NEGOTIATION AS PROBLEM SOLVING

Michael J. Prietula and Laurie R. Weingart

ABSTRACT

A view of negotiation is presented from the perspective of problem solving. In recent years, research in problem solving has shifted from “phenomena-driven” approaches to “architecture-driven” approaches seeking the fundamental invariants of intelligent deliberation across situations and then determining how responses to differing situations are composed by the invariants. The model of problem solving we present is based on Newell’s unified theory of cognition which specifies an underlying symbolic cognitive architecture and the principles by which it operates. The cognitive mechanisms underlying this formulation are seen as both universal and insensitive to the task at hand, driven predominantly by opportunistic adaptations dependent on the availability of knowledge to support search in problem spaces in service of goals. It is argued that negotiation is a knowledge-driven task and, consequently, is explainable in terms of the mechanisms proposed. We describe two studies which incorporate this perspective. We conclude with a discussion on computational models of negotiation theory.

Like it or not, you are a negotiator. Negotiation is a fact of life.
— Getting to Yes (Fisher, Ury & Patton, 1991)
Negotiation is a task in which two or more agents communicate, directly or indirectly, with the general objective of making decisions regarding resource allocation or activity. What makes negotiation interesting, or perhaps what makes negotiation interesting to researchers, is that people are often not very good at it. That is, when two people are presented a simple negotiation task, they more often than not come to an agreement that is less than the optimal one they both could have reached. The research question, then, is straightforward—why were suboptimal agreements reached? One route to an answer lies in explicating the reasoning methods of the negotiators themselves, that is, in explicating negotiator cognition.

Understanding negotiator cognition has been shown to provide insights into negotiator behavior (Bazerman & Carroll, 1987; Carroll & Payne, 1991). For example, in Carroll and Payne’s (1991) review, they incorporated the “schema” construct as a unifying theme around which several negotiation phenomena were discussed (e.g., why people often inappropriately assume that it is a zero-sum game). In this chapter we take a different perspective on negotiator cognition. Our unifying theme is not a memory or attentional construct, but a theory. Specifically, we start with a special type of cognitive theory, called a unified theory of cognition, and then derive from it explanations (and hence hypotheses) of phenomena documented in two-party, multiple-issue negotiation research.

A unified theory of cognition is one that posits “a single system of mechanisms that operate together to produce the full range of human cognition” (Newell, 1990, p. 1). As we shall see, the particular unified theory we have selected attributes variance in behavior to differences in (or in the use of) knowledge; therefore, differences in negotiation behavior presume differences in (or in the use of) knowledge. As we shall also see, this particular unified theory makes a distinct statement on the roles that the environment and limits on human information processing play on the acquisition and use of knowledge.

Why have we selected this approach as our starting point? As an adage in science goes, “Hypotheses must come from the theory and not the theorist.” Thus, by starting from such a theory, we are able to engage broad theoretical constructs and principles that serve as constraints on what we ask about cognition in negotiation (i.e., the hypotheses of interest), how we ask it (i.e., the methods of our studies), and how we might interpret the results. Such theories are quite useful things, as they succinctly state “what matters.”

This chapter is organized as follows. We first present the underlying theory of cognition which serves as the foundation for our work. The nature of the theory is both encompassing and unique. Our goal is only to present the fundamentals, the flavor if you will, of such a theory. We then describe two studies which incorporate components of the theory. These studies seek to answer a simple question in negotiation. Why do naive negotiators fail to
achieve Pareto optimality? We conclude with comments on a research agenda based on this approach.

**BEYOND INFORMATION PROCESSING THEORY**

In 1956 a group of psychologists, social psychologists, computer scientists (though the term was not yet defined), linguists, and engineers gathered to define a new way at looking at the human mind. As Simon (1980) has noted, 1956 was a very good year for the birth of information processing theory. George Miller published his seminal work on short-term memory (Miller, 1956); Noam Chomsky’s analysis of transformation grammars appeared (Chomsky, 1956); Bruner, Goodnow and Austin’s *A Study of Thinking* was in print (Bruner, Goodnow & Austin, 1956), and Newell and Simon reported on their program, the Logic Theorist (Newell & Simon, 1956). The Logic Theorist program is notable for three reasons. First, it was written as a symbolic, not a numeric, program. It manipulated *lists of symbols* rather than numbers, to reflect the symbolic associative nature of human reasoning. It was a theory of reasoning expressed as a program. Second, it introduced the concept of recursion to programming: functions being defined in terms of themselves. Third, it addressed a very interesting set of problems: it solved proofs in Whitehead and Russell’s *Principia Mathematica*. Gradually, the information processing perspective became the dominant paradigm in cognitive psychology (Lachman, Lachman & Butterfield, 1979).

Seventeen years later, Allen Newell wrote a critique of a symposium on information processing/cognitive psychology held at Carnegie Meilin University. That essay was entitled, “You Can’t Play 20 Questions with Nature and Win” (Newell, 1973). Although Newell was impressed with the efforts reported at the symposium, it was his concerns that yielded the central (or at least, lasting) theme of the paper. Newell was concerned at that time with cognitive psychology’s preoccupation with phenomena. Every new phenomena generated by a particular manipulation in a laboratory, would be promptly named and then would spur a flurry of experiments to further investigate it, requiring more manipulations, and producing more phenomena along the way, and so forth. To emphasize this point, he generated a list of 59 phenomena under investigation at that time. Newell’s conclusions reflected the lack of convergence:

But it didn’t seem to me to add up to much. What I wanted was for these excellent pieces of the experimental mosaic to add up to the psychology that we all wished to foresee. They didn’t, not because of any lack of excellence locally, but because most of them seemed part of a pattern of psychological activity that didn’t seem able to cumulate. (Newell, 1973, p. 293)
Scientific progress in general (and in psychology, in particular) is accomplished not by endlessly documenting phenomena. Rather, major progress occurs when the underlying mechanisms which produce the phenomena are described. Or, perhaps more accurately, progress occurs in cognitive psychology when a distinction can be made between the invariant phenomena of the mechanism (i.e., the mind) and how that mechanism yields classes of adaptive, variant phenomena (i.e., behaviors) as it generates and applies knowledge to achieve goals within tasks. One of the experimental paradigms Newell proposed as a solution was to build complete processing models rather than partial ones evidenced in phenomena-driven studies, where much is left underspecified and unconstrained. Newell argued that he “on balance, prefer(s) to start with a grossly imperfect, but complete model, hoping to improve it eventually, rather than start with an abstract but experimentally verified characterization, hoping to specify it further eventually” (1972, p. 375).

