

COMPUTABLE SOCIAL COMMUNICATION

A DISSERTATION SUBMITTED TO THE GRADUATE
DIVISION OF THE UNIVERSITY OF HAWAI'I IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

COMPUTER SCIENCE

MARCH 2007

By

David S. Pautler

Dissertation Committee:

David Chin, Chairperson

Martha Crosby

Dan Suthers

Kim Binsted

Ben Bergen

We certify that we have read this dissertation and that, in our opinion, it is satisfactory in scope and quality as a dissertation for the degree of Doctor of Philosophy in Computer Science.

DISSERTATION COMMITTEE

Chairperson

Dedicated to
Mom and Dad

Copyright 2007 David S. Pautler

ACKNOWLEDGMENTS

Pursuing my PhD has taken more than seventeen years, including two universities and three advisors. And for the last eight, I have worked remotely from California while working fulltime and having a new family. If I had known it would take even five years, I might never have started. But each of those years was full of learning experiences and new friendships.

I'm especially grateful to Bill Ferguson, a grad student full of a love for ideas and debate, who mentored me in my first two years, then at Northwestern University. Bill convinced me to go to a symposium at Stanford, where I met Alex Quilici. Alex was a student at the time, but later offered me a ticket out of Chicago to Hawaii as a researcher, and later, to a startup in the heyday of Silicon Valley. Alex has been simply invaluable as moral support, coauthor, boss, and friend.

David Chin, my advisor, has been a terrific guide and supporter over the past ten years, especially in keeping things on track while we collaborate long-distance.

Dan Suthers, my one-time boss and current graduate chair, has saved me more than once.

Kathy Kellermann, the author of CCT, has been very generous in listening to my ideas and suggesting relevant research in her field.

Kelly Aune in the Communication department, and Arne Jönsson and Martha Crosby in Computer Science have had great suggestions. I really appreciate their enthusiastic support.

I'm grateful to all the people at Northwestern who fired my initial excitement about cognitive science – Larry Birnbaum, Andrew Ortony, Dedre Gentner, the Eric's (Jones and Domeshek), Andrew Fano, Pete Prokopowicz, and Matt Brand. And I'm grateful to Roger Schank for creating the lab and providing full funding to me for many years.

My friends near and far -- Blanca & Julio Polo, Ben Bell, Mike Fiske, Jorn Barger, Steve Woods, Kirsten Wehner, Rick Cordes, John & Kari Embree, and Jeff Stark – for making time away from work so enjoyable.

And most of all, I'm grateful to my parents and siblings, Ed, Carla, Dawn, Dan, and Jill for a happiness extending from childhood into middle age. And to my wife Sarah and son Charlie, for starting my life on a new kind of adventure.

ABSTRACT

Computer programs that track personal information, that give advice, and that play other social roles are becoming embedded ever deeper in the fabric of everyday life. To perform in such roles, programs must be able to interact with people effectively and appropriately. Furthermore, the variety of social situations in everyday life precludes anticipating all the situations a program would need to know how to handle and coding that knowledge “at the factory”; instead, programs must have dynamic social intelligence. We propose a research strategy for AI and communication scholars to pursue for mutual benefit in building better social software and in formulating theories that can be tested through simulation.

In following this strategy, we have built a simulator based on a model in communication, and the simulator generates plans from a wider variety of persuasive goals and speech acts than any similar AI system. Our simulator, named Computable Social Communication (CSC), is based on Conversational Constraint Theory (CCT) developed by Kathy Kellermann and her students. CCT predicts how acceptable any of 56 speech acts would be for any of 13 persuasion goals, allowing for many variations in the situation, and the theory is strongly supported by several empirical studies done by its authors. CSC adapts CCT by making it computable and extends it to multi-goal, multi-act communicative plans. In human participant tests with CSC involving sixteen hypotheses about plan quality, ten hypotheses received at least 95% support. This result held in both an online unrestricted pilot with 54 participants and a formal study with 23 undergraduates.

TABLE OF CONTENTS

Acknowledgments	v
Abstract.....	vi
List of Tables	viii
List of Figures	ix
Chapter 1: Introduction	1
What We Are After	3
How We Plan to Get There	7
How We Will Evaluate Our Work	10
Summary	11
Chapter 2: Related Work	12
Linguistic Pragmatics	12
Social Psychology	16
Artificial Intelligence	22
Speech Act Planning	22
Multi-Agent Systems	31
Sociable Agents	31
Communication	38
Chapter 3: A Computable Model of Social Communication	50
Research Strategy	50
Adapting CCT	51
Extending CCT	56
Chapter 4: Evaluation 1 – Subjective Plausibility	66
Sample One: The Reckless Driver	67
Sample Two: The Reckless Driver and the Impolite Interface	68
Sample Three: The Embarrassing Husband	70
Chapter 5: Evaluation 2 – Predictive Power	72
Survey Design	72
Lessons from the Pilot	78
Survey Results	88
Chapter 6: Conclusion	94
Future directions	95
Appendix A: A Comparison of Scores Across Two Studies	102
Appendix B: Download the Simulator	106
Appendix C: CCT’s Factors Determining Minimum Preferred Levels	107
Appendix D: Transcript for “The Reckless Driver”	110
Appendix E: Transcript for “The Reckless Driver and the Impolite Interface”	116
Appendix F: Transcript for “The Embarrassing Husband”	123
Appendix G: Pilot Questionnaire	127
Appendix H: Survey Questionnaire	137
References	141

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1. D’Andrade and Wish’s plotting of speech acts.....	19
2. Bell and Daly’s affiliation-seeking.....	40
3. Proposed envisioning of relationships among theories.....	41
4. Persuasion goals used in 2004 CCT study	48
5. Speech act types used in 2004 CCT study	48
6. Redundant predictions where one found little or no consensus	86
7. Redundant predictions where one found strong disagreement	86
8. A comparison of scores across two studies	102

LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1. The “shape” of our desired solution	5
2. A conceptual map of influences	9
3. Wierzbicka’s analysis of “promising”	14
4. Kiesler’s Interpersonal Circumplex	17
5. Acts and effects in adjusting a relationship	28
6. Bickmore’s equations	30
7. A person interacting with REA	30
8. Facial expressions of Breazeal’s robot	37
9. Action selection in SIMs™	37
10. How CCT determines acts from chosen goal	43
11. How CCT adjusts thresholds	44
12. How CCT selects acts by overlaying thresholds	45
13. Four cases guiding design of the pilot	76
14. Pilot results	81
15. Pilot results with appropriateness threshold moved to 6.18.....	83
16. Pilot results with appropriateness threshold moved to 5.1.....	84
17. Survey design and results	90
18. Survey results with both thresholds adjusted.....	92
19. Pilot questionnaire	127
20. Survey questionnaire	137

CHAPTER 1

INTRODUCTION

If one stops to think about the challenges of everyday life, it is remarkable how many are social in nature: There are goal conflicts to resolve, difficult personalities to deal with, debts to pay and collect, relationships to maintain, and ambiguous situations to navigate. All these require intelligence and skill, and yet surprisingly, the field of Artificial Intelligence (AI), which is largely defined by its concern for problems that are challenging to human intelligence, has paid much less attention to this area than for others of similar scope. In particular, there is very little research attempting to simulate human social skills. The reasons for this neglect are not well-known, but it is time to do something about it. That is the aim of this dissertation – to begin to fill the gap.

Why is it that this has not been done before? We suspect it is because it is very hard to create a theoretical framework for capturing social phenomena. For simple frameworks, it has been fairly easy to find phenomena that do not fit, and more comprehensive frameworks quickly become too unwieldy to explain or use. This has certainly been true in our previous attempts to create a computable representation of social competency, and anyone who looks into research in social psychology, communication, and related fields would quickly come to a similar conclusion: There is a host of complex issues about how agents weigh situations and goals, including how they infer the goals and attitudes of other agents, how they perceive their relationships with those agents, and how they strategize for the goals that they adopt. But while there may be consensus across the social disciplines on what the issues are, there is little consensus on what theoretical

framework is most suited. By comparison, there has been a great deal of work in AI on navigating and manipulating the physical world, and the physical world seems just as complex as the social world. And yet, there is widespread consensus on how physics, chemistry, biology, and even human medical practice relate to each other in a reductive manner. Such a consensus framework makes it much easier to create an AI model of competence in those areas (although it is a model of expert competence rather than the naive competence that most humans have about the physical world). There is no comparable framework in the social sciences to tell us how to connect, say, the dimensional models of relationships in social psychology [Svennevig 2000] to the strategies for managing relationships in communication [Berger 1997]. It is understandable that AI researchers would prefer to focus on the physical world where borrowing from other disciplines is less fraught with the risk of misusing findings.

