mathematical studies are concerned. This will be part and parcel with
the discussion below.

§10.2. On Selecting One's Logics

The physical sciences have clearly acknowledged that no one model of the universe is absolute. This was an irrevocable albeit unexpected consequence of the Einsteinian revolution in physics. Logical studies, developed mostly in isolation from physics, have nonetheless produced an analogous conclusion: neither is any one mode of reasoning absolute. As articulated introspection, the various known forms of logic make it evident that a wide variety of modes of reasoning are viable; and inasmuch as these logics exist as formalized mathematical systems, it is evident that a wide variety of modes of reasoning also are available for use. Hence, in exactly the same way that science selects alternative geometries for modeling the universe, it can also select the logics by which it derives conclusions and organizes ideas.

This in effect is a realization of the prophecy by Lukasiewicz (1930), that science eventually would make direct application of multi-valued logics in its methods of research (cf. Gaines 1976). The freedom to select one's mode of reasoning, however, comes only at the expense of some new difficulties.

Soon after science clarified the distinction between a system and a system model, it became evident that models are always problem specific and observer- (or user-) dependent; that is, they have meaningful-ness only in a particular context as provided by a certain problem setting, and only to a certain collection of observers. But, in order for a model to be meaningful among a group of observers, it is necessary that it "make sense" in the manner of functioning effectively as a means of communication among that group's members. This in turn reflects back on the reasoning processes by which the model was constructed. Things

generally "make sense" to people only if they are expressed within those peoples' own logics. Such has been emphasized especially in a work by Maruyama (1974), which shows how basic epistemologies or "structures of reasoning" may differ from culture to culture, profession to profession, and so on. Consequently, it is important that, whenever a nonstandard logic is selected as the basis for a model, that logic must itself be described precisely in the context of some given, mutually comprehensible logic.

But this raises the question of what logics are mutually comprehensible. In light of Maruyama's remarks it is clearly unreasonable to expect that any one logic be comprehensible across all disciplines. Nevertheless, it may be argued that there is at least one logic which is comprehensible among the mathematical disciplines. Specifically, the claim is that all mathematical discussions--inasmuch as they are indeed mathematical--always at some level revert back to the classical bimodal mode of reasoning.

This view is supported in part by the Intuitionist philosophy of mathematics, which asserts that, even when studying a formal mathematics, the mathematician necessarily makes use of an informal, or "intuitive," mathematics which is not part of the formalism in question (cf. Kleene 1952, pp. 46 ff). At the same time, this is a departure from Intuitionism inasmuch as the intuitionist logic, as proposed by Brouwer, is three-valued. Brouwer's view of the foundations of mathematics was that logical propositions only obtain their ultimate verity in the context of somewhat deeper conception of "constructibility" (or "finite computability"). Briefly stated (cf. Kleene 1967, p. 196), a proposition $\exists x P(x)$, which says that there exists a number x having a certain property P, can only

be asserted as "true" if one can show how to actually compute such an x in finitely many steps. Furthermore, logical negation is interpreted "positively," meaning that the proposition ∃xP(x) can be asserted as "false" only if the proposition ∃xP(x) leads to a contradiction in finitely many steps. This is tantamount to a denial of Excluded Middle in reference to infinite domains, since there are propositions P about the completed totality of natural numbers for which no such computation procedure is known. Thus the intuitionistic logic embodies the three values: "true," "false," and "undecided."

But now let us consider this logic further. Brouwer strongly resisted all attempts to axiomatize his logic, since in his view, logic is grounded in mathematics, and not conversely, so that the possibility of such an axiomatization would be largely irrelevant. Nevertheless, Kleene (cf. 1952) bothered to carry out this exercise, and the nature of his accomplishment is an important case in point. Kleene's axiomatization takes the form of a formal logical system in exactly the same sense as defined in this dissertation--i.e., as having its formulas be either derivable or nonderivable from its axioms--and it counts as a correct axiomatization of Brouwer's logic in that a proposition is so derivable if and only if it satisfies the intuitionistic criteria of constructibility. Thus we obtain a clear illustration of the present view. Kleene's work implicitly demonstrates that, even though intuitionistic logic is three-valued, there remains another level at which it is two-valued. Namely, Intuitionism is bimodal inasmuch as its propositions are either "constructible" or "nonconstructible."

The claim that bimodality applies to all mathematics is of course much more broadscoped than a reference to selected classical or nonclassical

logics. Nevertheless, the general claim is also justifiable by the
(introspective) observation that there is always a bimodal logic which
the mathematician ultimately makes use of when he seeks to make sense to
himself (which also parallels Intuitionism, cf. Heyting 1964, but on a
different level). This fact of mathematical life admittedly cannot be
proved in any conclusive fashion. Yet it is an observation which each
individual mathematician may make as a personal assessment of the under-
lying basis of his own manners of reasoning. Granted this, it then
follows that the classical logic is necessarily comprehensible among
mathematicians and therefore provides a solid basis for intermathematical
communication.

Affirmation of the inherent bimodality of mathematics does not,
however, deny that other modes of reasoning might be more "natural" or
at least more appropriate for describing certain kinds of ideas. On the
contrary, the foregoing analysis clearly supports the validity of apply-
ing the theory of formal logical systems in developing models based on
nonstandard logics. For the theory of formal logics is itself expressed
in the formal, bimodal, and hence mathematically communicable logic,
while at the same time providing a precise context within which non-
standard logics can be described.

Note furthermore that, as the foregoing discussion indicates,
where formalized nonclassical logics are concerned, bimodality can
usually be located at the metalevel, i.e., at the next higher level
that the one which is occupied by the formalism itself. This is a
fact that will be of especial use in §11.5, where we consider the
empirical testability of models which are based on nonclassical modes of
reasoning.

§10.3. Quality Representation

Qualitative mathematics was described in Chapter 1 as that kind of mathematics which leaves numerical considerations in the background and focuses on general conceptual forms. Examples were cited from topology, abstract algebra, and logic. We may here expand this idea insofar as it pertains to the logical systems studied in this dissertation.

The theory of approximate reasoning is especially suited for modeling qualitative aspects of complex systems. In contrast with traditional uses of formal logic for representing expressions in mathematical languages (such as the language of arithmetic, or of algebraic groups), approximate reasoning is aimed at the reppresentation of expressions in natural languages (eg., English). The extent to which this has been achieved to date is shown in recent works (Zadeh 1978a and 1978b) which lay down "translation rules" for translating a wide variety of more or less complex natural language expressions into a formalism of the kind studied in Part I. This means that one can use the theory of approximate reasoning for the direct formulation of _verbal_ _models_ of systems in an internally coherent, and machine implementable, mode of reasoning.

Such an approach leads to somewhat more realistic models than do the standard modeling techniques. For, an analysis of a complex system can never be carried to the deepest level of detail. Indeed, practical experience shows that during the process of a real-world system's analysis, the complexity of the system even appears to grow before one's eyes. Hence, every model of the given system must inadvertently "blur" all details which lie beyond the point at which the analysis was stopped. This is emphasized also by Bunge (1973, cf. §11.2) who states that

models are always "approximations" of reality, and that mathematical models in particular are "idealizations" of the object it is intended to represent. The virtue of approximate reasoning is therefore that this vagueness which permeates our models of reality is formally taken into account. An example of this was given in Chapter 6. The theory of approximate reasoning allows for reasonably sophisticated system properties and situations to be described, and the formalized logic embodies the necessary principles for correct reasoning toward meaningful results.

The Kleene-Varela systems studied in Part II illustrate another use of formal logic for quality representation: the modeling of paradoxical, and hence traditionally unanalyzable, system properties. This highlights the significance of formal logic as providing the opportunity to "select" alternative modes of reasoning. While there is nothing emperical which affirms the existence of paradoxical aspects of reality, there is also nothing emperical which denies that such aspects are possible. The traditional mathematics, however, forbids a priori that paradoxical qualities be discussed. Thus, for purposes of even formally "assuming" the existence of such qualities, shifting to a nonclassical mode of reasoning is absolutely required.

The use of nonclassical logics for such purposes as these presents two important methodological questions which have yet to be addressed. These are (i) whether mathematical logic provides the "semantics for science" in the sense that formal systems can be interpreted as meaningful representations of the observable world, and (ii) whether models based on nonstandard modes of reasoning can have viability in the sense of satisfying the conventional criteria for emperical testability. The

foregoing has obviously assumed that both of these questions can be
answered in the affirmative. The justification for this assumption is
developed in the chapter to follow.

§10.4. Formal Logical Systems And Their Semantics:
A General Introduction

An axiomatization of a body of mathematical knowledge lays down in
concise terms exactly those ideas which that knowledge is thought to
entail. In some cases, such as with Peano's axioms for elementary
arithmetic (cf. §10.1), this isolates and condenses the essential ingre-
dients of a well-established collection of mathematical facts and
principles. In other cases, such as with group theory, or the theory of
two-person games, the axiomatization is a generic definition of a parti-
cular class of mathematical structures, e.g., the class of groups, or a
class of games. In all such cases, however, the axiomatization serves
as the foundation for a particular mathematical "theory," consisting of
all propositions, or "theorems," that the given axioms logically imply.

Formal logical systems go a step beyond such axiomatizations, by
furthermore making explicit (i) a precise language within which all the
propositions of the given theory may be expressed, and (ii) an axiomati-
zation of the underlying mode of reasoning, according to which proposi-
tions may be derived from the given axioms. By virtue of (i), formal
logical systems exist on a level of abstraction in which the original
mathematical structures (e.g., the arithmetic of natural numbers, the
groups, etc.) become "interpretations" of the languages of those systems.
For example, the formal theory of groups has as two of its interpreta-
tions the positive integers under addition and the rotations of a
square in the Euclidean plane. In different interpretations, one simply

takes the symbols of the formal language as standing for different kinds of mathematical entities. This, together with (ii), renders the formal logical system as a representation of the given theory in terms of which the theory may be examined as to its consistency, its completeness, its decidability, and other properties related to the kinds of propositions that the theory allows to be derived. Such a representation of a mathematical theory is generally known as a "formalization" of that theory.

Precise definitions of "formal logical system," "interpretation," etc., are given in §2.1. The following two subsections elucidate these same ideas in a manner which may be useful for the reader to whom these ideas are new.

§10.4.1. Basic Ideas And Terminology

A formal logical system is comprised of: (i) a formal language; (ii) a set of axioms, expressible in that language; (iii) a collection of inference rules (or rules of formal derivation); and (iv) a set of theorems, which consist of (a) the given axioms, and (b) all expressions of the given language that can be generated from those axioms by means of the given inference rules. The basic structure of a formal logical system is represented in the top half of Figure 2-1. The essential characteristics of the various components may be described as follows.

A formal language is specified in two steps. First one selects a set of objects to serve as symbols. Strictly speaking, symbols are purely syntactical objects--that is, they are just abstract, independently existing objects, having no preassigned meanings. At the same time, however, a particular choice of symbols is always guided by a predelection of the kind of ideas that one will want to use the language

to express. For example, if the language is to express the properties
of "equality of numbers," one will want to have at least one symbol for
use as an "equality symbol," in which case it would be customary to use
the sign = as its denotation. But note especially that this sign,
consisting of two parallel, horizontal dashes on the page, is merely a
device for <u>denoting</u> that symbol, and strictly speaking is distinct from
the equality symbol itself. Symbols as herein discussed can neither be
seen nor written down; rather, they are purely conceptual entities which
are taken as "given," as the members of some sufficiently large collection
of abstract things (e.g., the ordinal numbers, cf. §2.1.1).

Given a set of symbols, one then automatically has the collection
of all finite sequences of those symbols, called <u>strings</u>. The second
step in specifying a language is to select from this collection of
strings some <u>formulas</u> (well-formed strings, etc., cf. §2.1), these being
all those strings having one of various "forms." Deciding on appropriate
forms is always guided by an intuition about what it takes for a formula
to be meaningful, given some appropriate interpretations of the string's
constituent symbols. For example, in a language for a formal system of
arithmetic, a string of the form a=+ would probably not be selected as
a formula, whereas a string of the form a=b might be, since the latter
string becomes meaningful if one makes the natural interpretation of =
as representing "equality of numbers," and of a and b as standing
for particular numbers, e.g., a pair of non-negative integers. The
formal language is then the above set of symbols together with this
resultant set of formulas.

Axioms are taken from the set of formulas, and <u>inference</u> <u>rules</u> are
definable as mappings from the set of formulas into itself. Inference

rules are so-called because they show how certain "conclusions" may be "inferred" from certain "hypotheses." By the definition of a formal logical system, the set of <u>theorems</u> is uniquely determined by the given language, axioms, and rules of formal derivation. Hence, the choice of axioms and inference rules for a particular formalization is always guided by a consideration of the kinds of theorems that one will want the system to allow. Continuing the above example, to formalize a theory of "equality of numbers," one will want enough axioms and inference rules to ensure that the theorems include all, and only, the usual properties of equality.