Consequently, a sub-stream of research in this mold has emerged over the years based, in part, on this tenet. Current research into this approach has focused on articulating these complete processing models as cognitive architectures (VanLehn, 1991). Different architectures reflect different assumptions about the underlying invariant mechanisms of representation and deliberation, and what can be accomplished with them. What is interesting about this approach is a group of cognitive architectures that strive to be unified theories of cognition.

By “unified” we mean psychologically plausible mechanisms that can account for an entire range of intelligent behavior (perceptual, linguistic, reasoning, motor behavior, memory, learning, etc.) and for an entire range of tasks. This is a rather strong view and eliminates perspectives and theories that are less encompassing and/or less architecturally specific (e.g., Chomsky, 1980; Holland, Holyoak, Nisbett & Thagard, 1986; Minsky, 1986; Shank, 1985; Sternberg, 1985).

By “cognitive architecture” we mean the structure that provides the framework within which cognitive processing takes place. The concept of architecture is based on an important perspective developed to reason about, design, and implement computer systems on multiple levels of abstraction (Bell & Newell, 1971). As one can define a computer on several levels of abstraction (hardware and software), one can also define a mind on several levels of abstraction. Similarly, only a subset of levels (some argue only one) are relevant for the explanation (or production) of certain behaviors of interest. Our interest is specifically on cognition. Accordingly, Pylyshyn (1991) comments:

To put it another way, it [the architecture] is the level at which the system is representational, and where the representations correspond to the objects of thought (including percepts, memories, goals, beliefs, and so on). In other words, the semantic interpretation of these states figures in the explanation of the cognitive behavior. Notice that there may be other
levels of system organization below this, but these do not constitute cognitive architectures because their states do not represent cognitive content (p. 191).

There are (arguably) two groups promoting universal mechanisms for intelligent behaviors: connectionists and symbolicists. Connectionists (Rumelhart et al., 1986a,b; Touretzky, 1989) focus on constructing models of phenomena (e.g., memory) based on principles of neuroscience and approximating the behaviors of large numbers of neural components operating in parallel. Although aspects of micro-phenomena at the neural level may be modeled, the integration of the micro-models into a universal mechanism accounting for broad and general patterns of conscious deliberation cannot (yet) be explained (e.g., Skarda & Freeman, 1987, 1988). Regardless, full models have not been instantiated so the theory itself cannot be tested directly.

Symbolicists (e.g., Anderson, 1983; Minton et al., 1989; Mitchell, Keller & Kedar-Cabelli, 1986; Newell, 1990), influenced heavily by computer science and psychology, take the view (either explicitly or implicitly) that intelligent deliberation can be studied in a layered manner; that is, descriptions of system behaviors can be architecturally stratified on a biologically meaningful scale. Marr's (1982) classic characterization, for example, described three levels of analysis for understanding the computation of vision: (1) an abstract problem analysis which focused on decomposing the task into its primary constituents, (2) a formal procedural level in which algorithms could be specified which performed the task given correct inputs, and (3) a physical implementation level which could describe the mechanisms in terms of a specific technology. A key issue in Marr's influential proposal (that has had ramifications to all of cognition) is that a given level could be researched independent of the underlying level.

Pylyshyn (1980, 1984) and Anderson (1987) make similar hierarchical, but more general, characterizations using two levels. Essentially, on one level there are the mental algorithms (Anderson's term) referring to the "mental procedures and knowledge we possess that enable us to behave adaptively" (p. 468). These are then realized at the implementation level which describes the invariant mechanisms that "run" the mental algorithms (though this level is not as concrete as Marr's physical implementation level).

The theory of architecture on which we base our negotiation research is a similar, hierarchically-based, symbolic theory. The selection of this theory is based on its encompassing nature, as it promotes simple but universal mechanisms for representation, intelligent deliberation, and learning (Newell, 1990). The focus of the theory is that knowledge is the key to intelligent behavioral success and failure.
THE PROBLEM SOLVING ARCHITECTURE OF THE NEGOTIATOR

As a symbolic theory, computational mechanisms capable of defining and manipulating symbolic structures are fundamental. Incorporating symbols as the elemental entities upon which an intelligent representational system, such as ours, is based is not new (Whitehead, 1927). We deal with our own internal representations of what we see (or think we see), what we recall, and what we imagine. HOW such symbols may be represented is another matter (e.g., in neural circuits, in silicon). Our interest is only in THAT symbols are represented. The present approach extends and refines the notion of symbolism in deliberation, and makes an important distinction between two architectural levels—the manipulation of symbols (and symbol structures) and the formulation of knowledge in terms of symbols.

Manipulating Symbols

There are many types of “symbol manipulation systems” and we are all familiar with several, such as algebra and logic. These systems provide a basic language consisting of symbols, how they go together, and how they can be manipulated. Computers are also symbol manipulation systems, as we give the computer symbols (as commands) to manipulate other symbols (as data). The mind is also a symbol manipulation system. The representation and manipulation of symbols are physical events that occur in the brain.

There is an important general type of symbol manipulation system that can be defined (by special rules of representation and behavior) which creates, modifies, reproduces, interprets, and destroys symbols and symbol structures. This type of system is called a physical symbol system and “produce(s) through time an evolving collection of symbol structures” (Newell & Simon, 1976, p. 116). The interesting thing proposed about physical symbol systems is their potential, characterized as a hypothesis (Newell, 1980a):

*The Physical Symbol System Hypothesis: A physical symbol system has the necessary and sufficient means for general intelligent action.*

Thus, general intelligent activity (i.e., the nature of activity that we associate with human behavior) is dependent upon certain functional requirements of a symbol manipulation mechanism, but is independent of the exact specification of that mechanism (i.e., how symbols and their rules of manipulation are actually represented). Imposed upon the physical symbol system hypothesis is the notion that symbols form the basis of problem representations. The brain does not encode reality directly; rather, it operates in terms of representations of reality—symbols. Consequently, how problem
solving is fundamentally carried out can be expressed in terms of manipulating symbols. This is addressed by a hypothesis of Newell & Simon (1976):

_The Heuristic Search Hypothesis:_ The solutions to problems are represented as symbol structures. A physical symbol system exercises its intelligence in problem-solving by search—that is, by generating and progressively modifying symbol structures until it produces a solution structure.