Given this risk, why should AI pursue a model of social competence now? In short, because we have to. As computers become more embedded in human life, the more their design must be informed by an understanding of human social behavior. Take the example of in-car voice interfaces. Research by Cliff Nass and his colleagues at Stanford has shown that elderly drivers do significantly better on the road when their cars are outfitted with systems that use a recorded human voice to give warnings and suggestions [Jonsson 2005]. That study shows the potential benefit of a social interface, yet another study by the same group [Jonsson 2004] found that if the emotional tone of the voice did not match the current emotion of the driver (i.e., aroused voice/aroused driver, or relaxed voice/relaxed driver), then the social interface actually caused the driver to do worse! The

potential benefit of that in-car interface is the start of a wide range of services that automakers want to install in their cars, so it is not likely that the risk of a bad interface can be avoided by not using any interface at all; there is too much momentum in the industry. Instead, we have to ameliorate the risk of a bad social interface by making them smarter. Currently, the best way of doing that is to do usability tests of the interface in a lab setting and work out solutions for all possible situations before the product ships. For the currently small range of situations that in-car advisory systems deal with, such a ‘static’ solution works, but as the scope of such interfaces inevitably widens, there will be a need for ‘dynamic’ social intelligence.

What We Are After

The aim of this dissertation is to find an answer to this question: How can we simulate how a person pursues a social goal, or even several simultaneously...

- through the wide range of communicative tactics observed in everyday conversation...
- where the best option (or even just an acceptable one) will depend to a great extent on specifics about the people involved and/or the situation?

Our objectives are focused on the planning and acting stages of a complete observe-interpret-plan-act action cycle. Within the relatively small body of research in AI on modeling social behavior, the majority has focused on the “front” end of the cycle, on observing and interpreting human emotional and attentional states. For example, Rosalind Picard has built prototypes for sensing emotional state from the way users touch

keyboards and mice [Picard 2000]. Our aim is to provide a rich simulation of the other end of the action cycle, and in future work tie the two ends together.

Although our objective is a simulation, this research question follows the form of most quantitative studies, where the objective is to capture the important independent and dependent variables of an observed behavior and map out the relations from the first set to the second. In our case, social goals are the independent variables, communicative tactics are the dependent variables, and the dependency relations from goals to tactics are expected to have many contingencies based on the conversational context [Figure 1]. If one follows this analogy as a research strategy, one should identify what domain of conversation to focus on, what social goals and communicative tactics are representative of that domain, and what goals lead to what tactics under what conditions. But pursuing a simulation requires more than a quantitative study would; specifically, a simulation must show *how* the mapping of goals to tactics can recreate the observed behaviors that led to the mapping in the first place. More specifically, the openness of simulations to playing with scenarios encourages attempts to “break” the mapping, revealing where the theoretically-weak points are (if any): Perhaps the mapping fails to cover a certain scenario, or is found to be too vague to be simulated, and so on. There are many ways to implement a simulation, such as through thought experiments or paper prototypes, but the most stringent simulation test seems to be rendering it as a computer program, since computers have no preconceptions that might fill in or “paper over” unseen gaps in the model. For these reasons, our strategy is to make sure we create a model of *computable* social communication.

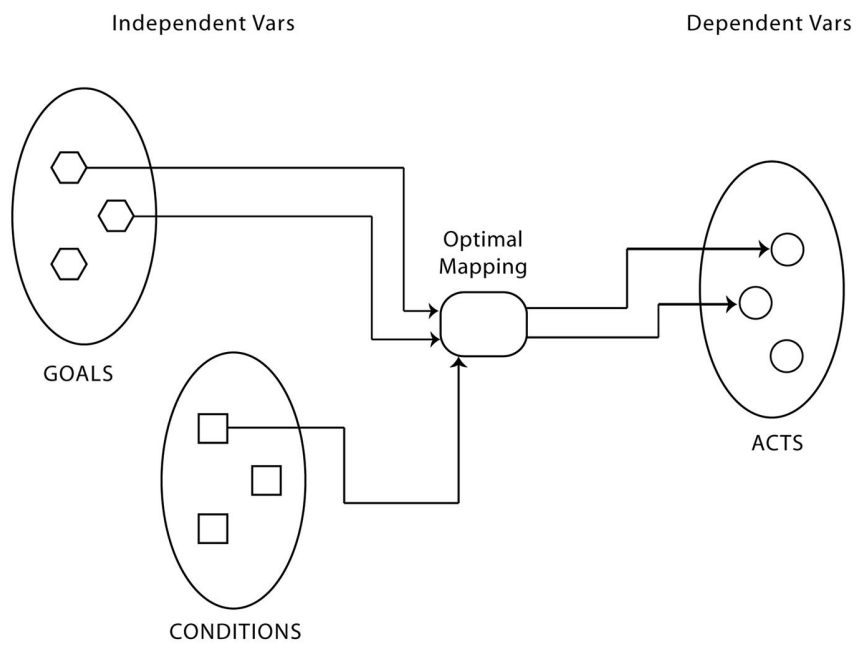


Figure 1. The “shape” of our desired solution

Although our model will be computable, and although we have discussed a need for a practical solution, our model is intended as a contribution to theory rather than for use in consumer-facing applications. We aim to create a cognitive model of the ‘naïve’ social competence a person uses when trying to persuade another person to do something or when trying to manage the relationship between the two people. It is a model of how one person interacts with another, rather than a model of how a computer should interact with users. More specifically, it will not be a direct solution to practical problems like those of an in-car driving advisor, for two reasons. First, an advisor interface needs to be an expert both in the domain of driving and in the area of persuasive communication; using naïve social competence could be ineffective or even dangerous. As an example of the ‘naïve’ view, most people are surprised that the relaxed voice is counterproductive for aroused drivers. Similarly, much of the research in social psychology and communication that we might use focuses on the social competence of college students (the typical subject group) which is certainly ‘naïve’ competence for most applications. So direct, practical application is limited by the data we have available. But the second and more important reason is that creating a model of naïve social competence has value in itself. In particular, it is hard to imagine how one could be an expert in persuasive communication without understanding the people one is trying to persuade without knowing their responsive chords and without knowing what buttons to push. That knowledge is knowledge of naïve competence. So, we have a clear idea of how we expect this research to bear practical benefits further along. However, the direct contribution of this dissertation is to theory.

Our first step is to identify the boundaries of the phenomenon we want to study. The scope of this project is intentionally delimited to social interactions that occur a) through verbal communication, b) between just two people, c) where one is consciously trying to persuade the other, and d) where those people are adult Westerners of average intelligence. We have adopted these delimitations for several reasons:

- Overt, verbal communication is the primary means of social interaction;
- While communication between pairs of interactants is different from communication within and across groups, studying pair-wise communication between individuals seems to capture something fundamental that should apply to larger networks;
- There is a wealth of research on how verbal communication within this demographic contributes to social interaction, much more than we were able to find for non-verbal interactions or for interactions within or across non-Western groups;
- Limiting the goals of conversation to persuasion allows us to tap into the rich literature in communication on “compliance-gaining.” There is less communication literature on other kinds of goals, such as “affinity seeking” (i.e., exploring the possibility of escalating a cordial relationship).
- The existing research within this scope indicates that the issues are already complicated enough that a wider scope would be unwieldy and difficult to evaluate.

How We Plan to Get There

We cast a wide net over the fields of AI, social psychology, communication, and pragmatics looking for approaches similar to ours, or ones that would argue against

pursuing it. Figure 2 represents our sense of the conceptual space of theories relevant to strategic interpersonal communication, and particularly theories that attempt to model reasoning in this problem space. We did not find any work that argues against our approach.

The next chapter describes the conceptual space in greater detail, but in brief, most of the related efforts have been about creating taxonomies of strategies rather than simulating the execution of strategies. One exception is Bickmore's work on a 3D computerized agent that attempts to sell real estate by mixing socially-appropriate small talk into its sales pitch [Bickmore 2003]. Another exception is Conversational Constraint Theory (CCT), developed by Kellermann and her students [Kellermann 2004, 2006; Kellermann & Park 2001; Kellermann & Cole 1994]. CCT targets the same types of communication as we described in the previous section, and shares our objective of mapping goals and conditions to relevant actions. But CCT does not describe a way of simulating the pursuit of a strategy (nor, obviously, multiple strategies simultaneously, as we plan to).