Now whereas the components of a formal logical system are purely syntactical entities--i.e., are symbols, strings of symbols, and rules for operating on strings of symbols--these components may obtain meanings by virtue of a <u>semantic interpretation</u> (alternatively, a <u>semantic structure</u>) for the language of that system. This consists of: (i) a universe of discourse, (ii) assignments from within that universe of specific meanings (or referents) for the symbols in the given language, and (iii) a precise definition of the sense in which the meaning of a formula, as determined by the meanings of the constituent symbols, is "valid" or "true."

The principal components of a semantic interpretation are represented in the bottom half of Figure 2-1. A <u>universe of discourse</u> is a set of objects, the members of which are called <u>individuals</u>. Strictly speaking, individuals are purely mathematical entities, e.g. numbers or sets of numbers. They may, however (cf. §11.2), sometimes be taken as representing concrete things, e.g., people, automobiles, or atoms.

The specific characteristics of meaning assignments depend on the linguistic roles that the symbols of the language are intended to play. For example, if the language contains some "individual constants" (cf. §3.1.1), then these symbols would be assigned as their meanings some specific individuals in the given universe, and if the language contains some "relation symbols" or "function symbols" (also cf. §3.1.1), these would be assigned as their meanings certain relations of individuals, or functions on individuals, in the given universe.

To illustrate: suppose that the language being considered is for a formal system of arithmetic, and that it contains individual constants, denoted by 0 and 1, an equality symbol, denoted by $=$, and an addition symbol, denoted by $+$. Then a natural interpretation would be one whose universe of discourse is the set of non-negative integers, and whose meaning assignments are: the numbers zero and one serve as the referents of the symbols 0 and 1, the referent of $=$ is the equality relation on the nonnegative integers, and the referent of $+$ is addition of nonnegative integers.

But note furthermore an important aspect of the relation between a formal language and its semantic interpretations. There is nothing here which requires that the domain of interpretation necessarily consist of numbers. On the contrary, an alternative interpretation could have as its universe of discourse a collection of sets, in which case 0 might be interpreted as the empty set, 1 as some maximal set, $=$ as equality of sets, and $+$ as set intersection. This reitterates a fact mentioned earlier, that a given formal language will in general have many different semantic interpretations, and it shows that the alternative interpretations

may in fact be quite different from the one (or several) which served
initially as the intuitive guide for that language's formal definition.

A notion of validity, or truth, is defined in terms of some truth
values and some truth valuations. Truth values can be any objects,
e.g., the letters T and F, or the numerals 1 and 0 for a bivalent inter-
pretation, or the set of reals in the interval [0,1] for a multivalent
interpretation, and so on. Truth valuations are mappings of formulas of
the given language into the set of truth values. These mappings may
sometimes be functions of the meanings of the formulas.
For example, if = is interpreted as the equality relation in the fore-
going illustration, then the valuations would be defined in such a way
that, for arbitrary individual constants a and b, a formula of the form
a=b is mapped to the value T (or 1) if and only if the given interpreta-
tion assigns the same individual as a meaning for both a and b. On the
other hand, truth valuations are the means by which the referents of
certain symbols are assigned. Specifically, these are the logical
connectives, such as ¬(not, negation), ∨ (or, disjunction), & (and,
conjunction), ⊃ (implies, logical implication), and ≡ (is equivalent
with; logical equivalence), which always have as their referents certain
functions from the set of truth values into itself. For example, in a
bivalent interpretation the meaning of & may be defined by: for any
formulas A and B, the formula of the form A&B receives the value T
under a valuation V--i.e., V(A&B)=T--if and only if both V(A)=T and
V(B)=T. By agreeing that all valuations in a given interpretation act
similarly on formulas built up with the symbol &, one thus captures the
sense in which, in classical logics, & is taken as representing the

English "and." In multivalent interpretations, of course, the meanings of the logical connectives are assigned differently.

A convention which has evolved for dealing with languages that contain individual constants and individual variables is to reserve the appellations "true" and "false" for only those formulas which make assertions about specific individuals in some universe (i.e., contain no individual variables which range freely over the universes of its interpretations). This convention is captured mathematically by having only such formulas be included within the domain of the truth-valuations. For example, the valuations would map the above formula a=b to the set of truth values, but not a formula of the form x=x. The latter type of formula is referred to as being "valid" or "invalid," where validity means that all possible substitutions of individual constants (which "name" the elements of a given universe, cf. §3.1 and §3.2) for the individual variables yield formulas in the domain of the valuations that are mapped to "true." If the formula is already in the domain of the valuations, then "true" and "valid" are used synonomously.

A semantics for a specific formal language is a class of interpretations for that language. Typically, a semantics is defined in terms of an arbitrary member of that class, in which case it is automatic that all the interpretations in that class have certain features in common. For example, they might all be bivalent, all assign the equality relation to the equality symbol (if indeed there is one in the given language), and all interpret the logical connectives in exactly the same way. A semantics for a class of formal languages is defined similarly. In this case it is customary for the languages in the class to have been defined generically--i.e., in terms of an arbitrary member of that class--and

hence to have a certain minimal set of symbols (sometimes called logical symbols) in common. Then an arbitrary semantic interpretation in the semantics for that class of languages may be defined in terms of the manner in which it assigns meanings to the logical symbols. The net effect is analagous with the situation of one language; the common symbols, which typically include a set of logical connectives, are interpreted in the same way by all interpretations for all the languages in the given class.

A formal logic consists of a formal language, or class of languages, together with a semantics for that language or class. In light of the above, this says that a formal logic is essentially determined by some specific meaning assignments for a specific set of symbols, in which case, use of the term "logic" signifies that this set includes some logical connectives. Hence this corresponds to the sense in which a logic is sometimes defined by some "truth tables" for logical connectives. Examples of logics appearing in this dissertation include the classical propositional calculus (§7.1), Kleene's three-valued logic for partial recursion (§8.2), and the logic of approximate reasoning (§3.1 and §3.2), which is a hybrid, having both bivalent and multivalent components.

Given the foregoing, a semantics for a class of formal logical systems is naturally defined as a semantics for the languages of these systems. In dealing with formal systems, however, it is generally the case that the systems in a particular class are defined analagously as a class of formal languages--i.e., in terms of an arbitrary member of the class--in which case the systems will have in common, not only a specific set of symbols, but also certain forms of axioms and inference rules, called logical axioms and logical rules. Then a semantics for

that class of systems will not only interpret logical symbols uniformly, but normally will also have it that the logical axioms are always valid (or true) and that the logical rules are validity- (or truth-) preserving in that they always yield valid conclusions from valid hypotheses. This in turn ensures that certain forms of theorems are always valid. For example, if a class of formal logical systems is to be based on the classical logic, then a semantics for that class would naturally be defined in such a way that all the tautologies of classical logic are semantically valid.

A semantic model of a formal logical system is a semantic interpretation for the language of that system in which all theorems of the system are valid. Evidently, by the definition of a formal logical system, in order for a semantic interpretation of its language to be a model of the system, it is sufficient that this be an interpretation in which all axioms of the system are valid, and all inference rules are validity preserving, since this will ensure that all formal derivations within the system lead only to valid theorems. Furthermore, it may happen that many interpretations for a formal system, even from within the same semantics, will not be models of that system. For example, a semantics for a formal system of arithmetic may have an interpretation for that system which assigns the meaning "one" to the symbol 1, "zero" to the symbol 0, and multiplication to the symbol +, in which case the theorem $1+0=1$ would not be valid. Finally, while it is immediate by the definition of model that the theorems of a formal logical system are always valid in every model of the system, the converse proposition need not be true--i.e., it is not necessarily true that a formula which is valid in every model of a formal system need be a theorem of that system.

This leads to the following important idea from the theory of formal systems.

Let F be a formal logical system, and let Σ be a semantics for F. Then F is said to be complete with respect to Σ if it turns out that a formula P in the language of F is valid in all models of F (in Σ) only if P is a theorem of F--i.e, if the abovementioned converse holds true. Accordingly, a class ℰ of formal logical systems is complete with respect to a semantics $ for ℰ if each formal system in ℰ is complete with respect to its interpretations in $.

The concept of semantic completeness makes precise the sense in which a class of formal logical systems is said to formalize a certain mathematical theory, or to characterize a certain collection of mathematical structures (i.e., the structures in the concern of that theory). If the formal systems are defined correctly, i.e., have the appropriate languages, axioms, and rules of inference, then the given mathematical structures will simply be the models in a properly defined class of semantic interpretations, in which case semantic completeness ensures that the axioms allow enough proof-theoretic strength to derive formally as theorems all formulas that express true propositions in the theory of those structures.

To avoid possible confusion with other works, here let us note also that a second usage of the word "complete" is sometimes employed. When expressed for systems whose languages include the standard logical connectives, a formal logical system F is proof-theoretically complete if, for every formula P of the language of F (or for every formula P of a certain form in that language) we have either that P is a theorem of F or that P is a theorem of F. It is known, for example, that every

axiomatization of the classical propositional calculus is complete in this sense. This proof-theoretic notion of completeness happens to not be used in this dissertation.

§10.4.2. Game Theory: An Example

At this point it will be convenient to illustrate the foregoing with an example that will also serve a purpose in the sections to follow. Suppose that it is desired to formalize the mathematical theory of games. (To date, no one appears to have carried out this exercise, but it is certainly reasonable to suppose that this could be accomplished with little difficulty). Here, "game" means any "two-person game" as studied by Rapoport (1966). Then it is implicit that the semantics of interest should include the entire class of two-person games; that is, these will be the "natural" interpretations of the formal systems we want to define.

To start out, therefore, we must select a symbol set having sufficient richness to express all the basic ideas of game theory. In particular, the symbol set should include the logical connectives denoted by, $\neg$, $\lor$, etc., the equality symbol denoted by =, some symbols to represent players, such as A and B, and so on, leading up to enough linguistic strength to describe strategy/outcome (or game) matrices and any functions which would represent each player's responses to the other player's moves. Among these would be a certain minimal set of logical symbols that are common to all the systems in the class being defined. Then, special ideas, unique to a specific system in the class, would be expressed by adding further symbols to this basic set.

Axioms and inference rules are chosen similarly. The axioms (or axiom "forms") and inference rules common to all the systems in the class should include (i) the axioms and rules for classical propositional calculus, (ii) the axioms for equality (i.e., standard axioms such as (a=b&b=c) a=c, which expresses transitivity of equality), together with (iii) enough further axioms and rules to describe the interrelations of players, strategies, rewards, etc. that must be present in <u>any</u> game. This provides us with a basic, or minimal, <u>formal theory of games</u>.

Further systems in the class may then be obtained by adding new symbols to the language, and by adding new axioms and/or inference rules defining how these new symbols interrelate with the basic symbols and with one another. For example, starting with the basic formal theory of games, one obtains the formal theory of zero-sum games by adding some axioms which require that the net outcome of any game be zero. On the other hand, starting with the same basic formal theory of games, one obtains the formal theory of nonzero-sum games by adding some axioms which explicitly state that the net outcome of any game should not equal zero. Furthermore, given the theory of nonzero-sum games, one may adjoin additional axioms which state that a game ultimately leads to the "prisoner's dilemma."

Note: A simple game matrix exhibiting the prisoner's dilemma is shown in Figure 10-1 (taken from Rapoport 1966, p. 128). This depicts a game wherein: (i) if the two contestants trust one another, they can cooperate and mutually choose strategies by which they mutually gain, while (ii) in a context of mutual distrust, and a desire on the part of each contestant to minimize his losses, the game becomes non-cooperative, and the contestants choose strategies by which they both lose.

Player B Stratagies

	B_1	B_2
A_1	5,5	-10,10
A_2	10,-10	-5,-5

Player A
Statagies

1st number is payoff to player A
2nd number is payoff to player B

Figure 10-1. Prisoner's dilemma game matrix.

The above remarks illustrate a process of obtaining formal systems of more and more specialized kinds of games by adding more symbols, axioms, and/or inference rules. But the fact of such specialization taking place is ensured only by virtue of a completeness result, i.e., by establishing that the class of all formal theories of games is complete with respect to the given semantics. As an illustration of this, let G stand for the minimal formal theory of games, let ZG stand for the extension of G which gives the formal theory of zero-sum games, let NZG be the analogous extension of G for nonzero-sum games, and let NZG+PD be the extension of NZG which furthermore contains axioms expressing the prisoner's dilemma (see Figure 10-2). Then a completeness result for the class of formal systems thus defined, with respect to a semantics which includes all possible games, would ensure all of the following: (i) a formula of the language of G is a theorem of G if and only if it is true for every game, (ii) a formula is a theorem of NZG if and only if it is true for every nonzero-sum game, (iii) a formula is a theorem of NZG+PD if and only if it is "true" for every game which leads to the prisoner's dilemma, and so on for ZG and any other systems based on G that one might want to define.