The realization of a general model of intelligence based on these two hypotheses embodies the concepts of problem spaces, operators, and search control. A _problem space_ can be envisioned abstractly as a set of related nodes (collections of associated symbol structures) representing various attainable states with, possibly, one or more distinguished states defining the solution to the problem: the goal(s). Any given state is defined by a subset of the total collection of symbol structures “in” working memory (i.e., attentionally salient). Manipulations of those symbol structures are accomplished through _operators_ that are also symbol structures activated from memory in service of (possibly intermediate) goals.

The notion of _search_ is a process of symbol manipulation which reflects deliberation and choice among alternative operators and goals. This occurs en route to detecting a particular goal state (symbol structure) under the heuristic search hypothesis. Problem solving as search also reflects proposal and consideration of potential solution states. Furthermore, this search is accomplished by a set of procedures comprised of search control mechanisms. These search control mechanisms engage and monitor problem spaces, states, and operators—that is, they too, interpret symbol structures. The lack of search control knowledge may result in such deliberation inefficiencies as an attempted exhaustive search of the problem space, the application of inappropriate operators, the inability to consider appropriate operators when necessary, the unnecessary consideration of previously encountered states, or a misconfiguration of the space.

Symbols that exist and can be formulated into complex symbol structures representing nodes in a state space are inherent properties of the architecture. A central issue is how such functionality can contribute to the formation of deliberation on a scale which could be judged as intelligent problem solving. The answer lies in the formulation of another level of functionality—the knowledge level.

Representing Knowledge

Up to this point the reader may well have observed that little has been said of negotiation. That is an important observation. The mechanisms described thus far are universal in their role in cognitive deliberation. These same symbol
manipulation mechanisms apply whether one is learning a language, driving a car, solving algebra problems, developing skill in any task (such as negotiation), or in developing expertise in general (Newell, 1990). What differs is the nature of the particular sets of symbol structures engaged in any one context at any one time. This is called the knowledge level (Newell, 1981):

Knowledge Level: There exists a distinct...level, lying immediately above the symbol level, which is characterized by knowledge as the medium and the principle of rationality as the law of behavior.

The knowledge level permits the creation of representational systems and arbitrary functions that denote aspects of itself and the external world—the task environment. A task environment refers to the problem as presented to the problem solver as viewed by an “omniscient observer” (Newell & Simon, 1972). It is important to realize that the task environment includes the (current) knowledge of the problem solver, the goals defined by the problem solver to achieve, and the knowledge of the devices and information at his or her disposal, and other people who may be active in the task (e.g., the negotiating partner). In order to attempt to solve a problem, the problem solver must create internalized representations of the task environment—problem spaces (Newell, 1980b).

The knowledge level defines “knowledge” as we generally use the term. These representations and functions are both interpretable and mutable; however, they are also subject to law (or principled behavior). The law embodies the notion of rational service toward goals, for goal-specification and goal-attainment are the cornerstones of intelligent action (Newell, 1981):

Principle of Rationality: If an agent has knowledge that one of its actions will lead to one of its goals, then the agent will select that action.

Rationality, in this sense, is not referring to optimization; rather, it reflects the inherent preference for specific alternatives within the context of problem space search. Thus, behavior (i.e., the application of knowledge) can be interpreted in terms of the set of dominant active goals and the knowledge available to be brought to bear on the goals.

Differing goals, differing knowledge, or differing availability of knowledge may result in distinctively different, but quite rational, behaviors. In addition, the availability of knowledge or salience of certain goals, is limited by the attentional capacity of the working memory of the individual in the context of the task. Working memory offers a limited capability (at the symbol level) to temporarily retain and focus on a subset of symbols in working memory, including externally perceived (i.e., sensory) as well as internally generated (i.e., memory, inferentially originated) symbol structures (Baddeley, 1992; Miller,
1956; Naatanen, 1992). This constrained capability, when taken along with the Principle of Rationality, recapitulates Simon's (1945) classic notion of "bounded rationality." The role of bounded rationality in understanding problem solving in general (and thus negotiation problem solving in particular) is not a peripheral one. Simon (1991) underscores the point:

Bounded rationality is what cognitive psychology is all about. And the study of bounded rationality is not the study of optimization in relation to task environments. It is the study of how people acquire strategies for coping with those environments, how those strategies emerge out of problem space definitions, and how built-in physiological limits shape and constrain the acquisition of problem spaces and strategies. (p. 35)

The general notion of intelligent deliberation is therefore expressible in terms of physical symbol systems that have the capacity to approximate knowledge level systems. It is at the knowledge level where intelligence and deliberation is defined in terms of tasks. While the physical symbol system is constrained by laws of physics and biology, the knowledge level is bound to obey the Principle of Rationality. However, within the confines of that Principle, much flexibility is permitted, as specifics of behaviors are a function of the knowledge available and the goals active for a particular problem. Behavioral invariants are difficult to assess independent of the task demands, for it is the interplay between the knowledge level, the physical constraints imposed by the symbol level on the knowledge level, and the task which (collectively) describe the deliberation alternatives available to the agent. With these principles intact, we can now turn to the task of negotiation.

A NEGOTIATION TASK

The construction of specific problem spaces, representations, and operators are determined by the task presented to the problem solvers. It is the task, then, that is a critical component in explaining behaviors: the goals that are important and the nature of the knowledge that must exist to achieve them. That is, HOW individual deliberation ensues is presented by the theory of the architecture; WHAT individual deliberation occurs is handled by the theory of the task and investigated at the knowledge level. What the theory tells us is that negotiation tasks are solved like any other problem—by the same mechanisms, in the same architecture, and with the same rationality principle in effect. What varies is the knowledge that might be available to the problem solver. Knowledge defines the immediate goals of the negotiator, how to view a negotiation, how to define and interpret the other party's behaviors, what assumptions to make, what options to try, and what to conclude under ill-defined or ambiguous circumstances. Knowledge imposes meaning on a situation, whether it involves introspective deliberation or
interaction with others. Knowledge defines successes as well as failures in negotiation.