We concluded from our review of related work that seeking out active collaboration with communication scholars would be more productive than creating our own theory, because:

- They have deeper experience with social communication phenomena;
- They have empirical data from real subjects in larger sample sizes;
- Their theories agree and disagree with AI theories in interesting ways;
- They lack precise propositional representation and/or simulations;
- Dividing up the work based on strengths should result in better, faster results.

We embraced Kellermann's CCT model in particular, with the aim of extending it to support computable simulations, particularly scenarios involving multiple goals and actions.

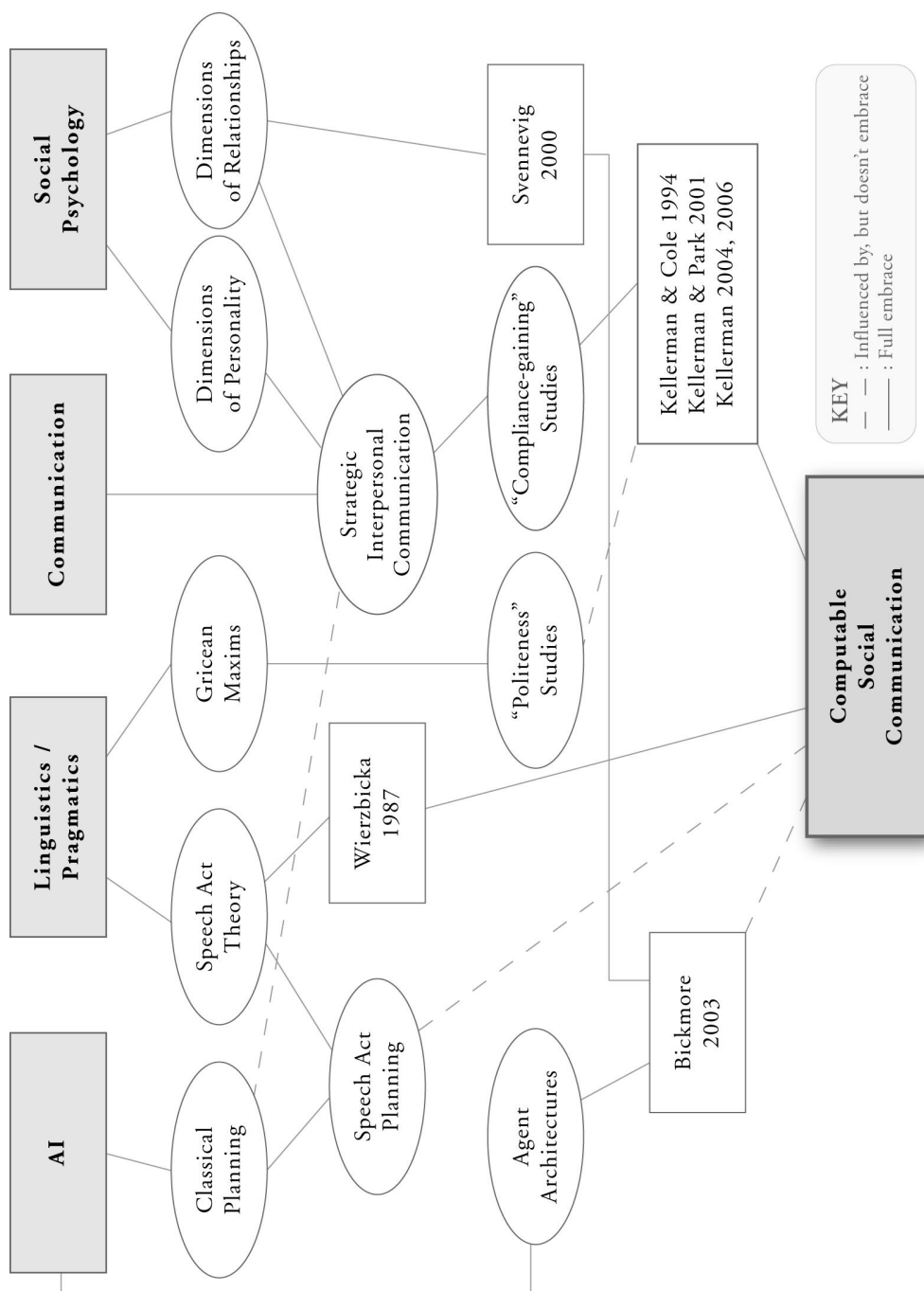


Figure 2. A conceptual map of influences on our approach

A consequence of our choice to collaborate is that we need to keep our computable model as similar to the original formulation of the theory as possible, so the model does not resemble traditional AI formulations of a problem space such as plan operators in speech act planning. Instead, we use a simple mathematical optimization technique for selecting actions for goals.

An advantage of the optimization technique is that it allows the model to weigh actions that only partially help or impede a goal, and to do so for multiple goals simultaneously, neither of which is possible with traditional plan operators. Bickmore's agent also uses mathematical optimization techniques to select actions for goals, instead of traditional planning, but his choice seems driven by the need for real-time responses rather than to promote collaboration or to model the partially-effective nature of real-world actions.

How We Will Evaluate Our Work

We will evaluate our model in two ways: First, through the traditional AI method of walking through the model's performance on a set of putatively representative sample problems, with the expectation that the reader would agree that the generated solutions are plausible. And second, we will do an empirical study where we use the model to generate two-goal plans that should be acceptable for both goals, only one goal, or neither, and do a survey to see how the acceptability predictions compare with human judgments of the same plans.

The survey results show a supportive trend for the model's ability to predict whether a certain set of speech acts would be an acceptable way to pursue a certain set of interpersonal goals when particular aspects of the situation are known.

Summary

Although there is research in AI and other fields on the social interactions between agents and humans, the problem of simulating how people communicate in a socially-appropriate and effective way is still largely unexplored. Understanding the range and subtleties of human skills is essential to building agents with comparable skills. And as software plays an ever greater part in personal life, it must become more socially-aware; socially-ignorant software has been shown to be counter-productive and even dangerous.

There is no generally agreed-upon theory of strategic social communication, much less one ready for formulation as a computable model. But one line of research in communication, Kellermann's Conversational Constraint Theory, seems especially promising. We have embraced this line of research and have extended it into a computable model that can generate plans for interpersonal communication. Furthermore, an empirical study shows a supportive trend for the predictions of the model.

CHAPTER 2

RELATED WORK

In our search for a way of generating socially-appropriate and effective communication, we cast a wide net over the fields of AI, social psychology, communication, and pragmatics. We were surprised that instead of finding a small set of shared approaches, the fields seem to have developed largely in isolation from each other, but with some intriguing family resemblances that suggest occasional intermarriages have occurred.

Linguistic Pragmatics

Linguistic pragmatics is a subfield spanning linguistics and the philosophy of language, and it focuses on language's relation to the world – for example, how a situation can oblige different levels of formal language in different cultures. (In contrast, the more widely-known subfields of syntax and semantics focus on language in isolation.) Another way that language and the world interact is when language changes the state of things, such as the way a promise might change the attitude or behavior of those who hear it. This way of acting on the world through language is the focus of Speech Act Theory in pragmatics. Because Speech Act Theory examines how language can be used intentionally to affect people, it is directly relevant to strategic interpersonal communication. It is also the earliest relevant work we know of, growing from work done in the 1950s by the philosopher of language J. L. Austin [1975]. The theory's influence has spanned several disciplines as well, influencing research in communication and AI.

Much of the research in Speech Act Theory (SAT) [Searle 1969, Searle & Vanderveken 1985], involves cataloguing performative verbs (i.e., those that fit the formula, “I hereby X,” whereby speaking such a sentence one effects some change), and situation conditions that are required for the performative to have its usual effect (e.g., a promise must appear sincere). The most comprehensive catalogue, Wierzbicka’s dictionary of English performatives [1987], has 270 entries, each representing a different way of acting on the world (e.g., announcing, baptizing, and reporting). The vast majority of the performatives in Wierzbicka’s dictionary have an obvious social dimension (e.g., thanking, threatening, and confessing). It is tempting to view the dictionary as a catalogue not just of performatives, but also of all the ways one person might influence another through language. Yet, Austin claimed that there is no practical limit to the number and variety of speech acts [Austin 1975]. Furthermore, the notion of a “speech act” is broader than that of a performative verb, since one can achieve such an act without actually using the associated performative (for example, one can promise by saying “I owe you one” without having to use the word “promise”). Nevertheless, speech act catalogues like Wierzbicka’s are rich sources of insight into the range of strategies and situations in which interpersonal influence can occur.