It follows that additional axioms for a logical system constrain the set of interpretations which can be models of that system. Models of NZG+PD are models of NZG, which in turn are models of G; models of NZG, however, need not be models of NZG+PD, and models of G need not be models of NZG. The latter models of G may instead be models of ZG. Thus the most general formal system--which may itself be called "game theory"--has the most models, while the less general systems have fewer
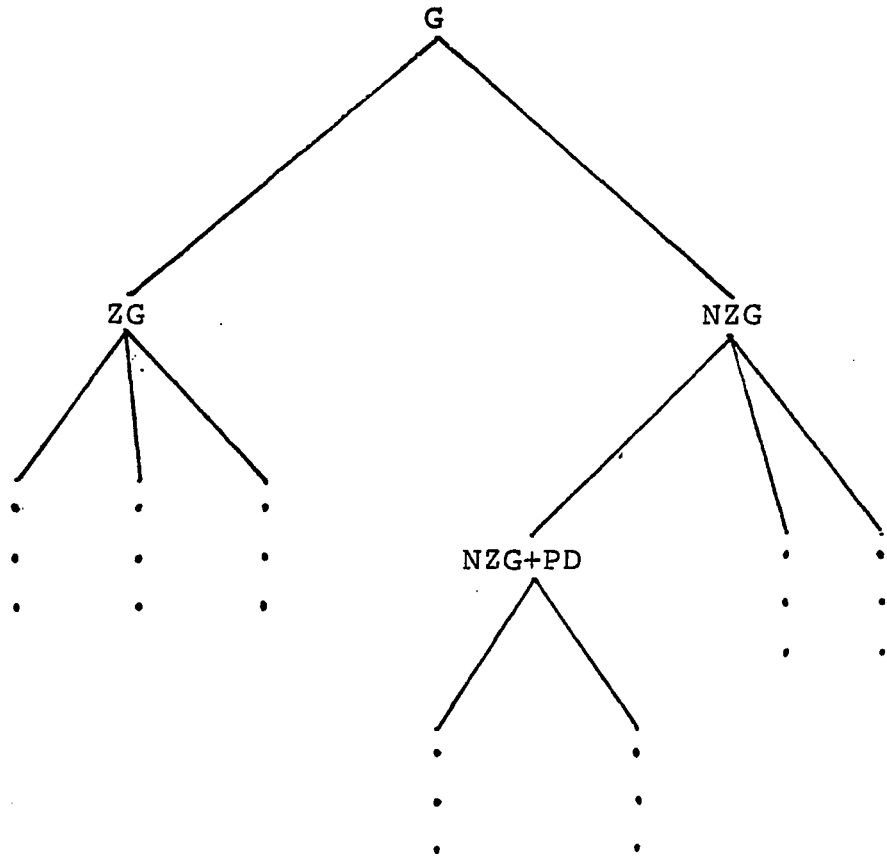
Figure 10-2: Adding new axioms to a formal system.

models. Today there is an extensive literature on techniques for determining the exact number of models of a given formal system.

§10.4.3. Further Terminology

Formal systems which have essentially only one model--i.e., all of whose models are "isomorphic"--are sometimes called categorical systems. A slightly weaker but sometimes more useful condition is that the system have essentially only one model of each finite or infinite cardinality-- i.e., for each finite or infinite cardinal number c, all the models whose universes have cardinality c are isomorphic. This property for a particular c is sometimes referred to as categoricity in c. For a brief discussion of these ideas, see §5.6.

A formal logical system is said to be consistent if its set of theorems does not include a contradiction (or an "inconsistency," cf. §2.1.5.2). In standard languages, i.e., languages which employ the usual logical connectives, a contradiction will be a formula of the form P&¬P, while in nonstandard languages, contradictions will normally have different forms (as in §7.3).

A formal logical system is coherent if there is at least one formula in the language of that system which is not a theorem of the system. The property of coherence leads to a broad division of all formal logical systems into two categories: Aristotelean systems, in which inconsistency always implies incoherence, and non-Aristotelean systems, which can be simultaneously inconsistent and coherent. A discussion of these ideas, and of how they relate to the literature, is given in §2.1.5.2 and §2.1.5.4. Examples of their use appear in §8.4.

CHAPTER 11

THE METHODOLOGICAL FRAME

Chapter 10 discussed the foundational aspects of using nonclassical logics as alternative guides to thinking. The present chapter continues this discussion in an investigation of methodological issues surrounding the use of formal logical systems and, particularly, nonclassical modes of reasoning for constructing models of real-world systems.

In §11.1 we review M. Bunge's work on the theory of modeling in the natural sciences, recounting basic definitions and explaining the concept of empirical testability.

§11.2 then takes up the views of Bunge and Suppes regarding the role of formal systems in providing the semantics for science, with the aim of clarifying certain points for purposes of the discussion at hand.

§11.3, makes use of §11.2 to show how Bunge's basic definitions (§11.1) can be rendered within the theory of formal systems.

§11.4 carries this one step further, by recalling the discussion of Chapter 10 and laying down the manner in which formal systems can enter into the relation of the observer to the observed.

§11.5 examines the question of empirical testability as it applies to models based on nonstandard modes of reasoning. It is concluded that Bunge's criteria apply, although in a more or less qualified fashion depending on the particular logic employed.

§11.1. Review Of Bunge's Theory Of Modeling

This section recounts the central ideas of Bunge (1973) on the use
of mathematical models in the natural sciences, insofar as these pertain
to the questions mentioned in the final paragraph of §10.3. Namely,
these concern the status of the theory of formal logical systems in
providing the "semantics for science," and the empirical testability of
models that are based on nonstandard modes of reasoning.

We begin with a summary of basic definitions. A _model_ here is any
entity which performs "a partial representation of a chunk of reality"
(Bunge, p. 19). A _model object_ is a model which serves as "a schematic
representation of a thing or of a situation assumed to be actual or
possible" (p. 97). A _theoretical model_ (or _specific theory_) is a model
which is a "hypothetico-deductive system concerning a model object"
(p. 97) and which is obtained by "grafting" a model object onto a general
theory--where a _general theory_ is a body of laws, or a "generic frame-
work" which is "concerned with the gross structure and behavior of
systems of almost any kind, regardless of their physical and chemical
constitution" (pp. 33-34). The interrelation of these ideas is
illustrated in Figure 11-1.

As examples of general theories Bunge cites "information theory,
game theory, systems theory (in particular cybernetics), automata theory,
and other extremely general black box theories, as well as nonspecific
mechanism theories such as general field theory, both classical and
quantal" (p. 33). Such theories are regarded as "generic sets" of
general descriptions and laws which describe the structure and behavior
of a very large (usually infinite) collection of set-theoretical struc-
tures (eg., game matrices). Alternatively, general theories may be

General Theory
(Non-Zero-Sum Two-Person Games)

Specific
Theory

Model Object
(Prisoner's Dilemma)

Concrete
Referent

Real-World
System

(Actual Game)
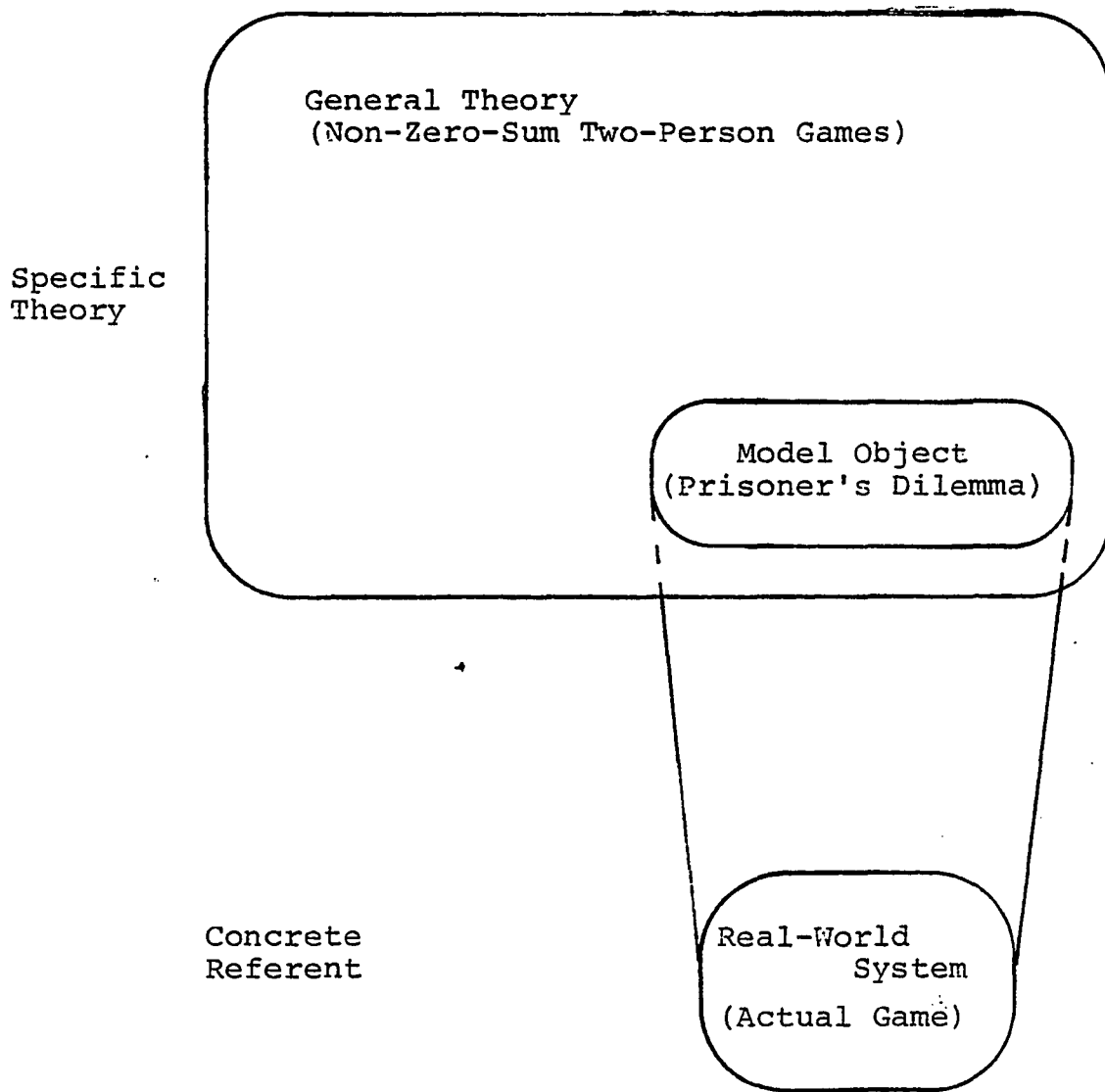
Figure 11-1. Structure of a theoretical model (specific theory).

regarded as containing the collection of set-theoretical structures themselves, in which case the general descriptions are taken only as technical devices by which the collection of mathematical structures is defined. The laws of course remain separate from the structures.

In either case, general theories are "stuff free" in the sense that their set-theoretic structures are not assumed to have any specific referents in the real world. They are merely uninterpreted abstract mathematical entities. It follows that general theories are not completely testable, since they do not have the requirement described by Popper of being refutable. According to this conception, a model is refutable only if it says something about a real-world system, in terms of which the model can be determined as true or false. But, general theories are neither "true" nor "false." Rather, a set-theoretic structure only either "applies" to a real-world situation, or it does not, according as that situation does or does not satisfy the definitions of the kinds of situations which are under that general theory's concern. For example, a real-world situation is either a two-person game, or it is not, depending on whether it satisfies the definition of a two-person game; and if it does not, then the general theory of two-person games remains unchallenged--one simply has here a situation in which two-person game theory is inapplicable as an analytical tool.

At the same time, however, general theories are partially testable in that they are confirmable, where confirmation is enacted through the testing of theoretical models (or specific theories) obtained by embedding specific model objects within that general theory's framework. As greater numbers of specific theories are verified within such a framework, the framework itself becomes increasingly more confirmed as a useful and

hence viable scientific theory. Complete confirmation, of course, is normally not possible since this would entail the testing of a usually infinite variety of specific theories.

A typical example of a model object would be a set-theoretic description of the prisoner's dilemma game matrix given in §10.4.2, applied to a real-world two-person game. In contrast with general theories, model objects are always assumed to have specific real-world referents. In Bunge's work, where M is a model object and R is a real-world referent, this modeling relation is denoted by the expression M≙R, and is referred to as an interpretative axiom or a semantic assumption (p. 96). Semantic assumptions are always implicit in informal scientific discourse, but must be made explicit in the foundations of science and in particular in the axiomatization of a scientific theory. Such explication is important since any given model object might be involved in multiple semantic assumptions. For example, (i) a pair of competing corporations who have the option of choosing between continued competition or a merger, and (ii) a pair of countries who have adopted a mutual status of detente, might constitute two referents of the same set-theoretic description of the prisoners dilemma. Expression of a semantic assumption thus makes clear which real-world referents are in concern.

A characteristic of model objects is that they are "idealizations" of their referents (p. 92). As such, model objects always omit certain details that are deemed to be irrelevant to the discussion at hand. Quoting Bunge (p. 92), "all the individuals of a given mice strain may be taken to be indiscernable and all ways of pressing a bar for food pellets may be assumed to be equivalent as well. In other words, the

real population, made up of different individuals, is modeled as a homo-geneous set (an equivalence class), and likewise the set of all possible events is partitioned into homogeneous (equivalence) classes". It is in this sense that model objects are always "conceptual" or "abstract."