After this circuitous route, let us return to our original question regarding negotiation: Why do naive negotiators fail to achieve Pareto optimality? The answer perhaps is now obvious: they fail because they cannot bring to bear the appropriate knowledge. This answer is somewhat deceptively simple, yet quite powerful. It tells us that negotiation is goal-driven and that knowledge is brought to bear in service of those goals. It tells us that aspects of the negotiation task environment (e.g., orienting instructions, negotiation materials, properties of the particular negotiation task, the other negotiator) are all potentially important in defining goals and engaging knowledge. It tells us where to look further.

What is Failure in Negotiation?

Negotiation can encompass a wide variety of situations. For our purposes, we focus on a typical type of laboratory negotiation problem: a two-party, four-issue case with integrative potential. This case involves two naive negotiators, a florist and a baker, who are trying to agree on four issues of concern before they commit to relocating their businesses to a new location. The four issues of concern are: hiring of clerks, maintenance of their area, joint advertising, and temperature of the area. The value of each alternative is provided by a “payoff schedule” which reflects the utility of each level of each issue for the particular negotiator (see Table 1).

Two of the issues (temperature, maintenance) are distributive, where one party’s gain is the other party’s loss. Two other issues (clerks, advertising), although also in inverse value to each negotiator, are worth differing amounts to each negotiator, thus allowing tradeoffs between issues. If the two negotiators simply “split the difference” (i.e., take the midpoints on all of the issues), their score would be a joint profit of 440 points. However, if they each conceded on one of two specific issues (clerks, advertising), their joint profit would be the best (i.e., Pareto optimal) solution available of 560 points.

Figure 1 illustrates the solution set for this particular negotiation task. Each point represents a potential agreement between the baker and the florist. As the instructions of the task require that an “agreement” is made only when the two negotiators decide on all four issue levels, the x, y-coordinates of an agreement point in Figure 1 are defined in terms of the total profit point values to each negotiator for that agreement. For example, suppose two negotiators agree on a solution (refer to Table 1): level C for clerks, level G for maintenance, level D for advertising, and level H for temperature. The overall value of the agreement is 425 profit points. The specific value of this agreement to the baker is 215 profit points (i.e., $50 + 30 + 30 + 105$) while the specific value to the florist is 210 profit points (i.e., $60 + 10 + 125 + 15$). The agreement is
Table 1. Payoff Schedule for Baker and Florist.

<table>
<thead>
<tr>
<th>Clerks</th>
<th>Maintenance</th>
<th>Advertising</th>
<th>Temperature</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 0</td>
<td>A 0</td>
<td>A 0</td>
<td>A 0</td>
</tr>
<tr>
<td>B 25</td>
<td>B 5</td>
<td>B 10</td>
<td>B 15</td>
</tr>
<tr>
<td>C 50</td>
<td>C 10</td>
<td>C 20</td>
<td>C 30</td>
</tr>
<tr>
<td>D 75</td>
<td>D 15</td>
<td>D 30</td>
<td>D 45</td>
</tr>
<tr>
<td>E 100</td>
<td>E 20</td>
<td>E 40</td>
<td>E 60</td>
</tr>
<tr>
<td>F 125</td>
<td>F 25</td>
<td>F 50</td>
<td>F 75</td>
</tr>
<tr>
<td>G 150</td>
<td>G 30</td>
<td>G 60</td>
<td>G 90</td>
</tr>
<tr>
<td>H 175</td>
<td>H 35</td>
<td>H 70</td>
<td>H 105</td>
</tr>
<tr>
<td>I 200</td>
<td>I 40</td>
<td>I 80</td>
<td>I 120</td>
</tr>
</tbody>
</table>

Florist

| A 80 | A 40 | A 200 | A 120 |
| B 70 | B 35 | B 175 | B 105 |
| C 60 | C 30 | C 150 | C 90  |
| D 50 | D 25 | D 125 | D 75  |
| E 40 | E 20 | E 100 | E 60  |
| F 30 | F 15 | F 75  | F 45  |
| G 20 | G 10 | G 50  | G 30  |
| H 10 | H 5  | H 25  | H 15  |
| I 0  | I 0  | I 0   | I 0   |

represented as the point (210, 215) in Figure 1 as the darkened point in the interior of the solution set. Because it is in the interior, it is not an optimal agreement—there are other agreements that can yield better overall totals for two negotiators. The best available (Pareto optimal) agreements for this negotiation task are shown in Figure 1 as the darkened set of points lying on the outer edge of the solution set. The goal of the negotiation is to achieve an agreement which lies on that Pareto optimal edge. In addition, each negotiator had a "walk away" value—an agreement whose value below which it is not profitable to negotiate—of 99 points. Within these constraints, there are a total of 5,683 possible agreements, 81 of which are Pareto optimal.

So how do negotiators perform on such a task? A rather consistent finding across various versions of this task is that they fail to agree on solutions which are Pareto optimal (Neale & Northcraft, 1986; Pruitt, 1981; Raiffa, 1982). Explanations of this "deviation from rationality" have generally been promoted in terms of behavioral decision theory and biases (Neale & Bazerman, 1991; Neale & Northcraft, 1991; Thompson & Hastie, 1990). Though strong arguments have been made regarding the insufficiency of such explanations to psychology in general (Berkeley & Humphreys, 1982; Cohen, 1984; Smith & Kida, 1991), our objective is to provide an alternative perspective that is not in opposition to the received view of negotiation failures, but rather it
promotes a stronger and more causally linked explanation based on an underlying model of human cognition.

What might have gone wrong in these negotiations? Our theory tells us that there are a few sources of error at the knowledge level: the goals the negotiator might be trying to achieve might be wrong, and the knowledge brought in service of those goals might be inadequate. Table 2 illustrates the strong interaction between knowledge and goals.

Goals are specified via knowledge; that is, what we know (the knowledge that is currently active for an aspect of a task) guides us in our deliberation efforts by asserting goals and working to achieve them. Therefore, the negotiators might start with a wrong assumption about the task (improper knowledge) and generate goals that are consistent with that assumption, but are inconsistent with the achievement of better negotiation outcomes. We are, in actuality, obeying the Principle of Rationality, but are not “being rational” in the broader view of the task.
Table 2. Possible Sources of Error in Negotiations.