As an example of the relevance of SAT catalogues to social communication, consider Wierzbicka’s analysis of promising [Figure 3]. The effect of the act is to *persuade* the hearer that the speaker is *obliged* to carry out what is promised. And for the promise to appear sincere, it must appear that the speaker’s *reputation* is at stake if he fails to follow

Conditions:

- I know that you want me to do [action] A.
- I know that you think I may not do it.
- I want to do it because you want me to do it.
- I want us to think that if I don't do it, people will not believe anything that I say I will do.

Dictum:

I say: I will do it.

Illocutionary Purpose:

I say this, in this way, because I want to cause you to be able to think that I have to do it.

Figure 3. Wierzbicka's semantic analysis of performative verb 'promising'

[Wierzbicka 1987, p. 205]

through. The concepts of persuasion, obligation, and reputation are inherently social; furthermore, their linkages to the act of promising suggest how failing in one or more conditions could have a variety of social outcomes, such as the speaker appearing to be desperate (through trying too hard to persuade) or untrustworthy (through failing to follow through, intentionally or unintentionally). That is, such a catalogue of acts suggests a much larger catalogue of socially-relevant concepts and linkages among them.

One limitation shared by all of the analyses since Austin is an emphasis on one of the earliest effects of acts on hearers, namely, uptake. Later effects on psychological, social, and practical concerns (categorized as "perlocutionary effects" in pragmatics) have been ignored, largely because linguists consider such effects to lie outside the subject of linguistics.

Politeness Theory is another subfield of pragmatics relevant to social communication [Brown & Levinson 1987]. Its main thesis is that in conversation, each participant must weigh a tradeoff between speaking *efficiently* and speaking in a *socially appropriate* manner. Speaking efficiently means adhering to Gricean maxims, which identify four primary dimensions of efficiency: informativeness, truthfulness, relevance, and perspicacity [Grice 1989: pp. 26-27]. Speaking in a socially appropriate way means attending to the *positive* and *negative* faces of one's audience and oneself, where "positive face" means the image one wants to project, and "negative face," the set of obligations one wants others to observe. Brown and Levinson claim that speaking appropriately requires considering each of one's possible paths forward in the conversation for its *face-threatening* potential, where the degree of threat is determined by cultural factors, the relative power between the participants, and their social distance.

Brown and Levinson claim that speakers who are concerned with being polite judge the degree to which their acts may be face-threatening before carrying them out, and that they use a set of strategies for averting escalating degrees of threat. These strategies are:

1. If the threat is low, simply do the act. (This is the same as following Gricean maxims.)
2. If the threat is minimal, perform the act but indicate to the recipient that no harm is intended, or promote his positive face in another way (e.g., compliment him).
3. If the threat is moderate, perform the act but find a way to promote the recipient's negative face (i.e., his independence) (e.g., offer to help him).
4. At higher levels of threat, the act should be communicated indirectly or abandoned.

Politeness Theory encompasses a rich set of concepts relevant to social interaction, allowing it to describe a wide range of conversational scenarios. And it fits well with

Speech Act Theory, since the former describes conversation at the strategic level while the latter addresses the tactical level (although we are not aware of any research that knits the two theories together into a single process model). Yet, some criticize Politeness Theory for oversimplifying; for example, it assumes that participants want to maintain the status quo in their relationship rather than perhaps enhance it or create distance. It also fails to take into account the role of personality in attending to and prioritizing issues of appropriateness or efficiency.

Social Psychology

Within social psychology there are a number of different frameworks for analyzing social interactions. One approach has been to develop a classification system for aspects of interpersonal behavior (e.g., mercifulness) based on a core set of behavior aspects. Several such classification systems have been investigated (see [Kiesler 1983] for an overview), although for their core sets they all use behavior aspects that are similar to a basic set of four. These four behavior aspects are: *friendliness*, *dominance*, *task-orientedness*, and *degree of emotion*. The core behavior aspects are usually envisioned as dimensions of a space in which non-core behavior aspects can be plotted as points. For example, the behavior aspect mercifulness might be represented as a combination of a high positive value for friendliness, a high negative value for dominance, and near-neutral values for task-orientedness and degree of emotion. Figure 4 illustrates Kiesler's plotting of many behavior traits for just the dimensions of dominance (on the Y-axis) and friendliness (on the X-axis), placing the trait of mercifulness in the lower right corner in accordance with the analysis already described.

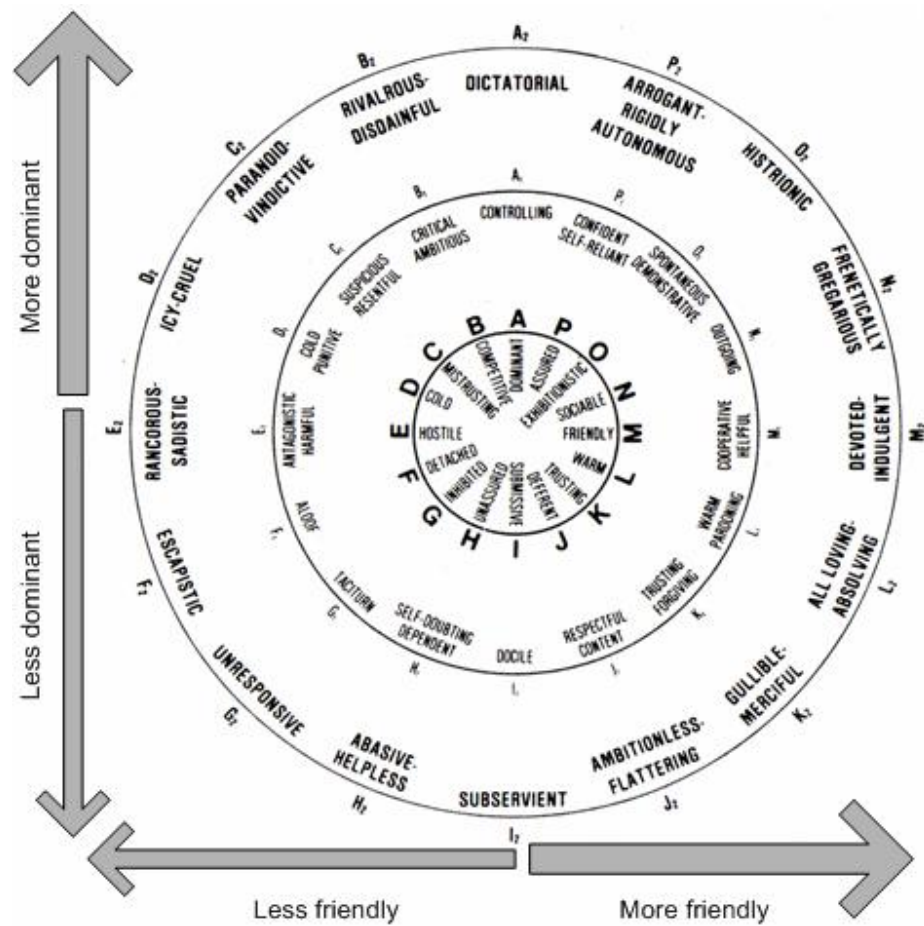


Figure 4. Kiesler's Interpersonal Circumplex

One objective of the behavior classification studies is to test a long-standing hypothesis in social psychology (and in communication, e.g., [Bateson 1958; Goffman 1967]) that the behavior of participants in social interactions is *complementary*. For example, Kiesler [1983] suggests that “complementarity occurs on the basis of (1) reciprocity in respect to the Control dimension or axis (dominance pulls submission, submission pulls dominance), and (2) correspondence in respect to the Affiliation dimension (hostility pulls hostility, friendliness pulls friendliness)” [p. 201]. Using this hypothesis, and a plotting of behavior aspects along dimensions of dominance and friendliness, predictions about particular complementary interactions can be generated. For example, merciful behavior (submissive and friendly) might be predicted to pull “histrionic” behavior (dominant and friendly). Kiesler also suggests that such complementary interactions are usually automatic and unintended. Such theories of circle-plotted sets of behavior traits and complementarity rules are known collectively as Interpersonal Circumplex Theory (IPC).