As with general theories, model objects are untestable in the sense of being irrefutable, but for a different reason. Although they are always assumed to have at least one real-world referent, model objects in themselves make no predictions. But predictions are necessary for refutability, since it is only in terms of a model's predictions that one ultimately determines whether that model is true or false.

Model objects do make predictions, however, when they are embedded within (or grafted onto) the framework provided by a general theory. In this manner the model object inherits a set of general laws according to which predictions can be inferred as logical implications of the model object. This in fact is what one tests in actual practice--i.e., a theoretical model, or specific theory, of the model object.

This highlights the major conclusion of Bunge's analysis: for a model to be completely testable it is necessary and sufficient that it be both refutable and confirmable. Hence only theoretical models are fully testable. The confirmability of theoretical models stems from their having specific real-world referents, and their refutability stems from their capacity to make specific predictions regarding their referents' behavior. Insofar as standard approaches to mathematical modeling are concerned, it may therefore be said that the criteria of empirical testability are fairly well secured.

§11.2. The Semantics Of Science

Let us now turn to a comparison of the scientific notion of "theoretical model" with the purely mathematical notion of "model of a formal logical system." The import of Bunge's remarks (pp. 110-113) is that an earlier view set forth by Suppes (1961) makes too strong of a case for the potential role of mathematical logic in scientific studies. Specifically, Bunge qualifies Suppes's view with the observation that the theory of formal logical systems and their semantics is not in itself enough to fully constitute the "semantics of science." Although Bunge does not provide an explicit definition of this phrase, it is clear from the context of his discussion that this semantics should include all of the linguistic and semantic apparatus necessary for the formulation and testing of specific theories. In this section we shall ultimately concur with Bunge on this point and then (in this and later sections) proceed to build on Bunge's analysis so as to better pinpoint the use of formal systems at least in serving the semantics of science even though not comprising its entirety.

It will be instructive to start with a brief summary of Suppes's contribution. The paper begins with a series of quotations, the first being Tarski's definition of a semantic model of a formal theory (i.e., of a formal system in the sense of this dissertation), and the remaining five quotes being descriptions of models used in the empirical sciences. The crux of his argument is then given as follows (pp. 165-166):

> I claim that the concept of model in the sense of Tarski may be
> used without distortion and as a fundamental concept in all of
> the disciplines from which the above quotations are drawn. In
> this sense I would assert that the meaning of the concept of
> model is the same in mathematics and the empirical sciences.
> The difference to be found in these disciplines is to be found
> in their use of the concept. In drawing this comparison between

constancy of meaning and difference of use, the sometimes diffi-
cult semantic question of how one is to explain the meaning of a
concept without referring to its use does not actually arise.
When I speak of the meaning of the concept of a model I shall
always be speaking in well-defined technical contexts and what I
shall be claiming is that, given this technical meaning of the
concept of model, mathematicians ask a certain kind of question
about models and empirical scientists tend to ask another kind
of question.

The support for this claim stems from some observations about the

nature of set-theoretical models of formal systems and their connection

with physical systems. These observations are illustrated by an example

as follows (pp. 166-167):

> It is true that many physicists want to think of a model of
> the orbital theory of the atom as being more than a certain kind
> of set-theoretical entity. They envisage it as a very concrete
> physical thing built on the analogy of the solar system. I
> think it is important to point out that there is no real incom-
> patibility in these two viewpoints. To define formally a model
> as a set-theoretical entity which is a certain kind of ordered
> tuple consisting of a set of objects and relations and operations
> on these objects is not to rule out the physical model of the
> kind which is appealing to physicists, for the physical model
> may be simply taken to define the set of objects in the set-
> theoretical model. . . .
> It is simple enough to see how an actual physical model in the
> physicist's sense of classical particle mechanics is related to
> this set-theoretical sense of models. We simply can take the
> set of particles to be in the case of the solar system the set
> of planetary bodies. Another slightly more abstract possibility
> is to take the set of particles to be the set of centers of mass
> of the planetary bodies. This generally exemplifies the situa-
> tion. The abstract set-theoretical model of a theory will have
> among its parts a basic set which will consist of the objects
> ordinarily thought to constitute the physical model. . . .

A summarization of these remarks is represented pictorially in

Figure 11-2. That this figure is implicit, although somewhat obscurely,

in Suppes's remarks is evidenced by (i) the statement that "the physical

model may be taken to define the set of objects in the set-theoretical

model," (ii) the statement that one can "take the set of particles to be

the centers of mass of the planetary bodies," and particularly (iii) the
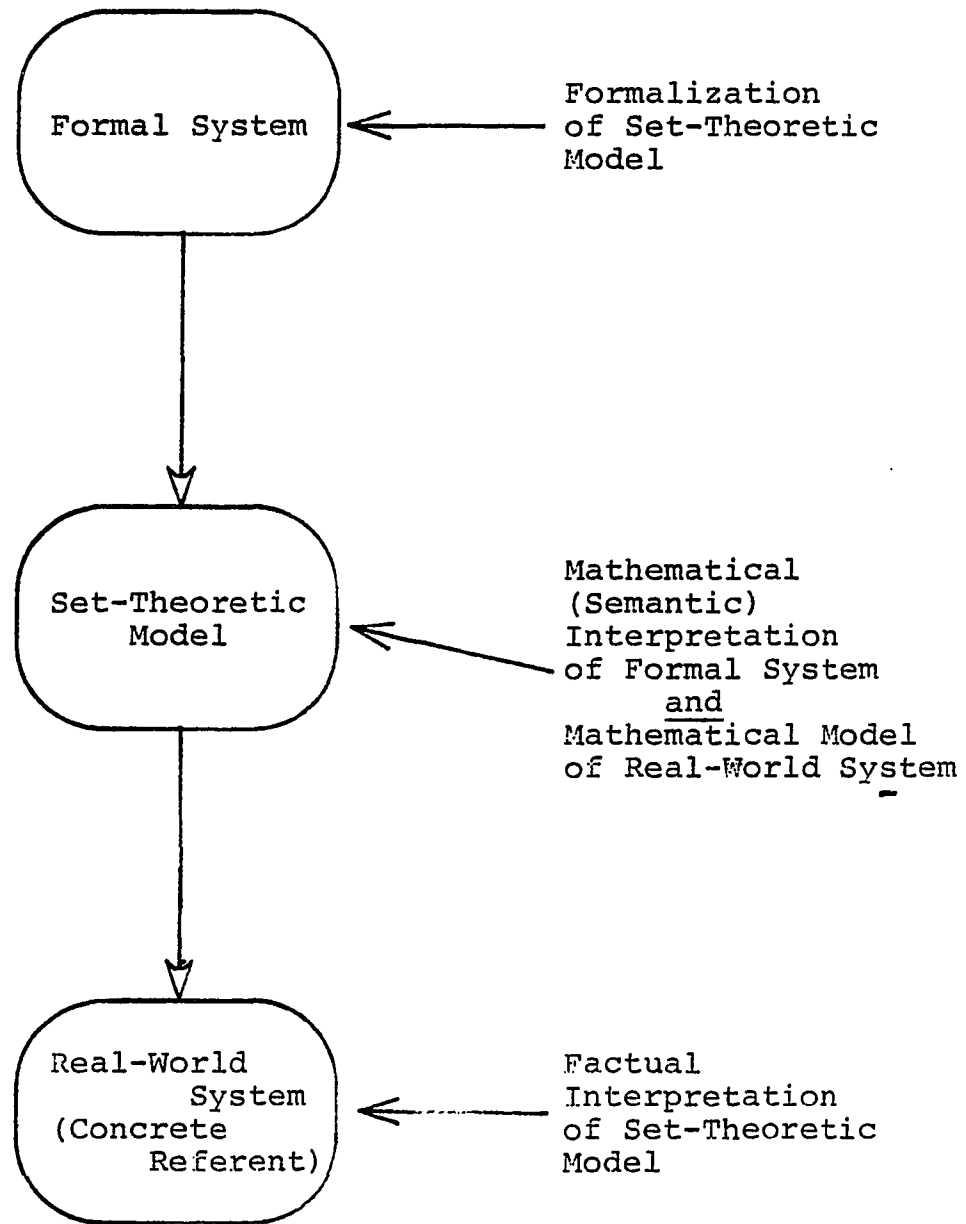
Figure 11-2. Suppes's view of the modeling relations.

statement that the "basic set. . . will consist of the objects <u>ordinarily</u> <u>thought</u> to constitute the physical model." In the context of §11.1, it is here evident that Suppes is referring to the fact that the elements of a model are "idealizations" of the elements of the physical system.

As it stands, this may be regarded as a fairly accurate description of the actual state of affairs. Yet, as Bunge points out, it lacks a small, but nonetheless important, matter of detail; namely, there is nothing inherent in formal systems or their semantic interpretations which connects them with the empirical world. While the upper half of Figure 11-2 is always an essential ingredient of the mathematics of the model, the lower half represents something which is "extramathematical," i.e., something which must be added to the set-theoretical structure in order that it become a bona fide theoretical model of a concrete (or supposedly existing) real-world system.

Observing this, Bunge proceeds to cast an alternate view in terms of some further examples. These shall be given here in detail, since they will serve a purpose also in later sections.

Consider the following system of symbols and axioms (Bunge, p. 111):

1. $S \neq \phi$.

2. (a)$F:S \rightarrow R$. (b)$G:S \times S \rightarrow R$. (c)$H:S \times S \rightarrow R$.

3. $s,s' \varepsilon S \Rightarrow H(s,s')=h \varepsilon R$.

4. (a)$\textbf{o}:R \times R \rightarrow R$. (b)$\square:R \times R \rightarrow R$.

5. $s,s' \varepsilon S \Rightarrow G(s,s')=h\textbf{o}[F(s')\square F(s)]$

In the following this system will be referred to as a "formal system" even though, strictly speaking, is only the proper part of a formal logical system which would furthermore embody some axioms characterizing an

underlying logic, together with some axioms which lay down the formal

interrelations of ϕ, =, ×, O, □, etc.

Bunge notes first of all that this system is "nonsignificant"

since we may assign the symbols as many different meanings as we please.

A specific semantics (as in §10.4.1) is provided, however, by an agreement

that under all interpretations we shall have that R stands for the real

line, O for the arithmetic product, □ for subtraction, and that the

remaining symbols and punctuation marks receive their conventional

interpretations. It follows that all the interpretations in this seman-

tics will have the properties (Bunge, p. 111):

1. S is a non-empty set.

2. (a) F is a real valued function on S. (b) G and H are real valued

    functions on the set of pairs of members of S.

3. H is the constant function with value h.

4. For every s and s' in S, $G(s,s')=h[F(s')-F(s)]$.

Different interpretations thus are obtainable by assigning different

sets to S, different functions to F, and different real values to h.

Given an assignment for S,F and h, the meanings of G and H are determined.

Any such set-theoretic structure, Bunge refers to as a set-

theoretical "formalism." His argument then proceeds (p. 112):

> This is a formalism interpreted within pure mathematics. It
> makes no sense outside mathematics. In particular, it is not a
> theoretical model in any metascientific sense of the term, for
> it does not concern anything extramathematical: the basic set S
> is an arbitrary (abstract) set and therefore the functions F, G,
> and H cannot represent any concrete properties. Precisely this
> renders the formalism valuable from a scientific point of view,
> for it is a read-made dummy that can be clothed in a number of
> ways.

Such "clothings" of the set-theoretic formalism are called "factual

interpretations," two examples of which are given as follows:

1. A "physical interpretation," wherein

   a. the individuals s of the universe of discourse are thought of as points on a d.c. electric circuit,

   b. F(s) is electrical potential at the point s,

   c. G(s,s') is intensity of current between s and s',

   d. H(s,s') is conductivity between s and s',

2. A "sociological interpretation," wherein

   a. the individuals s of the universe of discourse are thought of as countries,

   b. F(s) is enticement offered by s (eg., standard of living),

   c. G(s,s') is migratory pressure from s to s',

   d. H(s,s') is permeability of the border between s and s'.

This set of examples calls to mind a slightly different image than Suppes's, as shown in Figure 11-3. Here we have a formal system interpreted within mathematics as a certain set-theoretical structure, which is then further provided with "clothes" by which it becomes simultaneously a factual interpretation of the set-theoretical "formalism" and a theoretical model of a real-world system.

Yet, Bunge continues, "this is just a first approximation." Mathematical models of formal systems are always, by definition (cf. §10.4.1), required to be "true" for all the theorems of the system, while theoretical models are at best only "approximately true" of their designated real-world systems (since they are always "idealizations").