<table>
<thead>
<tr>
<th>Task-Relevant Knowledge</th>
<th>Task-Relevant Goals</th>
<th>Jain-Kahn Knowledge</th>
<th>Jain-Kahn Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present</td>
<td>Present</td>
<td>Necessary conditions for task performance</td>
<td>Relevant knowledge not there. Could be learned on-task.</td>
</tr>
<tr>
<td>Absent</td>
<td>Absent</td>
<td>Relevant Knowledge Not Engaged</td>
<td>Relevant knowledge not there. Unlikely to be learned on-task.</td>
</tr>
</tbody>
</table>

For this negotiation task, the task-relevant knowledge generally involves sets of tactics and assumptions about the negotiation. If the negotiator’s knowledge-base is appropriately tuned to monitoring the goal and strategy invoked, the negotiator may notice that they are failing. Two options are to engage another strategy for that goal, or perhaps consider alternative goals. In essence, the negotiator is engaging knowledge to monitor knowledge. This is called metacognitive knowledge (Flavell, 1979) and reflects learnable knowledge capable of governing goal specification and strategy use (e.g., Lundeberg, 1987; Pressley, 1986). On the other hand, if the negotiator does not have the relevant knowledge to achieve a goal, but still is not satisfied with the state of the solution, the negotiator may have to engage in a search (i.e., discovery) for alternative goals or strategies. That is, new knowledge must be generated as the task unfolds—the negotiator must learn.

A Problem Solving Perspective on Two Studies of Negotiation

In order to investigate negotiation from this problem-solving perspective, we conducted two experimental studies. The first study examined the role of knowledge in negotiation (Weingart, Hyder & Prietula, 1993). In that study, naive negotiators were provided with tactical information which, if implemented, could facilitate their performance on the task. However, the results of that study showed that while negotiators with knowledge outperformed those without knowledge, many did not achieve Pareto optimal solutions. In the second study, we further explored this phenomenon by focusing on task characteristics that might impede the discovery of optimal solutions (Hyder, Prietula & Weingart, 1993). We present each of these studies below and discuss them in terms of the theory. Figure 2 summarizes this discussion as a simple framework and how the two studies were used within that framework.
Figure 2. Multilevel Framework for Viewing Negotiator Cognition.

Study 1—The Role of Knowledge in Negotiation.

As noted, critical and central in the theory is the role of knowledge. Weingart et al. (1993) examined the extent to which knowledge of a set of tactics could account for variance in negotiation behavior and outcomes. In that study, one-half of the dyads were made familiar with a descriptive list of possible negotiation tactics and cues for application to employ within a single negotiation session. The list included integrative, distributive, and neutral tactics.

The descriptions of integrative tactics provided to the negotiators included "exchange information," "do not assume a zero-sum game," and "tradeoff across issues." The first tactic, exchanging information, is believed to increase
insight into the other party’s goals and interests and improve the chance that negotiators will find integrative agreements, if a zone of agreement exists (Pruitt, 1981; Putnam & Jones, 1982; Walton & McKersie, 1965). Exchanging information about priorities across issues represents the most integrative type of information exchange, whereas sharing information about preferences within issues, being a single-issue strategy, is more distributive in nature. For the second tactic, negotiators were informed not to assume a zero-sum game because negotiators often assume that one party’s gain on an issue equals the other party’s loss. This assumption interferes with the discovery of mutually beneficial agreements on tasks with integrative potential (Thompson & Hastie, 1990). The third tactic, trading off across issues, occurs when both parties make concessions on differentially important issues to reconcile their interests. When there are differing priorities on issues, one party will concede on one set of issues to gain advantage on another set (Pruitt, 1981). Negotiators can discover mutually beneficial tradeoffs by exchanging information about priorities and/or making multiple-issue offers (packaging).

In addition to the integrative tactics, negotiators were provided with descriptions of two distributive tactics, “appear firm” and “engage in persuasive argumentation,” and one neutral tactic, “set high aspiration level.” Appearing firm is a distributive tactic used to force the other party to concede by conveying the impression that the negotiator will not move very far in the other party’s direction (Pruitt, 1981). Persuasive argumentation, or substantiation of position, is a common distributive tactic used to change the other party’s attitudes towards the issues. A persuasive argument is usually stated using attributes the other party values (Pruitt, 1981). The neutral tactic of setting a high aspiration level can serve both integrative and distributive functions. A high aspiration level for individual outcome increases the likelihood that negotiators will achieve both high individual and joint outcomes (Neale & Northercraf, 1986; Pruitt, 1981).

Weingart et al. (1993) demonstrated that negotiators working on a typical mixed-motive task (i.e., the four-issue task described earlier) who had knowledge of these negotiation tactics reached agreements of higher joint outcome as measured by Pareto efficiency—a measure of how close the outcome is to Pareto optimality. In addition, negotiators with knowledge engaged in a higher proportion of integrative behavior: they asked for and provided more information about their priorities across issues than did negotiators without tactical knowledge. These behaviors were also significantly related to higher joint outcomes. Thus, the presentation of tactical knowledge accounted for the fundamental differences between more successful and less successful negotiator dyads. However, in this study only 10 percent of the 90 dyads agreed upon Pareto optimal solutions. Subsequent analysis of that data revealed no significant effect of knowledge on the discovery of Pareto optimal solutions. Even when dyads had relevant tactical knowledge regarding ways
to achieve optimal solutions, they did not, or could not, effectively apply the knowledge in the context of negotiation at hand. They did better, but could not reach the optimal frontier.

Study 2—The Role of Cognitive Processing Limits in Negotiation.

So why did the negotiators still not achieve Pareto optimality? Returning to the theory, we find that it also suggests that there are fundamental limitations on the mechanisms underlying the knowledge level at the symbol level, attentional and processing limits on deliberation. In a second study we explored the possibility that processing limits interfered with the discovery of optimal solutions (Hyder et al., 1993). That is, regardless of the level of knowledge we expected that Pareto optimality was elusive to unexperienced negotiator due to their inability to process all the information contained in the task. When processing demands were lessened, we expected negotiators to be able to reach optimal solutions.