The four dimensions of IPC have also been applied to speech acts by D'Andrade and Wish [1985]. In their research, test subjects used the dimensions as scales for rating a variety of speech acts performed by family members seen interacting in a documentary film. Because these ratings represent impressions made on hearers, although not the intended hearers, of particular speech acts, the ratings represent certain perlocutionary effects of the acts. Table 1 is an abbreviated listing of the act types found to produce the most extreme effects on each of the four dimensions. Each column of the table represents one of the four scales, with the speech acts producing the most extreme values on a scale appearing at the top and bottom of the associated column. Notice that ‘asserting’, which

Table 1. D'Andrade and Wish's plotting of speech acts on four dimensions

Bipolar rating scales	Speech act types	Bipolar rating scales	Speech act types
More friendly	S reacts to H with approval	More dominant	S makes a forceful assertion
	S makes a forceless request		S makes a forceful request
	S requests H's attention		S makes a forceless assertion
	S disagrees		S answers a question
	S refuses		S agrees
More hostile	S reacts to H with disapproval	More submissive	S gives H his attention
More task-oriented	S requests a commitment from H	More emotional	S evaluates self negatively
	S commits or complies		S expresses an attitude (includes many types)
	S requests direct action		S makes an assertion that is not a reaction to H
	S requests agreement		S makes an assertion or poses a question
	S reports an event		S makes a judgmental assertion
Less task-oriented	S evaluates self positively	Less emotional	S evaluates a third party negatively

is not typically considered to have social effects, is rated highly for dominance, especially when the act is performed forcefully. Notice also that ‘congratulating,’ ‘thanking,’ and ‘apologizing,’ which are typically considered to have social effects, are not listed. The authors explain that such expressions of a speaker's feelings are highly conventionalized and that their research shows that such forms occur much less frequently in ordinary interactions than expressions of approval and disapproval. As a consequence of this finding, the authors collected all conventionalized expressions of attitude into a single category, appearing in the table as the second most emotionally-laden act type.

Although Politeness Theory and Interpersonal Circumplex Theory have seemingly little to say to each other, at least at first glance, several studies have noted how they can be useful together for analysis. For example, Oakman, Gifford, and Chlebowsky [2003] point out:

Politeness Theory and IPC Theory have some remarkable similarities. Positive face is clearly related to the warmth dimension of the IPC. Positive face-work strategies (e.g., use in-group slang, use nicknames, compliment, joke, amplify agreement) are designed to communicate affiliation. Negative face seems related, though perhaps not so obviously, to dominance concerns. It seems likely that a person's dominance status could influence whether his or her desire to act autonomously and without interference is gratified. Moreover, Brown and Levinson (1987) describe behaviors oriented to the maintenance of negative face (e.g., apologize, use honorifics, minimize the imposition) as the core of deference behavior, which is clearly associated with dominance relations.

It might be possible to array positive and negative face-work strategies on the IPC axes of dominance and warmth. Face-work strategies can be exploited to convey hostility as well as warmth, and dominance as well as deference. One can “do the opposite” (for example, use insults instead of in-group markers), or notably fail to perform relevant face-work (for example, refuse a request in a blunt, bald-on-record fashion). A lack of

expectable warmth, then, can convey hostility, and a lack of expectable deference behavior can convey dominance.

It seems that if one accepts a dimensional analysis of interaction, where dominance and affiliation are the primary dimensions, then it should be possible to use PT and IPC together to make predictions about behavior at both a macro-level and at a detailed level: At the detailed-level, one could predict how a speaker would handle a potentially face-threatening (or face-enhancing) situation, and at the macro-level, what kind of response behaviors the audience would enact.

The dimensions of power and social distance appear frequently in the literature not just in analyses of *behavior patterns* but in analyses of *social relationships* as well [Brown & Gilman 1972; Svennevig 2000] (although Haslam [1994] argues for categories over dimensions). This is not too surprising, if one views relationships as being defined primarily by recurring patterns of behavior that are specific to that type of relationship. Svennevig makes a connection to Politeness Theory as well, noting like Brown and Levinson do that the degree of threat of any speech act depends on the relationship of the participants, but going further to operationalize the relationship dependency through four dimensions: power, plus a breakdown of social distance into three dimensions: solidarity, familiarity, and affect. Svennevig's dimensional model of relationships is the basis of Bickmore's AI agent described in the next section.

Artificial Intelligence

Speech Act Planning

There is a long history of research in AI on strategic interpersonal communication, although the social aspects of such communication have been largely neglected until recently. Speech act planning (SAP), in particular, has modeled how an agent with a goal to influence another agent can use knowledge of how different speech acts affect different hearer mental states, and thereby select acts whose effects match the goal. For example, the following construct is a “plan operator” that connects putative effects of requesting (i.e., inducing the hearer of the request to intend to do the requested action) to putative conditions that must be present for the request to succeed (i.e., the speaker of the request must actually want the hearer to do the requested action, and the speaker must sincerely believe that the hearer is capable of doing the requested action):

Name: Requesting(Speaker, Hearer, ActionRequested)

Effects: intendsToDo(Hearer, ActionRequested)

Conditions: hasGoal(Speaker, intendsToDo(Hearer, ActionRequested)) and
believes(Speaker, isCapableOfDoing(Hearer, ActionRequested))

This style of analysis is virtually identical to that of Speech Act Theory. For example, the conditions for requesting here focus on what it means to perform the act sincerely, just as the conditions listed for promising by Wierzbicka in Figure 3 do. However, SAT representations of act effects are generally limited to impacts on the audience at the moment of *uptake*, while SAP representations usually represent effects on audience *intentions*, further along a causal chain. Capturing effects on audience intentions allows

SAP to simulate conversations with multiple turns, such as a pair of people working together on a task [Grosz & Sidner 1986], since the operators represent how the goals of the participants change during the interaction.

Plan operators originated in classical AI planning, where a typical planning algorithm would take a propositional representation of a goal as input, then search for plan operators having an effect that could be matched to the goal. If a plan operator were found to match, there would be another search for known facts about the current situation that match the conditions of the operator. Any unmet conditions would be treated as sub-goals and the process would proceed recursively. Speech act planning adopts this algorithm largely without question, in order to focus on the operators – the objective is to capture the tacit knowledge needed to communicate in elementary ways such as informing and requesting. Such tacit knowledge turns out to be fairly difficult to formulate (which is a frequently-learned lesson in almost all domains that the subfield of Knowledge Representation in AI attempts to tackle). For example, the operator above represents the condition of sincerity as simple belief, even though sincere requests have other implications, such as not knowingly leading the hearer into danger.

Similar to the narrow coverage of act types in SAP, its analysis of act conditions has also been slim, generally focusing on whether the act was sincere (under the assumption that acts that are sincere also appear sincere, and that apparently sincere assertions are always convincing).

These early systems [Cohen & Perrault 1979; Allen et al 1994] addressed only the goal of persuading someone to do some act, where the agents were assumed to be so naturally cooperative that all one needed to do was inform someone that one needed something and that agent would comply without hesitation. The set of speech act types modeled contained just two kinds: informing and requesting. A notable exception is the early work of Bertram Bruce. Even though Bertram Bruce's work on speech act planning [Bruce 1975a, 1975b] preceded that of Perrault and Allen, it covered a wider range of speech act types, including such socially relevant acts as warning, threatening, and thanking. Interestingly, despite including such acts, and mentioning that some acts affect emotions, Bruce still represented only the effects of these acts on beliefs and goals. Bruce claimed that people use at least three different plan types to generate acts: goal-seeking, fulfilling an obligation, and acting on a habit-like predisposition. The relative richness of Bruce's theory may be due to the fact that it was not developed for a narrow task such as simply providing information; instead, it was meant as a framework for modeling all kinds of conversations. For example, he used his representations of acts and plan types to develop schemas of typical conversational exchanges, such as the tendency of requests to be followed by promises, suggestions, or refusals. Some schemas were even specialized to exchanges between particular roles, such as between teachers and pupils. But Bruce's work seems not to have been carried forward by others.

There is much less work in AI based on Politeness Theory compared to Speech Act Theory; in fact, we know of only one such project. Nancy Green implemented the first three strategies of PT for dealing with face-threatening acts as part of a computational

system for generating explanations that are both informative and appropriate [Green 1994]. Green does not explicitly represent face or levels of threat to drive her system's choice of strategy. Instead, her strategies are triggered directly by conflicts between the recipient's assumed goals and the system's planned acts. For example, Green's system simulated the role of an academic colleague to the user and generated the response (R) shown below when given a question from the user (Q) [Ibid, p. 81]:

Q: Did you manage to read that section I gave you?