> Consequently every theoretical model is, in the best of cases, a quasimodel in the sense that its formulas are (at best) satisfied only approximately by reality. Therefore the model-theoretic concept of a model does not coincide with either of the two metascientific notions of model. Which shows that model theory is not enough to consitute the semantics of science, and suggests
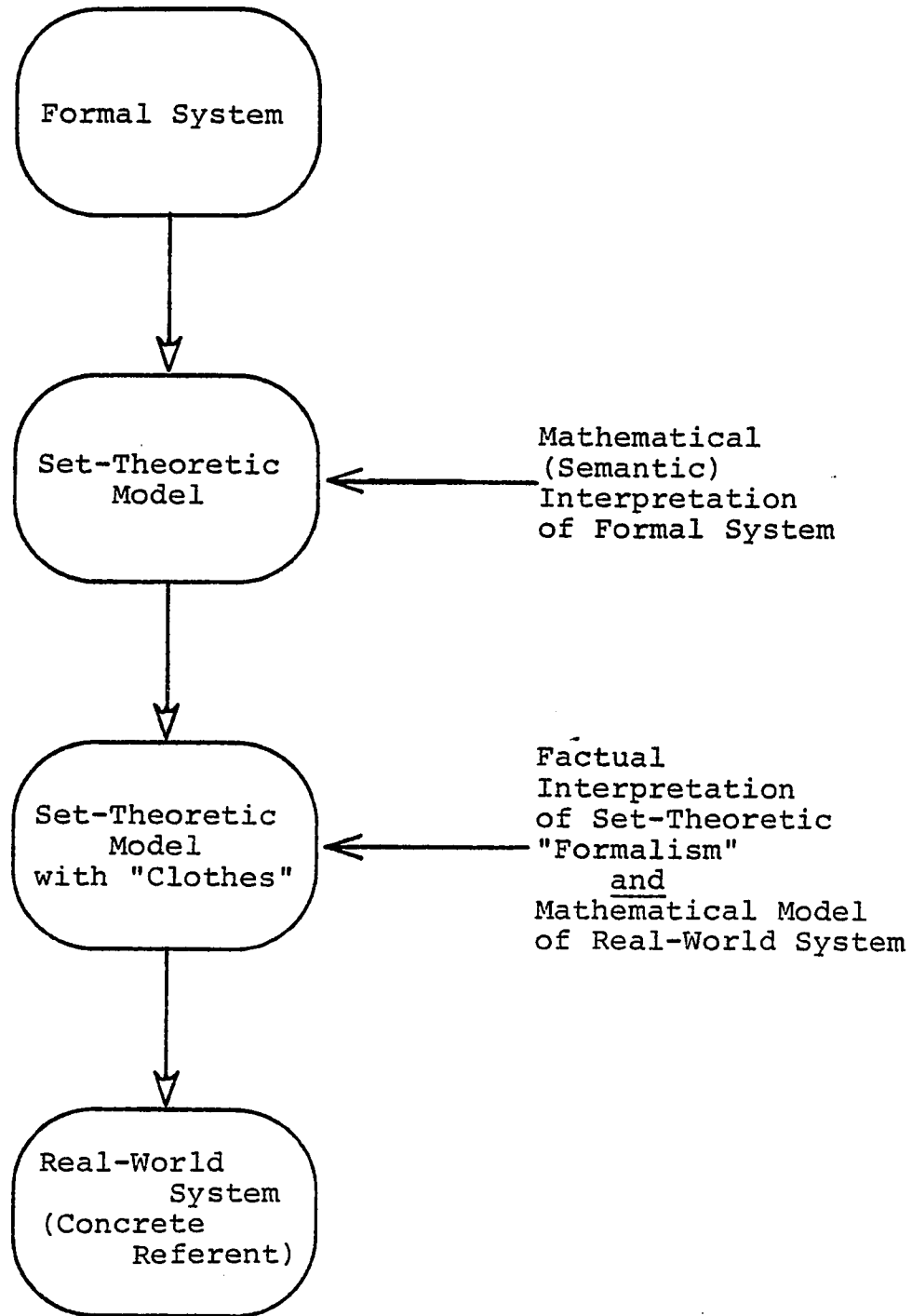
Figure 11-3. Bunge's view of the modeling relations.

that the very term 'theoretical model' (and also 'mathematical model') would be advantageously replaced by 'specific theory.'

Now on these points Bunge's position evidently stands firm. In fact, it is likely that this would gain the concurrence of Suppes himself. Indeed Bunge's rebuttal of Suppes may be somewhat strongly stated, since Bunge merely extends the analysis in a way which augments but does not necessarily contradict Suppes. That Suppes overlooked the above distinction between the "exactness" and "inexactness" of the separate modeling relations apparently stems from his focusing on the set-theoretical entity itself, rather than on the nature of the implicit semantic assumptions. Nevertheless, Bunge's point that development a specific theory out of set-theoretical structure always requires something "extramathematical" is significant.

Regarding the nature of theoretical models thus construed, however, it is worth noting also that there is a further difference between the relations of factual interpretation to set-theoretical structure and of set-theoretical interpretation to formal system. This may be unraveled as follows. In the latter relation, the set-theoretical structure is an entity which is completely separate from the system which it interprets-- i.e., it contains no symbols or symbolic elements, but only set-theoretical objects together with some interpretative mappings (cf. §10.4.1) from symbolic elements into set-theoretical entities. In the former relation, on the other hand, the set theoretical structure (or some fragment thereof) is properly contained within the factual interpretation--i.e., correctly envisioned, the theoretical model consists of at least a part of the set-theoretical structure together with some further mappings which link set-theoretical entities with objects in the real world. On

this count, it may be argued that one could with no loss revert back to a three-leveled diagram such as Figure 11-2, as long as one keeps in mind that it is the lower part, i.e., the semantic assumption, which indeed qualifies the semantic model of the formal system as a representative of some real-world system. This three-leveled image will be used in the work below.

Last, while Bunge's work clearly does not deny the usefulness of formal systems to science, his rendering of "theoretical model" obscures a point that is central to all that follows here. Namely, that the presence of such a system, either implicitly understood or explicitly formalized, is essential to the model for providing both the linguistic and the logico-deductive mechanism. That this is so stems from there being no symbolic or logical apparatus at the level of the set-theoretical model. These features, by which the factual model obtains "formulas" and implies "predictions," exists only at the level which is represented by the formal system. To practicing scientists, of course, this distinction is seldom relevant, since first, the underlying logic is usually the classical logic, whose verity is normally unquestioned (cf. §10.2), and second, even in standard mathematical theories the possibility of separating symbols from their meanings is unimportant.

Nevertheless, as a further step toward making the notion of theoretical model complete in this context, it is worth observing in addition to the above that not only do such models properly contain a part of their set-theoretical "formalism," but they also contain at least a fragment of some formal system, namely that part which expresses the properties and behaviors of the given set-theoretical model. This understanding will be tacit in the sections to follow.

§11.3. A Formalization Of The Basic Modeling Ideas

Now that we have clarified the role of formal logical systems with respect to the semantics of science, we are in a position to lend further detail to this role by reinterpreting the analysis of Bunge (§11.1) within this frame. Here it will be shown that the theory of formal logical systems contributes a new level of precision to the concepts of model object, theoretical model and general theory, and, more importantly, that the "exactness" of the relation of formal system to its semantic models is precisely what lends theoretical models their property of refutability.

In §11.1 it was seen that the idea of a general theory is exemplified by the theory of two-person games, and that this theory may be regarded as the collection of all possible set-theoretical structures that satisfy a generic set of definitions and laws. Furthermore, it was seen that the idea of a model object is exemplified by a set-theoretical structure which represents a prisoner's dilemma situation in the real world, whereas a specific theory of that model object is obtained by "grafting" that structure onto the general theory of two-person games. As part of the following, we make this notion of "grafting" formally explict. It will be convenient to take as our theory the theory of nonzero-sum two-person games which, although somewhat less general than the full theory of two-person games, nonetheless qualifies as an example of a general theory.

We may proceed to translate Bunge's definitions into the terminology of formal systems in the following manner. Consider the formal system NZG of §10.4.2, and let the collection of all possible semantic

interpretations of NZG be the semantics Σ(NZG). Then the general theory of nonzero-sum two person games may be defined as NZG together with the collection of all interpretations in Σ(NZG) which are models of NZG (in the "exact" model-theoretic sense). Note that this merely formalizes the generic set of descriptions and laws which (§11.1) normally comprise the informal general theory, and all of the set-theoretical entities thus described are correctly characterized, i.e., it is assumed that NZG is semantically complete with respect to Σ(NZG). The sense in which a general theory is "stuff free" is here represented by the fact that the semantic interpretations in Σ(NZG) are not assumed to have any particular referents in the real world.

The manner in which a general theory obtains its connection with the real world, through the use of a model object, can be elucidated by similarly translating the definitions of model object and specific theory. A model object of a real-world system which exhibits the prisoner's dilemma may be defined as the axioms PD (described in §10.4.2), together with any part of a semantic model in Σ(NZG) that satisfies those axioms, and furthermore together with any semantic assumption which maps that part of the semantic model into the real world. Note that here the model object per se usually does not utilize the full semantic interpretation, but only a certain subcollection of functions and relations on the universe of that interpretation. Furthermore note that there might well be more than one model in Σ(NZG) that satisfy the axioms PD (i.e., in which PD is "true" thereby possibly giving rise to more than one model object. Such model objects may or may not be equivalent, either from the standpoints of the axioms PD or that of and the concrete referent.

A <u>specific</u> <u>theory</u> of a prisoner's dilemma model object may then be defined as the formal theory NZG+PD together with the smaller semantics Σ(NZG+PD), and furthermore including any semantic assumptions in which the interpretations in Σ(NZG+PD) might be involved (i.e., as containing model objects). This illustrates that the process of "grafting" a model object onto a general theory may be represented formally as a process of "adjoining" some additional axioms to a formal logical system. In case there is only one such model in Σ(NZG+PD)--i.e., if NZG+PD is categorical in the sense of §10.4.3--then the specific theory is just that model; otherwise the specific theory consists of a collection of more or less similar models.

This collection of definitions has been summarized schematically in Figure 11-4. Here is illustrated the formation of the theory NZG+PD as the composition of NZG and PD. Further it is indicated that the semantics for NZG will in general be larger than the semantics for NZG+PD (as was discussed in §10.4.2) and that the set-theoretical part of the model object is normally a part of a particular interpretation, here denoted by I_4, within the semantics Σ(NZG+PD). The presence of I_3 in the diagram covers the possibility that there might be similar but nonisomorphic models of NZG+PD--i.e., that NZG+PD might not be categorical. The lower part of the diagram shows the semantic assumption which links the set-theoretical part of the model object with the real world, and furthermore suggests that the specific theory of that model object will normally pertain to some larger system within which the concrete referent is embedded. Thus we see that all the essential aspects of Bunge's definitions are taken into account.
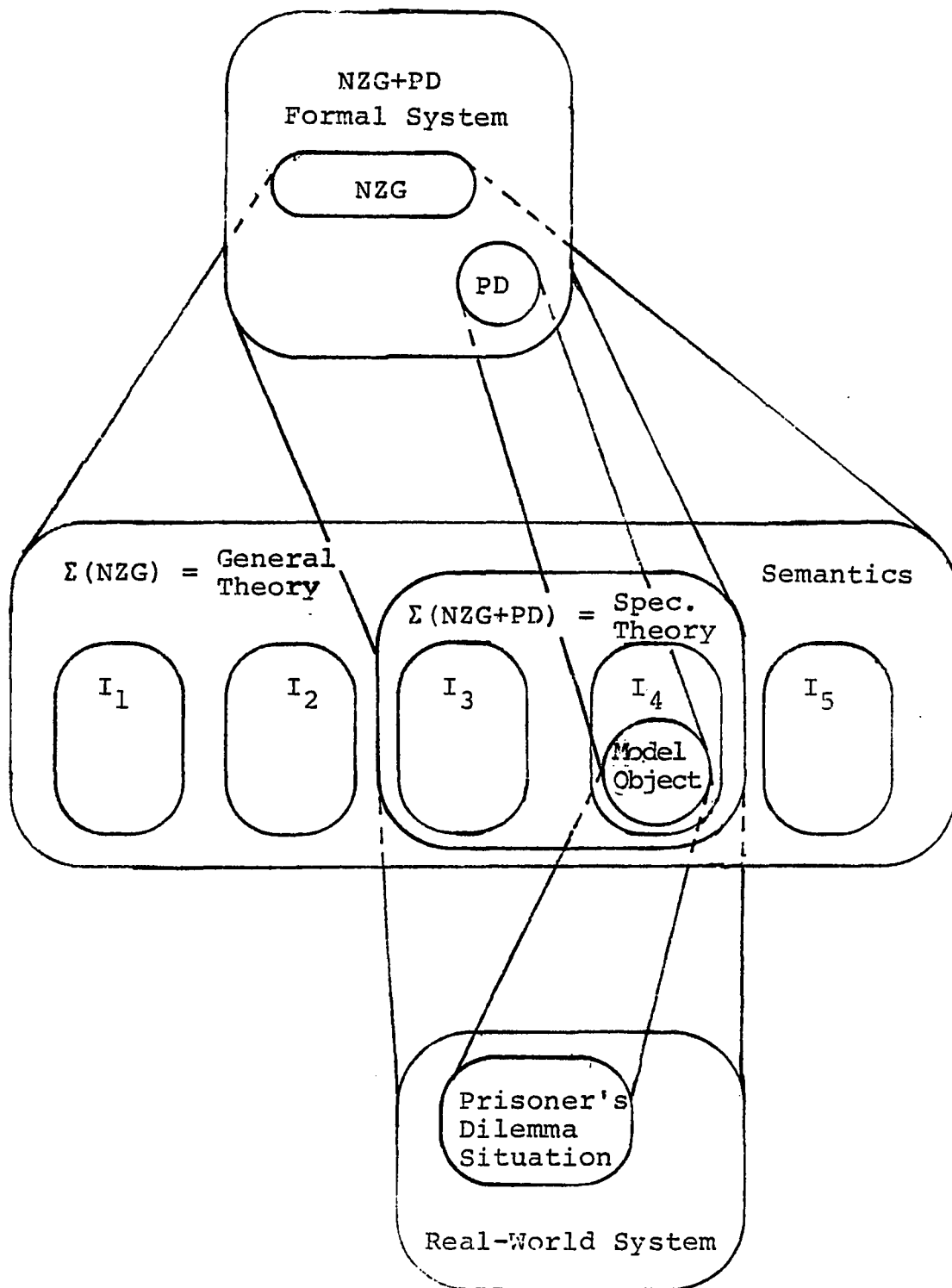
Figure 11-4. Formal representation of modeling ideas.