Two findings relating to the use of a dominant distributive tactic substantiation, in the Weingart et al. (1993) study motivated the processing limitation approach taken in the Hyder et al. (1993) study. First, tactical knowledge did not alter the proportional use of substantiation in the dyads. Regardless of the presence of tactical knowledge, the use of substantiation was much higher (24 percent of total behavior) than was the use of all integrative behaviors combined (7.23 percent of total behavior). Substantiation in the Weingart et al. study generally took the form of content-based argumentation that is, justifications for an offer or position were made based on the specific domain of the task, as defined in the negotiation materials. This finding is consistent with the general observations that distributive behaviors (often a consequence of an inappropriate zero-sum assumption) are strongly held default methods of naive negotiators (Thompson & Hastie, 1990). Second, substantiation behaviors were reciprocated in both groups. That is, occurrence of substantiation was often followed immediately by subsequent substantiations. Typically, a negotiator would substantiate a position on an issue (e.g., “I need this because…” which would be followed by the other part developing a counter-substantiative argument (e.g., “Well, I need this because…” that may involve not only the construction of an argument to support his or her promoted position, but the construction of additional arguments that would counter the other party's position. Thus, substantiation would lead to more substantiation.

Together these results indicate that when substantiation occurs, it leads to additional substantiation, and that tactical knowledge does not lessen the frequency nor dominance of this behavior. It appears that regardless of whether tactical knowledge is present, substantiation behaviors (an affiliated distributive behaviors) dominate negotiation behaviors and thus
have the potential to interfere with discovering optimal outcomes (Pruitt, 1981).

In a 2 × 2 design, tactical knowledge was crossed with two versions of the negotiation task reported in Weingart et al. (1993) and described earlier. The original version of the task was retained for one task content condition. Two store owners (a baker and a florist) were to negotiate a move to a common location. The negotiation task contained four issues (temperature, maintenance, hiring clerks, advertising) and each issue had nine option levels. Furthermore, two issues involved complex options—options that subsumed more than one choice parameter (e.g., a clerk issue involved hiring, supervision, and payment parameters). This version of the task was refereed to as “high-content” since it contained domain-specific information, potentially resulting in high task demand.

The second version of the task was deemed “low-content” and was similar in structure to the first, but with one significant change—all semantic components of the roles and options (i.e., the task “content”) were removed. No reference to the setting of the negotiation, the role of the florist or baker, nor labeling of issues was included. Yet the basic components of the problem structure (i.e., number of issues, levels per issue, relative points) were retained, so the two negotiation problems were isomorphic. Removing the task content, while preserving the fundamental structure of the problem, eliminated reference to the specific negotiation domain. This also permitted all negotiation assumptions, strategies and tactics to be applied except for one—content-based substantiation. Without a particular domain represented, there can be little basis for incorporating substantiation based on domain content, which greatly simplified the task. This alteration of the task simplified the task environment, which in turn reduced the processing demands placed on the negotiators.

We expected that although naive negotiators would probably enter the negotiation with a zero-sum assumption, the inability to engage in substantiation, a key behavior often associated with that assumption, would reduce the cognitive load. This reduction in load would release cognitive resources such that alternative approaches could be considered. With this dominant and inappropriate behavioral option inhibited, we expected negotiators to reach more Pareto optimal solutions.

The results of Hyder et al. (1993) replicated the knowledge effects found in Weingart et al. (1993). First, the presentation of negotiation tactics to naive negotiators was sufficient to significantly increase two of three integrative behaviors that were related to high Pareto efficiency: asking questions and providing information about priorities across issues. Second, the presence of knowledge did not change the frequencies of any of the distributive behaviors. Third, whereas the presence of knowledge improved Pareto efficiency, it did not influence the achievement of Pareto optimality.
In contrast, reducing the content of the task improved both Pareto efficiency and Pareto optimality. The results were dramatic: 24 of 31 dyads discovered Pareto optimal solutions as compared to 3 of 30 dyads working on the high content task. This occurred because negotiators working on the low-content task altered their use of both distributive and integrative behaviors. As expected, their use of substantiation dropped substantially (from 32 percent to 2 percent). Other distributive behaviors also dropped significantly, including "providing information about preferences within an issue" and "procedural comments regarding discussing issues one at a time." All three of these behaviors were negatively related to the achievement of Pareto optimality. Increases in integrative behaviors were evidenced in information exchange regarding priorities across issues (through providing and asking for this information), procedural suggestions regarding delayed reciprocity (e.g., "I concede on this issue now, if you promise to concede on something else later"), and multi-issue offers. These integrative behaviors were positively related to Pareto optimality.

It is not surprising that multi-issue offers were important to the discovery of Pareto optimal solutions. When issues are packaged together in an offer rather than considered sequentially, it is easier to arrange trades or concession as negotiators search for packages that are mutually beneficial. In contrast, when issues are dealt with individually, negotiators tend to compromise on each issue sequentially. This makes the discovery of potential tradeoffs more difficult, resulting in lower joint benefit (Thompson, Mannix, & Bazerman, 1988; Weingart, Bennet, & Brett, 1993).

The increased use of multi-issue offers by negotiators working on a low content task might be related to the availability of cognitive resources in that condition. Crafting multi-issue offers is a complex event. A negotiator must at least, consider the differential importance of the issues to him or herself when deciding how to construct the offer. Negotiators who are also concerned about making an offer that might benefit the other party must also estimate the relative importance of issues to that person. The elimination of content might have reduced the overall cognitive load such that more resources were available to engage in these thought processes.

What Did We Learn?

In short, we learned what makes this negotiation task difficult. The model of cognition told us that knowledge is a critical component of problem solving task performance. In the first study, we discovered that knowledge does indeed significantly improve the quality of solutions reached by the negotiator. Having task-specific knowledge is better than not having it. Knowledge matters. Negotiation is a problem solving task and to improve you need task-specific knowledge.
The model of cognition also told us that there are limits as to what the knowledge level can do within a specific task context, because of processing constraints inherent at the symbol level. In the second study, we discovered that what was impeding optimal performance (in this group) was not a lack of knowledge, but a lack of processing capacity given the nature of the task. In fact, the reduction in task demands actually permitted groups without task-specific knowledge to achieve (i.e., discover, learn) the optimal solutions. Thus, knowledge can be traded-off against task demands.