R: I've read the first couple of pages.

This response follows the positive face politeness strategy (i.e., strategy #2) because it indicates that the system/colleague respected the user's wishes enough to make a start at his request, even though it has not finished the job yet. The strategy to attempt to appease the user was triggered by a pair of facts: (1) the system believed that the user had wanted some job to be done that was not done yet (and the system was responsible for doing it), and (2) the system believed that the user wanted something else done, and it had been done. The completed job is used as the focus of the appeasement strategy; that is, the speaker responds by talking about what he has done rather than about what he has not done.

There is also very recent work on generating and detecting emotion in text (see Belz [2003] for a review) and in speech [Pentland 2005], and on ascribing credit or blame [Mao & Gratch 2003]. But almost all the work on planning speech acts has relied on

certain assumptions about planning in general, seemingly rooted in its blocks-world origins:

- An act either completely satisfies a goal (or subgoal), completely thwarts it, or is not relevant to the goal;
- Solving a set of goals means finding a set of acts that satisfy each goal without interfering with each other's contributions to the goals.

These assumptions underlie all the modern flavors of classical AI planning: planners that permit hierarchies of acts, that minimize commitment through partial ordering of acts in solutions, and that protect act contributions through protecting causal links. But these assumptions do not have a good fit with social communication, where it is rarely true that any act can completely satisfy or thwart any goal. For example, consider one's range of options when asking for a date -- no approach is sure of success, and each usually involves a partial promise of success mixed with risk of partial or complete failure. It turns out that the classical style of planning does not fit well in many physical domains, either; consider the problem of a Mars rover trying to work its way out of a sand pit -- probably no option is clearly the correct one, and every alternative carries promise and risk in partial measure.

But the planning approach is not completely misguided. In fact, its influence on theories in communication and social psychology has been growing steadily since the 80s [Cody 1986; Daly & Wiemann 1994; Berger 1997]. Early studies tended to produce theories that correlated a set of dimensions such as personality or behavior traits [Kiesler 1983], or features of the situation such as relative social status. The dimensional nature of these

theories seems to be a factor of both the kind of analytic instruments available (e.g., statistical analysis) and that there was no dominant theoretical framework. Researchers in these fields have noted that comparing such theories and building on them has been difficult, but that an information-processing, goal-directed framework seems more fruitful [Cody 1986].

The influence of AI planning concepts on communication and social psychology came full circle in two AI projects, LetterGen [Pautler 1997; Pautler & Quilici 1998] and REA [Bickmore 2003], that attempted to synthesize goal-directed frameworks from those other fields into richer, more true-to-life simulations. LetterGen gave advice on how to phrase socially-appropriate letters and e-mail by asking the user about his interpersonal goals and by proposing candidate plans for those goals. It simulated the effects of each of 40 different types of speech acts on the beliefs, emotions, and goals of the hearer, as well as the relationship between the hearer and speaker. It would plan by matching speaker goals against these effects and tracing the causal chains backward to acts, which became the acts of its plans. It would also trace forward from its selected acts looking for effects that would threaten its goals, and plan mitigating acts. The causal chains were synthesized from many different lines of research in the social sciences, and covered many kinds of reasoning (such as attributing motivations and monitoring obligations), which made it difficult to evaluate as a cohesive framework. Also, there was no empirical comparison with human subjects, so it was left to the reader to judge the plausibility of the plans it generated. A representative set of LetterGen's causal chains is illustrated in Figure 5.

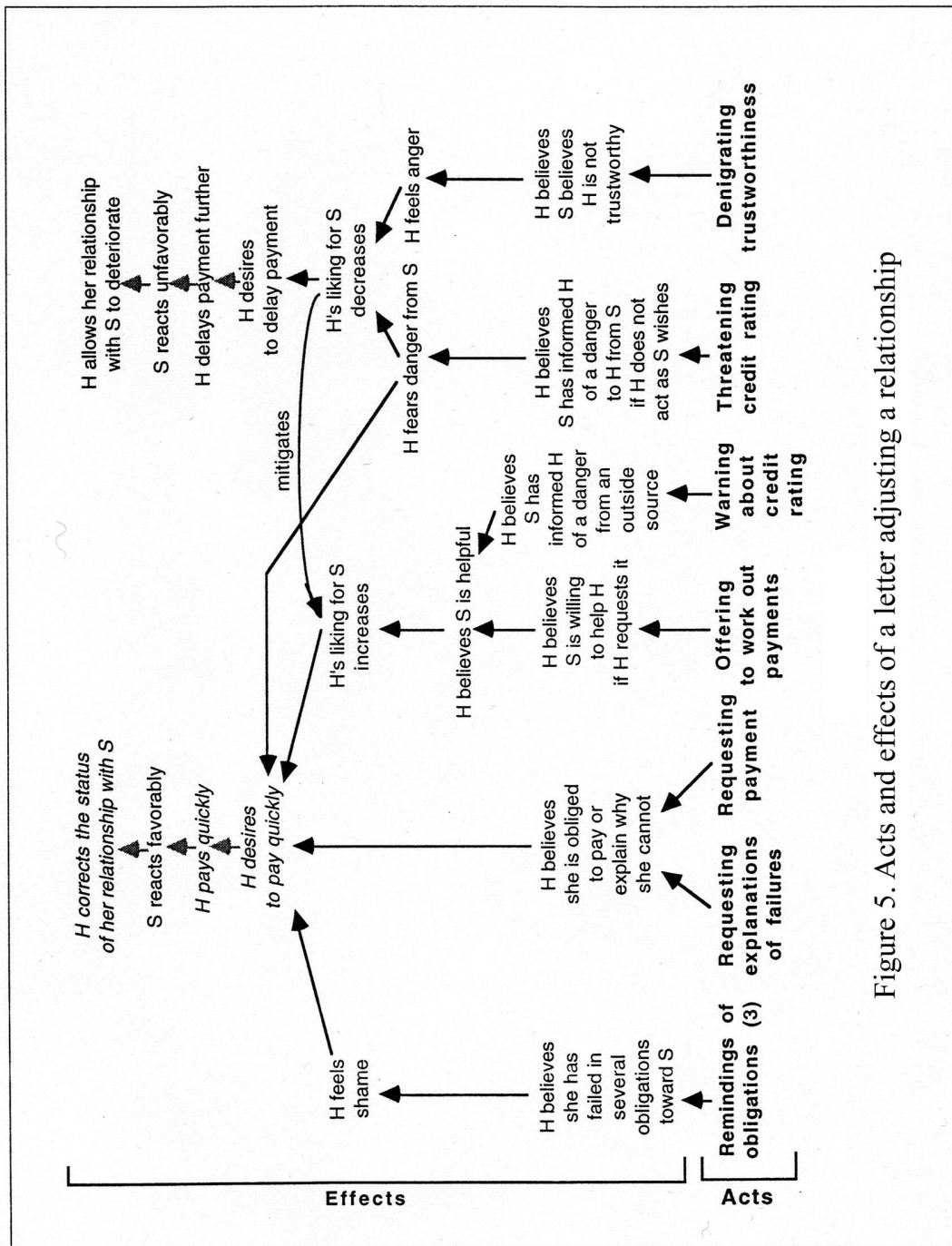


Figure 5. Acts and effects of a letter adjusting a relationship

Bickmore's work took the synthesizing approach further, using a wider range of social science research to motivate representing relationship strength in four dimensions instead of one (using Svennevig's set of dimensions), and proposing mathematical formulae for evaluating candidate plans for their ability to change relational dimensions by desired amounts [Figure 6]. This reasoning engine drove the behavior of 3D animated characters who interacted with the user through speech and gesture- and gaze-tracking [Figure 7]. Bickmore tested his REA agent with human users and found them to be slightly successful in fostering trust (a major part of one of his four dimensions), albeit under narrow conditions contrary to the initial hypotheses, and where it turned out to be difficult to explain the contrary findings. The large architecture involved (i.e., 3D animation, speech recognition and synthesis, real-time synchronization, etc.) makes it especially hard to attribute credit or blame for these outcomes. In particular, it is difficult to determine from his study results whether his set of dimensions or his formulae could be improved.