Next let us consider the concept of refutability in this context.
Recall from §11.1 that the requirements for a model to be refutable are
(i) that the model have a specific referent in the real-world, and
(ii) that it allow for making predictions about its referent. Then the
fact that a specific theory of a model object has an empirical referent
will always be--as we have seen--"extramathematical," since the intro-
duction of a referent is based only on a choice by the modeler. The
nature of a prediction, on the other hand, is intrinsically mathematical
and is based on the characteristics of the general theory. We may
therefore define this concept in a continuation of the above example as
follows.

By a prediction of the above specific theory (of the prisoner's
dilemma) is meant the semantic interpretation (i.e., the meaning) of any
theorem of NZG+PD in the semantic model of NZG+PD which contains the
given model object. In other words, a prediction is simply a logical
consequence of axioms of the formalized general theory, interpreted
within that theory's semantics. This clearly makes precise the signi-
ficance of general theories in providing logico-deductive systems. And
the concept of a refutable model thus becomes precise via a precise
rendering of the concept of prediction.

Actual verification or refutation of a given specific theory is,
of course, always a matter of scientific investigation. One tests the
predictions of the theory against the behavior of the concrete referent
and thereby determines whether the predictions are "true," "false,"
"approximately true," and so on. Clearly, these predictions will normally
be only approximately true of the real-world system. However, the
"exactness" in the relation of the semantics of Σ(NZG+PD) to the formal

system NZG+PD is mandatory for the investigative procedure to be correct.

For it is only because a prediction is the "meaning" of a theorem in

NZG+PD that one is assured that that prediction is indeed a valid logical

inference from the given model object. Without the presence (either

implicitly or explicitly) of a formal logical system, no such correctness

can be ensured.

Semantic completeness of the formal logical system is also an

implicit necessity, in order to ensure that the formal system indeed

characterizes the given general theory. Categoricity, on the other

hand, might or might not play a role. If the formal system is cate-

gorical, then one knows that the specific theory utilizes the only

semantic model which satisfies the axioms of the given model object, in

which case a formula of the language of the system is derivable from the

axioms of the system if and only if it is true for that particular

specific theory. But even if the formal system is not categorical, we

still have that all derivable formulas are true for the specific theory,

by the definition of semantic model; only we do not have the converse.

In terms of testability, this means that simply because a formula is

semantically valid, we do not know if it qualifies as a prediction,

since only the formally derivable formulas are bona fide logical

inferences within the context of the general theory. In actual practice

however, this fact of mathematical life is of no real consequence, since

scientific testing is always restricted to only those predictions which

are indeed formally derivable. Thus, categoricity is not generally

needed for testability, even though it might be useful in some situations.

Finally, it is true that in practice the underlying logical system

is rarely made formally explicit; and in most cases this explication

serves no useful purpose. The usual approach in scientific work is to simply assume a given, commonly known, body of knowledge and to explicitly present only the "proper axioms" which describe the specific theory in concern. This is what was done, for example, in Bunge's description of a formal system in terms of the axioms 1 through 5 (§11.2). But, this does not negate the fact that there is always some formal logic, indeed usually the classical logic, permeating the entire investigation.

§11.4. The Use Of Formal Systems

With these ideas in hand we may now bring forward the discussion of §10.1 and §10.2 and bear more carefully on the manner in which formal systems may enter into the relation between the observer and the observed. Let us begin with a brief review of the way in which models are formed.

We have seen that the use of a general theory is always necessary for the modeling exercise, for it is only when a model object is grafted onto a general theory, thereby forming a theoretical model, that it becomes fully testable. At the same time, however, general theories enter into the modeling procedure in another way. For when undertaking to study a given real-world system, some general theory is needed a priori as a guide in prescribing the kinds of data and measurements one is to obtain. For example, the kind of research conducted will be quite different if one regards the system as an instance of information theory as opposed to the theory of games. Oftentimes, of course, the characteristics of interest are not adequately representable with the existing theories, in which case there develops an interest in formulating a new

theory. But in any case, the existence of some general theory, either actual or possible, is implied.

Given a general theory, the main steps toward model construction may be outlined as follows. First, observational data is gleaned from the real-world system, by some appropriate means--visual observation, electron scattering, and so forth. This data is perceived by the observer in terms of measurements, interrelations of behaviors, and so on. Second, these perceptions are formulated within the language of the general theory, thereby creating the abstract (i.e., set-theoretic) component of a model object. Third, this formulation is hypothesized as having the given real-world system as its referent, thereby becoming a full-fledged model object. Fourth, this model object is taken together with the general theory so as to constitute a testable theoretical model of the system in concern.

Now, as discussed in the foregoing, when adopting a general theory, there is always implicit an underlying mode of reasoning. This mode of reasoning is usually not made explicit, since it is almost always the classical bimodal mode with which all mathematicians are assumed to be familiar (cf. §10.2). If the logic is made explicit, however, then one winds up having the general theory being represented by a formal logical system, complete with axioms and rules of formal derivation in the manner described in §11.3. Accordingly, the model object becomes represented by special symbols and axioms within this formal system, and the theoretic model itself takes the form of a semantic model of this system, which, in the manner of §11.2, has the real-world system as a factual interpretation.

This state of affairs is summarized in Figure 11-5. Here is illustrated the observer's perceptions being developed into a model object by the process of "formulated observation" (cf. §10.1), and the observer's mode of reasoning (for now, assumed to be bimodal) first being used internally as an organizing principle for thinking about his perceptions, and second being developed into an axiomatized formal logic through the process of "introspection articulation" (also cf. §10.1). Thus formalized, the logic then becomes an external organizing tool and a mechanism for drawing inferences.

It is furthermore indicated in Figure 11-5 that (i) the same formal logical system might have different semantic models whose referents are the same real-world system (although typically representing different aspects or parts of that system), and (ii) that the same semantic model might simultaneously serve as a representation of more than one real-world system (i.e., be involved in different semantic assumptions). Each of these situations may be exemplified as follows.

Consider the latter case first. Assume that our formal system is Bunge's axiom system as described in §11.2, i.e., we take the five axiom schemes as being formalized in the explicit context of an axiomatized logic, together with all the necessary axioms for multiplication, subtraction, equality, and so on. Now consider Bunge's two factual interpretations: the physical system consisting of a dc circuit, and the social system consisting of a collection of countries. As was shown in §11.2, each of these concrete systems may be modeled with semantic models of our formal system. Here we may make the further observation that, if it is assumed that these two semantic models are such that their universes of discourse have exactly the same number of elements,

A/I = Articulated
      Introspection

F/O = Formulated
      Observation

S/I = Semantic Interpretation

F/I = Factual Interpretation
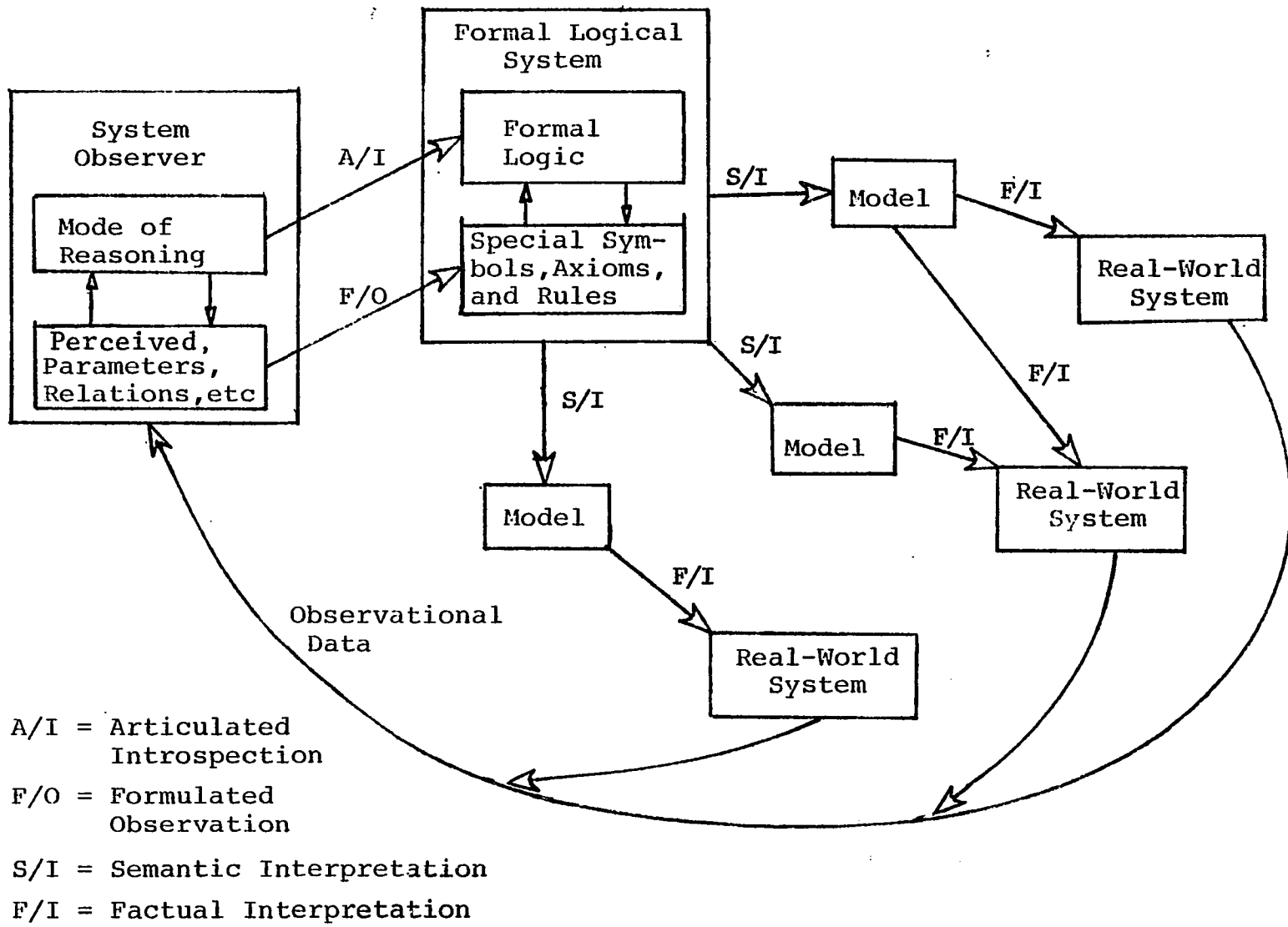
Figure 11-5.   The use of formal systems.

then these models are for all practical purposes identical. Thus we here have one semantic model serving simultaneously as a theoretical model of two distinct real-world systems.

To exemplify the situation of two distinct theoretical models of the same real-world system, we can make use of a tongue-in-cheek example provided by Bunge (p. 112). In this case, the real-world system consists of all persons currently employed in academe. Then the symbol S is assigned as its semantic meaning a universe of discourse consisting of one abstract representative of each professional academic; F is assigned as its meaning the function which gives each person's number of publications; and H is assigned a real-number measure of "natural hatred." Then G respresents a function which gives the amount of "professional jealousy" between the individual academics.

To develop a second semantic model, of the same formal system, and having the same body of academics as its referent, consider the collection of all professional societies within academe. Let S represent the collection of subsets of the previous universe of discourse, each of which subsets has its members being the representatives of all the members of a particular professional society; let F be assigned the function which gives the number of elements in each such subset, thereby having as its real-world referent a membership count of each professional society; and let H represent a measure of the average pairwise overlap in interests among the societies. Then G represents a function which gives a pairwise measure of the total amount of fruitful collaboration that can be expected to occur among academics which simultaneously belong to both societies in that pair. As a semantic interpretation of the formal system, this is obviously distinct from, and not isomorphic

with the one which models "professional jealousy." Nevertheless it qualifies as a theoretical model of the same real world system.