CONCLUSION AND A NOD TO FUTURE EFFORTS:
COMPUTATIONAL MODELS OF NEGOTIATION

At the onset, this chapter was based on a general approach to science, a particular approach to psychology, and a specific investigation of dyadic negotiation. The general approach to science was, in fact, to be general: to select a (mostly) complete theory of a complex mechanism and use that theory to guide an investigation into the mechanism’s behavior under certain circumstances. The psychological territory within which we roamed was based on a unified theory of cognition that, in part, asserted that the primary source of variance in problem solving task performance was task-specific knowledge, though attenuated by computational limits. And our obligations to explanation in psychology were, indeed, toward that which accounts for most of the variance of a phenomenon.

With this theoretical armament, we then explored the foundations for failure and success in a particular task—a dyadic, mixed-motive negotiation. From the results we concluded that poor performance is indeed driven by a lack of knowledge, but optimal performance cannot necessarily be achieved when the demands of the task interfere with the application (or discovery) of knowledge. On the other hand, the reduction in task demands facilitated the application, as well as the discovery, of knowledge. Knowledge matters, but so do tasks.

The story, however, is not yet quite complete. What remains to be investigated is how these naive negotiators avoid errors in different types of negotiations, how they become better negotiators in negotiations, and how we can help do both. In other words, we have to understand how these naive negotiators learn.

Learning is not a collective event, though it may be situated in a collective setting and involve interactions with multiple persons. Learning is a very personal event. What is needed, therefore, is a fundamental shift from nomothetic perspectives of the task, to idiographic perspectives of individual behaviors. Working from a universal theory of cognition allows us precisely to do that. Working from this particular universal theory of cognition allows us to do that precisely. The mechanism for that precision is the construction
of computational models of individual negotiation behavior and studies of individual learning.

A primary goal of our research, then, is to promote a theory of negotiation in the form of a computer program serving as a simulation of the cognitive processes involved in negotiation. The benefits of the unified theory on which we have based our empirical studies is that there already exists a computational instantiation of that theory as a program, *Soar* (Laird, Newell & Rosenbloom, 1987). Consequently, movement from laboratory model to computer model preserves the integrity of the theory.

The use of computational models as tools for modeling the reasoning components of human deliberation has yielded interesting insights into the foundations of human reasoning (Anderson, 1983; Posner & McLeod, 1982; Simon, 1979). In addition, this approach has been useful in explicating task specific aspects of reasoning in several domains—extensively in medicine (Clancey & Shortliffe, 1984) and to a lesser extent in business (Bouwman, 1983) as well as in reading (Just & Carpenter, 1987), planning (Wilensky, 1983), algebra (Brown & Van Lehn, 1980), particular models of skill acquisition (Neves & Anderson, 1981), and even scientific discovery (Langley, Simon, Bradshaw & Zytkow, 1987; Shrager & Langley, 1990).

Such simulations represent instantiations of process description models of behavior reflecting the similarity between cognition and computation (Pylyshyn, 1984). The nature of computational metaphors allows for the expression of functional models on a certain level of abstraction, independent of a specific form of materialistic interpretation (Gunderson, 1985; Pylyshyn, 1978). A program (as a simulation) is a tool which allows psychological hypotheses to be formulated and tested. Computer models, as experimental data, can confirm or contradict hypotheses about reasoning (Johnson et al., 1981). This permits working with concepts that are insufficiently explicit for mathematical expression or too complex to be described by tractable mathematical models. Simon (1981) notes the role of computational models in science:

> Testing a theory of computer problem-solving by constructing a computer program that solves problems—presumably in a humanized way—at first sounds strange. However, it is no stranger than testing our theories of the RNA code by synthesizing macromolecules from nucleotides and then by observing their efficacy in synthesizing proteins. What we are testing, at the outset, is the sufficiency of the mechanisms we have postulated to produce the phenomena (problem-solving) in which we are interested. Thus, a simulation program for solving problems demonstrates, if it is successful, that problems of the variety under study can be solved with a program of the characteristics that have been assumed (p. 300).

Given the goals of the proposed research, we propose that there are for primary reasons for incorporating knowledge-based modeling. First, knowledge-based model permits an examination of the kind of reasonir
phenomena which "go together" in solving a negotiation problem: a set of mechanisms operating dynamically. This requires examining how the mechanisms behave individually and collectively over time to produce negotiation problem solutions. The resulting description must proffer mechanisms and organizations that explain the cognitive events occurring in the time frame of the negotiation. Such an endeavor could, in theory, be done without a computer model (i.e., the rules of reasoning could be written in English or some artificial "simulanguage") but the level of complexity required in real world tasks precludes this tactic (Schank commentary in Searle, 1980).

A second reason, related to the first, is the interest in describing the processes of problem solving as well as the final product. A knowledge-based model not only contains the posited mechanisms derived from the constructs of cognitive psychology, but they are formulated in such a way as to generate a trace of the processes produced by the mechanisms as problem solving ensues (Johnson et al., 1981). That is, the particular mechanisms are invoked, a record is maintained and made available for subsequent review and analysis. Thus, in one sense, there is a reductionist bent to the approach which specifies how problem solving may be explained by an organized set of sub-phenomena (Newell & Simon, 1972). Yet, reductionism is a fundamental and useful approach to understand complex, dynamic systems and especially in the explication of complex problem solving. We not only must have a theory of the communication, but we must have a theory of the communicator. Furthermore, the theory of the communicator must account for ALL relevant behaviors (and constraints on behavior) at the particular level of representation selected. In essence, we propose you cannot have a "theory of negotiation;" rather, you must have a theory of a human "negotiating" with the fundamental machinery of thought used in everyday deliberations. Negotiation is what the human is doing; deliberation is the mechanism by which the human is doing it.

Third, the use of a knowledge-based model enforces a rigor and uniformity in the description of a set of phenomena. It provides descriptive uniformalisms which tend to promote theories (in the form of programs) which are "homunculus resistant." If a mechanism or process is articulated in theory, even if it is unobservable, it must be articulated in code.