More formally, given the set of agent goals $G = \{g_1, g_2, \dots\}$, the degree of satisfaction, $S_G: G \rightarrow 0..1$, and priority, $P_G: G \rightarrow 0..1$, for each goal, each move is assigned the following activation energy during each update cycle.

$$\alpha_i^0 = 0$$

$$\alpha_i^t = \text{decay}(\alpha_i^{t-1}) +$$

$$\begin{aligned} & E_{GOAL}(m_i) * G_{GOAL} + E_{ENABLED}(m_i) * G_{ENABLED} + E_{FORWARD}(m_i) * G_{FORWARD} + E_{BACKWARD}(m_i) * G_{BACKWARD} + \\ & E_{RELEVANCE}(m_i) * G_{RELEVANCE} + E_{FACE THREAT}(m_i) * G_{FACE THREAT} + \\ & E_{TOPIC ENABLED}(m_i) * G_{TOPIC ENABLED} + E_{SPENABLE}(m_i) * G_{SPENABLE} \end{aligned}$$

Figure 6. Bickmore's equations for calculating activation energy in the spreading activation network used to generate communication plans



Figure 7. A person interacting with Bickmore's real estate agent avatar, REA, via speech, gaze- and gesture-tracking, and 3D visualization

Multi-Agent Systems

Research on *multi*-agent systems (MASs) has paid much less attention to the human example of appropriate social interaction than has classical work in AI or work on speech acts (and Bickmore's agent interacts with people rather than with other agents). In most MA systems, the user provides a goal and the primary job of every agent is to find a way to cooperate with the others in order to achieve that goal. Consequently, the dominant view of what it means in MAS for an agent's behavior to be socially appropriate is that the behavior be optimal with respect to the agent's cooperation with other agents [Rosenschein & Zlotkin 1994; Shoham & Tennenholtz 1992]. This notion is very similar to Grice's maxims for cooperative conversational contributions; it is the notion that cooperation should be pursued for its own sake, regardless of how it affects individuals. Castelfranchi [1992] argued that cooperation-for-its-own-sake is likely to be a bad strategy for both MAS and speech act systems as such systems are applied to (more realistic) tasks in which each agent must behave in a more human-like manner. He gives an example in which two agents do not agree about how they should cooperate and each must argue to try to convince the other. Such conflicts are even more likely when different collections of agents with different goals come into contact.

Sociable Agents

On another strand of work with computerized agents, they have been designed to be social in a more recognizably human way, either as simulation tools like ours or to provide entertainment. For example, the SIMS™ game by Electronic Arts allows human players to mix with agents (a.k.a. Non-Player Characters (NPCs)) by dancing, flirting,

partying, and so on. There are similar agents in AI research on “interactive drama”, storytelling systems, and even sociable robots:

- **Kismet** [Breazeal 2002] is a robot that monitors the gaze, proximity, and vocal tone of a human interactant (when seated face-to-face) and reacts through head posture, facial expressions, and simple vocalizations. It encompasses a full action cycle by interpreting human actions in terms of whether they seem threatening, entertaining, and so on, by weighing these interpretations against its drive state (e.g., loneliness and sleepiness) and emotional state (e.g., joy and fear), then adjusting those states, and then expressing those states through action. While Kismet’s behavior is quite rich, it makes evaluation of the planning component difficult because it is not clear how to assign credit to any of its components, especially given its charming appearance and animation [Figure 8].
- **The SIMs™**, as already mentioned, uses agents that engage in a variety of social activities. Like Kismet, each agent has a drive state and an emotional state. Agents select behaviors by taking the state most in need of satisfaction (e.g., hunger) and looking in the immediate environment for any object or agent that could best satisfy that need [Forbus 2002] [Figure 9]. For example, agent Dwayne might be somewhat lonely but very hungry, and see both a refrigerator and agent Tanya nearby; objects in the system have name-value pairs representing their utility for a variety of drives, and refrigerators are represented to be much more useful for hunger than fellow agents. So, Dwayne would approach the refrigerator first, but once his hunger drive dropped in intensity below that of his loneliness, he would begin to approach Tanya. (Presumably, Tanya is not offended that

Dwayne approached the refrigerator first, or Dwayne is not capable of modeling other agents.) The extent to which such agent behaviors appear human-like and realistic to game players seems to depend a great deal on the drive intensities and object utilities being “in tune”; for example, if loneliness never exceeded hunger, or if female agents were no more entertaining to male agents than other objects, the game would fail. We cannot be sure how the weightings in the game were determined, since it is proprietary, but we suspect it was through a great deal of trial-and-error since we are not aware of any relevant data from human participant studies. (Kismet’s weightings appear to be non-empirical as well.)

- **ChimpWorld [Hughes 1992]** and **Affective Reasoner [Elliott 1992]** are research projects that aim to be “wind tunnels for the social sciences”, as Hughes puts it. ChimpWorld’s agents are simulated chimps who live in a simulated savannah a few yards square containing food and other resources. Affective Reasoner’s agents are customers and taxi drivers in a simulated TaxiWorld. Agents in both systems construe events and react to them in the emotion-laden manner one would expect in such domains. And both systems use empirically-based causal rules connecting drive and emotion states to actions, but they leave open the weightings of states in each agent, and the initial scenario, to the user. This openness is an invitation to experiment. Our simulation is also an invitation to experiment, but we aim to start with causal rules and state weightings that are both empirically-derived and leave open just the initial scenario (although we expect experimenting might show that the rules or weightings might need more empirical testing).

- **Oz Project** [Reilly and Bates 1992] is an “interactive drama” project and a forbearer of multi-agent social games such as The SIMs™. The “drama” emerges from the agents following their own simulated drives and emotions, which lead to action which sometimes involves other agents. (The basis for the emotion model is the work of Ortony, Clore, and Collins [1988], which is also the basis for the model in Affective Reasoner.) There is an optional “drama manager” component that attempts to guide agents toward more interesting interactions. In the Oz Project (and in Affective Reasoner), the agents not only model emotional reactions to their own surroundings but also reason about the likely emotional reactions of other agents. Furthermore, the agents represent beliefs, goals, and changes in personal relationships as these relate to emotions. Because these agents have some knowledge of how their particular acts may affect the emotions and relationship attitudes of other agents, they have some idea of how to make their actions polite and not merely generally cooperative. The agents also know how to be impolite in order to adjust or terminate a cordial relationship. One implementation used animated blob-shaped characters called Woggles who could express themselves through eye movements, by stretching their bodies or hopping, and through text bubbles. Another implementation, Lyotard, was a running stream of text describing the action.
- **MEXICA** [Pérez y Pérez 1999] is one of the most recent projects in the subfield of story generators, which stretches back at least as far as Meehan’s TALESPIN [Meehan 1976]. Like the text-streaming presentation of action in the Oz Project’s Lyotard, MEXICA and TALESPIN both use agents to embody characters in their

dramas. And like all the social agent systems discussed so far, MEXICA uses weightings to tune its plans toward greater realism; for example, one rule is: *“If a warrior is ill or wounded and a lady cures him, it is natural that he will develop very positive emotions towards the lady. In this way, when defining the action A CURED B, post-conditions might include that B develops an **Emotional Link** of type 1 **intensity** +3 towards A.”* [Ibid., page 54, emphasis mine] Although the weightings seem to play an important role in making agent behavior realistic, the weightings, like those in the other systems so far, are not based on studies with human participants.

- **Homunculus** [Gullahorn and Gullahorn 1963] is the grandfather of all sociable agent research, as far as we can tell. It was an AI project that implemented some of the theories of Social Exchange psychology, which focuses on the expected utility of asking for a favor or giving help and the “social debt” one may accrue or pay off by doing so. Homunculus was designed to simulate how one participant of an interaction might reason about social costs and benefits as a way of planning his actions. For example, one simulated interaction involved two coworkers where one, Ted, asked the other, George, for help with a project. Homunculus simulated both agents. George's cost consideration in such a case is that he would lose time for his own project if he helped Ted. Ted's cost consideration is that he is allowing himself to be dependent on George, thereby reducing his social standing relative to George. George must weigh his loss of time against a gain in status relative to Ted, and Ted must weigh his loss of standing against receiving help. Each agent weighs these costs and benefits using a combination of three factors.

These factors are: (1) how frequently has the agent been rewarded for such a favor in the past? (2) how intrinsically valuable to him is the typical reward? and (3) how much does he need this reward at the moment? In this example simulation, the "Ted" agent decided to accept a loss of status and ask the "George" agent for assistance. However, George decided that helping directly would take away too much of his time, but he was still able to gain some increase in status by recommending another agent to help Ted. The factors used to make these decisions relied on representations of other agents' tendencies to be helpful or difficult, the degree to which the other agents were already in debt, and the standing of the agents in certain social groups. Homunculus also simulated how a person might feel after helping and then getting nothing in return (i.e., anger), or how he might feel after being helped and then not returning the favor (i.e., guilt).