Now, while we have thus far assumed that the underlying logic of our formal system is bimodal, Figure 11-5 makes it obvious that alternative modes of reasoning can be chosen. In this case, the arrow labeled A/I represents the selecting of an alternative mode of reasoning, in the form of an explicitly defined formal logic. Here note furthermore, that this alternative logic need not be consciously present within the thinking processes of the observer. Rather, it is here implicit that the alternative logic is being implemented merely as a tool for model construction, while the actual mode of reasoning of the system observer may remain bimodal, but now in the slightly more sophisticated sense of being operative at the metalevel, for thinking about the nonstandard logic (cf. §10.2). By the same token, this also includes the possibility that part of the metalanguage even be made formally explicit, as in the theory of approximate reasoning studied in Part I. In any case, the explicitly defined logical system thus becomes available to the observer for ensuring that he obeys the laws of his chosen mode of reasoning.

Thus we have at least the possibility of using the theory of formal systems for developing real-world system models based on non-classical modes of reasoning. In turn, this raises a question regarding their ultimate usefulness for producing testable models of real-world systems. This question is taken up below.

§11.5. Empirical Testability Of Nonstandard Models

The two criteria of testability, as discussed in §11.1, are the Popperian critereon of refutability and Bunge's critereon of confirmability. It has been shown that only theoretical models satisfy both criteria. Moreover, it was seen that the relation of a model to its factual referent is at best one of approximation, so that Bunge's semantic assumption M≅R is always at best only approximately true.

As such, this analysis opens a pathway for the application of the logic of approximate reasoning, which has as its main objective the formulation of a concept of approximate truth. Indeed, an illustration of this idea is given in Chapter 6, where the "certainty" of predictions is linked directly to a measure of the "imprecision" in the model.

Yet at the same time, it must be asked to what extent such models are testable, especially insofar as this concerns refutability. Are we to allow here that models be acceptable even if they are "not very true"? Surely, merely down-grading the certainty of the predictions is no substitute for discarding a poor model and striving for a model which is "more true."

And the question of refutability becomes even more exacerbating when we consider applications of Kleene-Varela logics to formulation of empirically untestable properties, as in §9.4.1. Here, by our very interpretation of the logic, we appear to automatically rule out the possibility of Popperian refutation.

These questions become tractable, however, if we look more carefully at refutability in light of an earlier discussion. First, the Popperian notion of refutability only apparently rests on an assumption of semantic bimodality. This is true because, in semantically complete

classical systems, a proposition is derivable from the axioms if and only if it is semantically true. Thus, saying that a prediction is true is here equivalent to saying that the proposition which expresses that prediction is formally derivable from the axioms. However, on closer scrutiny, we can see that it is actually the condition of "derivability" and not "semantic truth" which is of primary significance. For it is only insofar as the proposition is a valid logical consequence of the axioms (describing the model in question) that that proposition may be regarded as a prediction. Propositions which do not follow from the model are of no concern, regardless of their truth value.

Second, even though bimodality is not essential for refutability, it is clearly necessary at some level of analysis. For the criteria of testability itself is meant to serve for deciding whether a model should or should not be accepted. This problem is ultimately resolved by the discussion of §10.2. There it was argued that, even when dealing with nonstandard logics, the mathematician's intuitive reasoning about those logics is always bimodal at some, perhaps higher than the formal, level. In terms of refutability, this means that one need only seek out this particular level as the one at which the criteria of testability are to be applied. Furthermore, since in all formal systems (as defined in this dissertation) a proposition is either derivable or nonderivable from the system's axioms, it turns out that the appropriate level is easily isolated. This may be illustrated by some examples.

The theory of approximate reasoning allows for the formulation of such propositions as

"JOHN is young,"

wherein the truth value of the proposition is determined as the degree

of compatibility of JOHN with the fuzzy set of ages of young persons.
Suppose that this truth value is 0.5, indicating a degree of equivocation
between "young" and "not young." Certainly one would not discard this
proposition as a correct representation of reality simply because it is
not even approximately true. This would be to miss the point of approxi-
mate reasoning. In effect, we here have a situation in which the truth
value itself has become a part of the model, and having its own factual
interpretation, i.e., as a degree of compatibility. The concept of
refutability applies, therefore, not to the proposition as stated, but
to the higher-level, and somewhat more complex, proposition

<p style="text-align:center">"'JOHN is Young' is 0.5 true."</p>

Such a proposition is in fact refutable in the same sense as classical
models, since its correctness as a statement about the factual world
depends only on a particular measurement, i.e., the measurement of
compatibility, and it is certainly reasonable to suppose that appropri-
ate compatibility measures can be crafted for whatever objects are in
concern.

A consideration of refutability for Kleene-Varela systems leads to
a similar conclusion. Consider again Wienberg's "perfect system"
(§9.2.4.1) and suppose that the proposition P in the formula

$$P \cong \neg P$$

asserts the property of "perfectness." Suppose further the same seman-
tics as before, with truth values taken as representing "verifiable,"
"refutable," and "neither verifiable nor refutable." Then once again,
Popperian refutability does not apply to the proposition as stated,
since it is again implicit that the truth values themselves are a part
of the model. What one must consider is the proposition

"P is neither verifiable nor refutable,"

which can at least in principle be either a correct or incorrect statement about the factual world.

Thus we conclude that models based on nonstandard modes of reasoning can indeed provide usable perspectives on the observable world. It is simply a matter of locating the appropriate semantic level at which the criteria of testability are to be applied. Moreover, as the above examples suggest, this appears to be exactly that level at which the truth values themselves become a part of the model.

Underlying all of this is of course the assumption that the logics employed be at least coherent logics, if they are inconsistent. Although this has nothing to do with the question of refutability, it is an important requirement in order that the formal system which uses that logic have meaningful semantic interpretations. Incoherent system will in general have no semantic models in the sense defined in this dissertation. Given a coherent general theory, however, one always has essentially the same situation as in traditional science, where consistency of the general theory is ensured by the nature of its mathematical formulation. Logical coherence will be the minimum condition under which rational viabilty is ensured.

CHAPTER 12

CONCLUSION

Recent years have evinced a small but rapidly growing interest in using nonclassical logics for the analysis of complex systems. This dissertation has succeeded in bearing this idea a few more steps out of the realm of mere conjecture and into the realm of pragmatic actuality.

On the whole, the farest reaching result is the demonstration that nonclassical logics can indeed be used as a basis for modeling. Since this settles the relevant issue of empirical testability, it shows that in principle any logically "coherent" mode of reasoning qualifies for the formation of usable perspectives on the world. Thus a solid methodological foundation is established for implementing an immense variety of multivalent logics, modal logics, temporal logics, and so on. It is simply a matter of furthering the development of such logics to a sufficient linguistic richness that they can express the system properties in concern.

In this sense, the Kleene-Varela logic discussed in Part II shows promise for modeling paradoxical, and hence classically intractable, system properties. That many systems do exhibit logically anomalous behavior is nowdays patent, and the prospect that certain systems bear paradoxical attributes per se has recently surfaced in the literature. An internally coherent formal context is therefore needed in order that such behaviors and attributes can be discussed in a meaningful way. In this case the explicitly defined formal logic serves as a guide for

correct reasoning where ordinary mathematical intuition no longer applies.

A similar purpose is served by the semantically complete axiomatization of approximate reasoning, developed in Part I. In light of the rather extensive literature on this topic, and the plethora of known applications, this particular result is likely to be of widespread theoretical significance. The present axiomatization lays down in concise terms most of the principles of reasoning which any application of fuzzy logic must obey.

Hand in hand with its providing answers to a small collection of questions, however, this work also points to several other issues which have yet to be explored. A major theoretical problem left untouched by Part I is the extension of the concept of linguistic theory to accommodate fuzzy-logical quantifiers, i.e., generalizations of the classical "for all" and "for some" to expressions like "for many" and "for few." By analogy with the corresponding problem for classical systems, this presents a challenge of considerably greater difficulty than the semantic completeness results established here. In addition is the issue of linguistic approximation which, although requiring a more sophisticated formal language, could conceivably also be provided with a concise formalism.

It is also evident that the full range of potential applications of approximate reasoning has hardly begun to be explored. The example developed in Chapter 6 represents only one of many directions that could be undertaken, i.e., the use of dynamic models for forecasting. Overall, the cutting edge of fuzzy set theory is its power as a model of natural languages, together with its ultimate implementability on the modern

computer. Based on this conception it is likely that fuzzy sets will play an important role in data base management, interactive question-answering systems, computer aided instruction, business decision analysis, mechanized control systems, and pattern recognition, as well as in higher level systems applications, such as semantic modeling and complexity approximation.

The discussion in Chapter 6 concerning the use of fuzzy logic for establishing a direct link between model imprecision and certainty of predictions illustrates a more general class of problems for which fuzzy sets may also become a vital tool. In particular is the proposed extension to full implementation of fuzzy logic for incorporating a concept of linguistic "Certainty." And this may in part involve, or even require, an explicit fomalization of Zadeh's generalized modus ponens for fuzzy sets of type 2 (§5.5.3). Further, the present formalization of approximate reasoning is implicitly an extension of the Kleene-Varela logic to a logic having a substantially richer linguistic structure, which means that the linguistic theories developed in Part I may well provide exactly what is needed for expressing paradoxical notions which go beyond the capacity of the languages studied in Part II.

On a higher level is a more broadscoped project of the kind mentioned in the preface. Given a certain body of known laws and principles regarding an aspect of reality, if one formalizes those into a semantically complete formal system, that system will normally have a large collection of semantic models other than the ones which initially served as the intuitive guide for developing the formalization. Such models will then represent alternative "worlds" in which the given principles and laws hold true, and therefore may conceivably stimulate new avenues of emperical investigations.

Last, while there nowdays exist a wide variety of nonclassical logics which scientists might choose to employ in model construction, it is also clear that emperical research may inspire the development of totally new modes of reasoning. On this count, the present work may be deemed a success if it has only planted the idea that such an approach is both feasible and in fact promising of a deeper understanding of the world.

REFERENCES

The notation in brackets following each entry indicate subsections where that reference is cited in the text.

Bateson, G., Jackson, D. D., Haley, J., and Weakland, J. 1956. Toward a theory of schizophrenia. Behavioral Science, 1:251-267. [1.2]

Bellman, R. 1977. Large systems. In H. A. Linstone and W. H. C. Simmonds (1977), pp. 100-103. [1.1]

Bellman, R. and Giertz, M. 1973. On the analytic formalism of the theory of fuzzy sets. Information Sciences, 5:149-156. [2.2.8]

Bellman, R. and Zadeh, L. A. 1977. Local and fuzzy logics. In J. M. Dunn and G. Epstein (1977). [1.1, 2.2.9,5.5.2]

Bunge, M. 1973. Method, model and matter, D. Reidel, Boston, Mass.. [1.1.3, 11.1-11.5]

Cavallo, R. E. (ed.) 1979. Systems research movement: characteristics, accomplishments, and current developments, General Systems Bulletin, Special Issue--Summer 1979, Vol. IX, No. 3. [1]

Cull, P. and Frank, W. 1974. Flaws of form. Int. J. General Systems, 5:201-211. [1.2]

Dubois, D. and Prade, H. 1979a. Fuzzy sets and systems: theory and applications, Academic Press, New York. [1.1]

Dubois, D. and Prade, H. 1979b. Operations in a fuzzy-valued logic. Information and Control, 43:229-240. [5.5, 5.5.3]

Dunn, J. M. and Epstein, G. (eds.) 1977. Modern uses of multiple-valued logic, D. Reidel, Dordrect, Holland. [1.3, References]

Frege, G. 1879. Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. In J. van Hiejenoort (1967), pp. 1-82. [2.1.5.1]

Gaines, B. R. 1976. Foundations of fuzzy reasoning. Int. J. Man-Machine Studies, 8:623-668. [1.1, 1.3, 2, 2.2, 2.2.8, 4.5, 4.5.2, 10.2]

Gaines, B. R. and Kohout, L. J. 1977. The fuzzy decade: a bibliography of fuzzy systems and closely related topics. Int. J. Man-Machine Studies, 9:1-68. [1.1, 1.3]

Gale, R. M. 1968. The language of time, Routledge and Kegan Paul, London. [1, 9.2.4.2]

Giles, R. 1976. Lukasiewicz logic and fuzzy set theory. Int. J. Man-Machine Studies, 8:313-327. [1.1]

Gödel, K. 1931. On formally undecidable propositions of Principia Mathematica and related systems, I. In J. van Heijenoort (1967), pp. 596-616. [10.1]

Goguen, J. A. and Varela, F. J. 1979. Systems and distinctions, duality and complementarity. Int. J. General Systems, 5:31-43. [1]