This is the philosophical *sine qua non* of computational models. One is obliged to move from construct to phenomenon when the theory (and the science) permit it. Essential to this concept is the realization that levels of abstraction of theory constrain (and direct) the levels of representation and form of performance of code (Anderson, 1990; Marr, 1982; Newell & Simon, 1976; Pylyshyn, 1984). Thus, if we are attempting to understand how two parties negotiate together, we must first have a strong model of how two parties reason individually. The knowledge and mechanisms of reasoning are expressed solely in terms of the constructs supported by the underlying computational theory.
Finally, the mechanisms embodied in the knowledge-based model may be manipulated in order to test subsequent effects on reasoning behavior. While humans, it is quite difficult to tamper with the underlying mechanisms of thought. Such attempts are generally made by manipulating the experimental environment (e.g., through orienting instructions, goals, task materials, or deception). There is always a need for explicit manipulation checks and the specific effects of manipulations may not be properly anticipated, detected, or measured. The use of a computational model, however, allows such manipulations. Furthermore, manipulations may be used to make predictions and propose explanations of behavior (Johnson et al., 1981).

As we have noted, predictions may be made concerning behavior, but tests of actual model performance must be made to verify them. This is directly related to the issues underlying model validation. Three general, and increasingly powerful, methods can be brought to bear: sufficiency testing, process tracing, and component analysis.

* Sufficiency testing* is the weakest form of validation and focuses entirely on the outcome of the behavior. In essence, it states that a model should at least be able to produce the behavior it purports to explain. A program that models negotiation behavior should be able to generate a negotiation solution, though perhaps an inadequate one.

* Process tracing* makes a stronger statement. Going beyond the task of showing that an unspecified set of mechanisms produces a certain result, this test demonstrates that particular mechanisms (or knowledge) can produce the behavior. In this test, comparisons are made at some level of abstraction between the model and a referent (a proposed gold standard). The level of abstraction at which our models are written are consistent with the underlying theory of cognition; that is, the negotiation task is subsumed within the computational model of intelligence.

* Component analysis* examines specific contributions of the mechanisms or knowledge represented in the reasoning events. In particular, our hypotheses about negotiation center around the availability of specific forms of knowledge and specific attentional constraints on deliberation. We do not posit "new" cognitive mechanisms.

Clancey (1984) offers a strong view of the role computers play in theory development and testing, arguing that "too often experimental analysis seems to fall short by not being precise enough to be programmable" (p. 78). The critical goal of computational model development is the construction of a program which addresses the entirety of the task on a level of abstraction which permits testing of the theory.

The development of a computational model of negotiation, however, is not done in isolation of (more traditional) human experiments. In fact, the development requires data from such studies for design, validation and verification. Such studies have demonstrated and documented important
phenomena occurring in negotiation. If our computational model is to contribute to the understanding and explanation of negotiation cognition, it must be able to reproduce these phenomena. But it must reproduce these phenomena in terms of the underlying model of cognition embedded in the theory. Estes (1978) identifies the significance of this distinction:

Doubtless no one would disagree that the first step toward theory construction in any scientific area must be an adequate description of facts, accomplished by means of terms that do not presuppose the theory. What is not so obvious is that there may be alternate frameworks for the description of facts and lower level empirical laws, and the decision made at this choice point may have profound implications for the theory. (p. 1)

Consequently, future efforts must incorporate two sympathetic forms of inquiry: experimental human studies as well as computational models.

Incorporating computer simulation methods to understand the dynamics of human negotiation is a relatively new endeavor; however, researchers are indeed incorporating this approach to explicate the nature of the processes and outcomes of negotiation. For example, Darling and Mumpower (1990) have begun to model the progression of offers which occur in mixed-motive negotiations.

We conclude our chapter with the assertion that the ultimate goal, as we see it, of understanding negotiation problem solving is to be achieved by explaining how individuals use knowledge to solve tasks. As such, insight is gained by studying explicit individual behavior and not by studying the “average” behavior of groups. Individual negotiation cognitive behavior is seen neither to be an ineffable quality of the mind, nor a social epiphenomenon, but a legitimate and critical target for investigation. And computer models are a legitimate and critical method for conducting that investigation. In 1956, cognitive psychology was spurred by Newell and Simon’s Logic Theorists. Perhaps the next major step in negotiation research is a Negotiation Theorist.

NOTES

1. Definitions of terms and the description of the negotiation task will be presented later in this chapter.
2. Of particular interest, it discovered a new proof to their Theorem 2.85. In their publication of the results, however, Newell and Simon failed to convince the editors to place the Logic Theorist as a co-author.
3. The specification of levels or layers in research are ubiquitous in science. Chomsky (1980) and Simon (1969) offer specific discussions of the role that levels play in theory formulation and testing in psychology.
4. Connectionists, on the other hand, define algorithms and procedures which are an inherent part of the computation theory comprising the approach (Pylyshyn, 1989), thus minimizing the independence (or even definition) of conceptual levels of abstraction. Perhaps these two approaches are not in opposition, but are currently theoretically incommensurable as their linkages or relationships not yet sufficiently specified (cf. Fodor & Pylyshyn, 1988; Harnad, 1990).
5. The reason for "symbol" (or "representation") structures is based directly on what we know of neurobiology: neural stimulation (persistence) and plasticity (change) are based on biochemical and structural changes in response to patterns of excitation (Squire, 1987) embedded in cortical and subcortical neural circuitry (Sheppard, 1988, Chapter 30). Thus, the only communicative and representational events that can occur in the brain are based on sensing or communicating patterns (functioning as symbols). Consequently, symbolic processing serves as a basic universal mechanism for composing general functions for distal access of associated, but distributed, symbols (Newell, 1990). There is strong equivalence between the potential behavior (and descriptions) of different types of physical mechanisms operating as symbol systems (Pylyshyn, 1989).

6. Note that rationality is defined in terms of the behavior of the architecture with respect to the knowledge it contains and the goals that are most salient (i.e., are most attended to) at a given time. This is different, for example, from Anderson's (1990) conceptualization of rationality defined in terms of experience ("to optimize the adaptation of the behavior of the organism" p. 28).

7. By naive, we mean that these negotiators have not had any formal training or experience in negotiation.

8. This is also referred to as the BATNA—Best Alternative To a Negotiated Agreement.

9. The concept of a homunculus has a long history in psychology, philosophy and religion. Essentially, in an attempt to explain a phenomenon or property, one posits a "little man" (i.e., an entity) inside of something that grants it the phenomenon or property of interest, thus actually explaining little and setting the foundation for an unfortunate series of regressions (i.e., what is it inside of the homunculus that grants IT the phenomenon or property?). In psychology, for example, this concept has arisen, directly or indirectly, in the attempts to explain the mechanisms that control information processing (Barsalou, 1992).

REFERENCES