Gupta, M., Ragade, R., and Yager, R. (eds.) 1979. Advances in fuzzy sets theory and applications, North-Holland, New York. [1.1]

Gupta, M. M., Saridis, G. N., and Gaines, B. R. (eds.) 1977. Fuzzy automata and decision processes, North-Holland, New York. [1.1]

Haack, S. 1979. Do we need fuzzy logic? Int. J. Man-Machine Studies, 11:437-445. [5.3, 5.3.1]

Hartman, R. S. 1967. The structure of value: foundations of scientific axiology, Southern Illinois University Press, Carbondale, Ill. [9.2.4.3]

Heyting, A. 1956. Intuitionism: an introduction, North-Holland, Amsterdam. [10.1]

Heyting, A. 1964. The Intuitionist foundations of mathematics. In P. Benaceraff and H. Putnam (eds.), Philosophy of mathematics, Prentice-Hall, Englewood Cliffs, New Jersey, pp. 42-49. [10.2]

Hilbert, D. and Ackermann, W. 1928. Grundzüge der theoretischen Logik, Springer, Berlin. In English as R. E. Luce (ed.), L. M. Hammond, et. al. (trans.), Principles of Mathematical Logic, Chelsea, New York, 1950. [3.3.1, 7, 7.1]

Howe, R. H. and von Foerster, H. 1975. Introductory comments to Francisco Varela's calculus for self-reference. Int. J. General Systems, 2:1-3. [1.2]

Kandel, A. and Lee, S. C. 1979. Fuzzy-switching and automata; theory and applications, Crane-Russak, New York. [1.1]

Kandel, A. and Yager, R. R. 1979. A 1979 bibliography on fuzzy sets, their applications, and related topics, Florida State University, Tallahasee, Fl. 32304. [1.1]

Kaufman, A. 1975. Theory of fuzzy subsets, Academic Press, New York. [1.1]

Kleene, S. C. 1938. On notation for ordinal numbers. J. Symbolic Logic, 3:150-155. [1.2, 8, 8.2]

Kleene, S. C. 1952. Introduction to Metamathematics, Van Nostrand, New York. [1.2, 8, 8.2, 9.1.1, 9.1.2, 10.1]

Kleene, S. C. 1967. Mathematical logic, Wiley and Sons, New York. [10.2]

Kohout, L. J. and Pinkava, V. 1980. The algebraic structure of the Spencer-Brown and the Varela calculi. Int. J. General Systems, 6:155-172. [1.2]

Lendaris, G. G. 1980. Structural modeling: a tutorial guide. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-10, No. 12:807-840. [1, 6.1]

Linstone, H. A. 1973. On discounting the future. Technological Forecasting and Social Change, 4:335-338. [1.1]

Linstone, H. A. 1977. Confessions of a forecaster. In H. A. Linstone and W.H.C. Simmonds (1977), pp. 3-12. [1.1]

Linstone, H. A., Hays, J., Lendaris, G. G. Rogers, S. D., Wakeland, W., and Williams, M. 1978. The use of structural modeling for technology assessment (2 Volumes), National Science Foundation Grant No. ERS76-21040, Institute for Futures Research, Portland State University, Portland, Ore. 97207. [6.2, 6.4]

Linstone, H. A., Lendaris, G. G., Rogers, S. D., Wakeland, W., and Williams, M. 1979. The use of structural modeling for technology assessment. Technological Forecasting and Social Change, 14:291-327. [6.1]

Linstone, H. A. and Simmonds W.H.C. (eds.) 1977. Futures research new directions, Addison-Wesley, Reading, Mass. [1.1, References]

Lukasiewicz, J. 1930. Philosophical remarks on many-valued systems of propositional logic. In S. McCall (ed.), Polish Logic 1920-1939, Clarendon Press, Oxford, 1967. [10.2]

Martin, R. L. 1970. The paradox of the liar, Yale University Press, New Haven, Conn. [9.2.3]

Maruyama, M. 1963. The second cybernetics: deviation-amplifying mutual causal processes. American Scientist, 51:164-179. [1.2]

Maruyama, M. 1974. Paradigmatology and its application to cross-disciplinary, cross-professional and cross-cultural communication, Part I. Cybernetica, 27:136-156; Part II, 27:237-281. [10.2]

Maturana, H. B. 1970. Biology of cognition, B.C.L. Report No. 9.0, Biological Computer Laboratory, University of Illinois Press, Urbana, Ill. [1.2]

McLean, M., Shepherd, P., and Curnow, R. 1976. Techniques for analysis of system structure, SPRU Occasional Paper Series No. 1, Science Policy Research Unit, University of Sussex, U.K. [1.1, 6, 6.2]

Mendelson, E. 1964. Introduction to mathematical logic, Van Nostrand, New York. [2, 2.1.5.4]

Mizumoto, M. and Tanaka, K. 1976. Some properties of fuzzy sets of type 2. Information and Control, 31:312-340. [5.3.1, 5.3.2]

Moisil, G. 1975. Lectures on fuzzy logic, Scientific and Encyclopedic Editions, Bucharest, Rumania. [1.1]

Negoita, C. V. 1980. Fuzzy systems, Abacus Press, Ormond, Beach, Fl. [1.1]

Orchard, R. A. 1975. On the laws of form. Int. J. General Systems, 2:99-106. [1.2, 7, 7.5]

Post, E. L. 1921. Introduction to a general theory of propositions. In J. van Heijenoort (1967), pp. 264-283. [2.1.5.4]

Rapoport, A. 1966. Two-person game theory, University of Michigan Press, Ann Arbor. [10.4.2]

Rasiowa, H. 1974. An algebraic approach to nonclassical logics, North-Holland, Amsterdam. [3.3.1]

Rasiowa, H. and Sikorski, R. 1963. The mathematics of metamathematics, Polish Scientific Publishers, Warsaw. [4, 4.1, 4.1.2]

Reichenbach, H. 1944. Philosophical foundations of quantum mechanics, University of California Press, Berkeley. [1.3, 10.1]

Rescher, N. 1969. Many-valued logic, McGraw Hill, New York. [1.3, 7.1]

Rescher, N. 1973. The coherence theory of truth, Oxford University Press, London. [1.2, 2.1.5.2]

Rescher, N. and Brandom, R. 1980. The logic of inconsistency, Basil Blackwell, Oxford. [1.2, 2.1.5.2]

Robinson, A. 1963. Introduction to model theory and to the metamathematics of algebra, North-Holland, Amsterdam. [2]

Rogers, H. Jr. 1967. Theory of recursive functions and effective computability, McGraw Hill, New York. [9.1.1]

Russell, B. 1902. Letter to Frege. In J. van Heijenoort (1967), pp. 124-125. [1.2, 9.2.3]

Russell, B. 1908. Mathematical logic as based on the theory of types. In J. van Heijenoort (1967), pp. 150-182. [1.2, 9.2.3]

Sahal, D. 1976. Homeorhetic regulations and structural stability. Cybernetica, 19:305-315. [1, 1.2]

Sahal, D. 1977. The conception of futures in a systems framework. In H. A. Linstone and W.H.C. Simmonds (1977), pp. 158-163. [1.2]

Sahal, D. 1978. A formulation of the Pareto distribution. Environment and planning A, 10:1363-1376. [1]

Sartre, J. P. 1956. Being and nothingness, H. E. Barnes (trans.), Simon and Schuster, New York. [1.2]

Schwartz, D. G. 1973. Free-variable theories, Masters Thesis, Simon Fraser University, Burnaby, B.C., Canada. [4]

Schwartz, D. G. 1977. Toward a global spirit. In E. Laszlo and J. Bierman (eds.), Goals in a global community, Pergamon, pp. 315-322. [Preface]

Schwartz, D. G. 1981a. Isomorphisms of Spencer-Brown's laws of form and Varela's calculus for self-reference. Int. J. General Systems, 6:239-255. [Preface]

Schwartz, D. G. 1981b. Forecasting in a fuzzy environment. In General systems research and design: precursers and futures, Proceedings of the twenty-fifth annual North American meeting of the Society of General Systems Research, Toronto, Canada, January 6-9, pp. 20-26. [Preface]

Shoenfield, J. R. 1967. Mathematical logic, Addison-Weseley, Reading, Mass. [2, 3.4.1, 4.6.3]

Skolem, T. 1923. The foundations of elementary arithmetic established by means of the recursive mode of thought, without the use of apparent variables ranging over infinite domains. In J. van Hiejenoort (1967), pp. 302-333. [10.1]

Smith, H. 1976. Forgotten truth: the primordial tradition, Harper and Row, New York. [1.2]

Smullyan, R. M. 1961. Theory of formal systems, Princeton University Press, Princeton, N.J. [2]

Spencer-Brown, G. 1969. Laws of form, George, Allen, and Unwin, London. [1, 1.2, 7, 7.3-7.5, 8.1]

Suppes, P. 1961. A comparison of the meaning and uses of models in mathematics and the empirical sciences. In H. Freudenthal (ed.), The concept and role of the model in mathematics and natural and social sciences, D. Reidel, Dordrect, Holland. [1, 1.3, 11.2]

Swami, P. (trans.) 1977. The Bhagavad Gita, Random House, New York. Earlier edition, Faber and Faber, London, 1935. [Opening Quote]

Tarski, A. 1949. The semantic conception of truth and the foundations of semantics. In H. Feigel and W. Sellars (eds.), Readings in philosophical analysis, Appleton, New York. [10.1]

Thom, R. 1975. Structural stability and morphogenesis, D. Fowler (trans.), Benjamin, Reading, Mass. [1]

van Hiejenoort, J. (ed.) 1967. From Frege to Gödel a sourcebook in mathematical logic, 1879-1931, Harvard University Press, Cambridge, Mass. [References]

Varela, F. J. 1975. A calculus for self-reference. Int. J. General Systems, 2:5-24. [1, 1.2, 8, 8.1, 8.5, 9, 9.2.3]

Varela, F. J. 1979. The extended calculus of indications interpreted as a three-valued logic. Notre Dame J. Formal Logic, XX:141-146. [1.2]

Varela, F. J. and Goguen, J. A. 1978. Arithmetic of Closure. J. Cybernetics, 8:291-324. [1]

von Foerster, H. 1973. On constructing a reality. In W. Preiser (ed.), Environmental design research, Dowden, Hutchinson, and Ross, Stordsburg, pp. 33-46. [1.2]

Weinberg, G. M. 1975. An introduction to general systems thinking, Wiley, New York. [1, 1.2, 9.2.4.1]

Whitehead, A. N. and Russell, B. 1910-1913. Principia mathematica, Cambridge University Press, Cambridge; Vol 1, 1910, Vol. 2, 1919, Vol. 3, 1913; 2nd ed. Vol 1, 1925, Vol. 2, 1927, Vol. 3, 1927. [10.1]

Yager, R. R. 1979. A procedure for ordering fuzzy subsets of the unit interval. Technical Report RRY 79-12, Iona College School of Business, New Rochelle, N.Y. 10801. [2.2.7]

Zadeh, L. A. 1965. Fuzzy sets. Information and Control, 8:338-353. [1.1]

Zadeh, L. A. 1972. A fuzzy-set-theoretic interpretation of linguistic hedges. J. Cybernetics, 2:4-34. [1.1, 1.2.4]

Zadeh, L. A. 1975a. Fuzzy Logic and approximate reasoning. Synthese, 30:407-428. [1.1]

Zadeh, L. A. 1975b. The concept of a linguistic variable and its application to approximate reasoning. Information Sciences, Part I, 8:199-249; Part II, 8:301-357; Part III, 9:43-80. [1,1, 2, 2.2, 2.2.6, 2.2.9, 2.2.10, 5.3.1]

Zadeh, L. A. 1978a. PRUF--a meaning representation language for natural languages. Int. J. Man-Machine Studies, 10:395-460. [1.1, 2, 2.2.5, 2.2.10, 5.4, 10.3]

Zadeh, L. A. 1978b. Fuzzy sets as a basis for a theory of possibility. Int. J. Fuzzy Sets and Systems, 1:3-28. [1.1, 2.2.10, 5.4, 10.3]

Zadeh, L. A. 1979. Liar's paradox and truth-quantification principle. Memorandum M79/34, Electronics Research Laboratory, University of California, Berkeley. [1, 9.2.4]

Zadeh, L. A. 1980. Fuzzy sets versus probability. Proceedings of the IEEE, 68,3:421. [2.2.10]

Zadeh, L. A., Fu, K. S., Tanaka, K., and Shimura, M. (eds.) 1975. Fuzzy sets and their applications to cognitive and decision processes, Academic Press, New York. [1.1]

Zwick, M. 1978. Quantum measurement and Gödel's proof. Speculations in Science and Technology, 1:135-145. [1.2]

Zwick, M., Schwartz, D. G., and Lendaris, G. G. 1978. Fuzziness and catastrophe, IEEE Proceedings of the International Conference on Cybernetics and Society, Tokyo/Kyoto, Japan. [1]

INDEX

The numbers following each entry refer to principal sections of the text where the indexed term is defined or discussed.