Such systems lie outside the dominant cooperative paradigm in multi-agent systems and evince much more of the self-motivated behavior that Castelfranchi argues for. They share similar architectures as well (but judging from their citations, the similarity appears to be accidental): Drive/urge states such as hunger and loneliness drive emotion states which in turn drive behaviors. Some behaviors have immediate effects on drives, such as eating and hunger, while others depend on the participation of other agents or humans, such as flirting and loneliness. But as mentioned already, much of what makes the agents' behavior realistic seems dependent on the weightings given to emotion strength, the urgency of a drive, and the impact of actions on drives and emotions. Because the weightings in these systems appear to be derived from trial-and-error tuning, some of

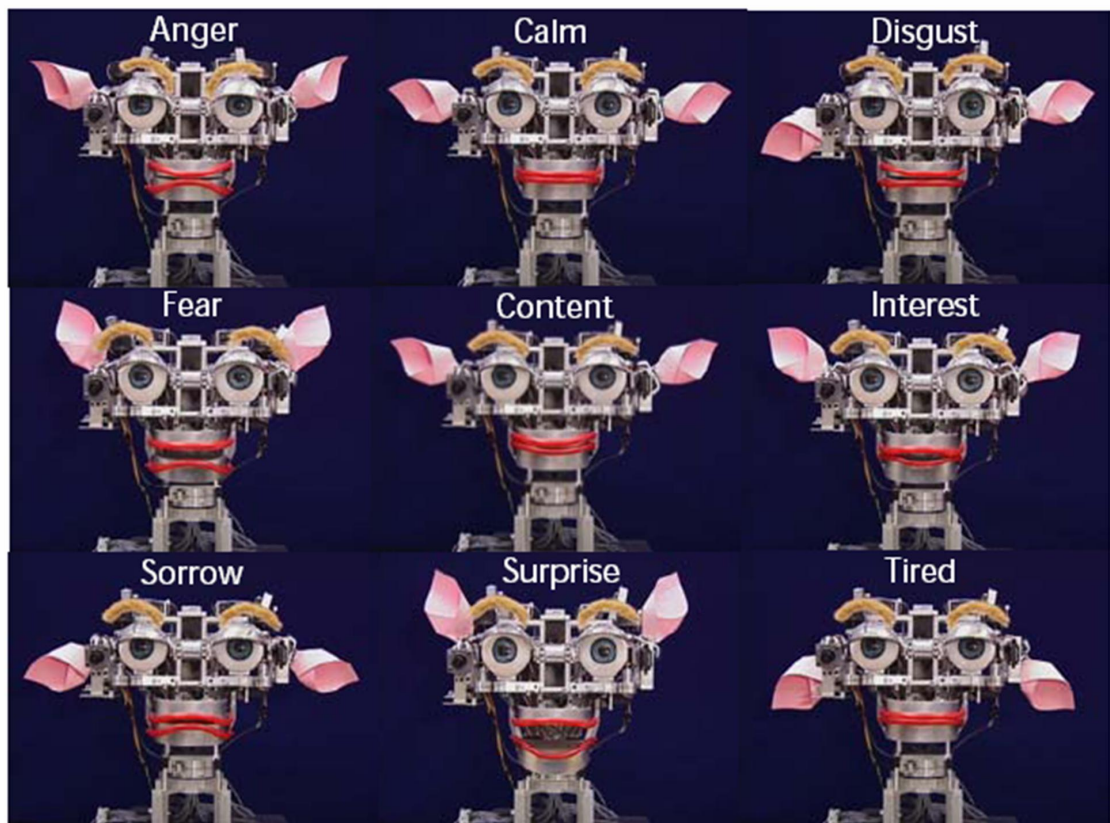


Figure 8. Nine of the facial expressions that Breazeal's Kismet robot is capable of

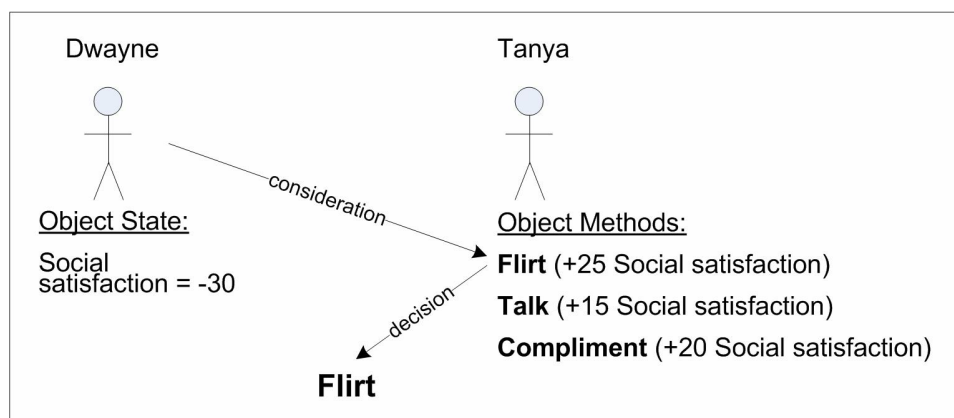


Figure 9. Action selection in The SIMs™

(according to game creator Will Wright via Ken Forbus' online lecture notes)

them may be very reliable and may even be good indicators of how similar states would be weighted by humans. But very few weightings from any of the systems have been published. Thus, one could make a contribution by publishing one's weightings and making them an important part of the discussion.

Regarding the trend toward interactive drama and believable game characters, John Lasseter, the director of *Toy Story*, argues that if animated characters expressed emotions (and personality traits) in the same way humans do in everyday situations, the action would not be very entertaining because typical emotional reactions to realistic scenarios are too subtly expressed. He argues that communicating the inner world of a character to an audience requires some form of exaggeration [Lasseter 1987]. Although the trend in sociable agents is to “embody” them by providing a graphical appearance, voice, and so on, there is a danger that either user expectations will be raised too high for the relatively “flat” personality of agents to succeed, or agent personalities will have to be exaggerated, which complicates the evaluation of the agent because it becomes hard to assign credit to either the social planning component or the exaggerated persona.

Communication

In communication, one finds yet another approach to understanding social communication: strategy taxonomies. Generally, each study in this literature starts with one of a handful of high-level goal categories and attempts to list comprehensively every strategy for pursuing that high-level goal that the study found in some corpus. There is

consensus about the set of high-level categories – they are the social goals of seeking affiliation with others, maintaining affiliation, avoiding conflict, and gaining compliance – but less agreement about what should count as a lower-level category.

The edited collection *Strategic interpersonal communication* [Daly and Wiemann 1994] contains studies for each of the high-level categories already mentioned. In one of the studies described in the collection [pp. 118-119], the communication scholars Bell and Daly identified 25 strategies for affiliation-seeking, and Table 2 reproduces the first 10 of these. Some of these strategies resemble strategies already mentioned. Others, such as facilitating enjoyment, have not been mentioned at all. The conversational rule-keeping strategy actually combines the often conflicting substrategies of being cooperative (e.g., Grice's maxims) and being polite (e.g., Brown and Levinson's face-saving strategies).

Similar to the taxonomies of speech acts, these taxonomies of strategies cover a wide range of social behaviors and provide insight into the rich variety of those behaviors. But these taxonomies do not provide details about how goals lead to action, nor how such plans might depend on the situation. Furthermore, sometimes a taxonomy resulting from one study differs significantly from a taxonomy for the same high-level goal described in another study, and there is no agreed-upon method for deciding among them [Kellermann & Cole 1994].

Despite the limitations of taxonomy-based approaches already described, we believe that such taxonomies play a vital role in mapping out, albeit roughly, the range of social

Table 2. 10 of Bell and Daly's 25 affiliation-seeking communication strategies

Strategy	Definition
Altruism	The affinity seeker strives to be of assistance to the target in whatever he is currently doing.
Assume Control	The affinity seeker presents himself as a person who has control over whatever is going on.
Assume Equality	The affinity seeker strikes a posture of social equality with the target.
Comfortable Self	The affinity seeker acts comfortable and relaxed in settings shared with the target.
Concede Control	The affinity seeker allows the target to assume control over relational activities.
Conversational Rule-Keeping	The affinity seeker adheres closely to cultural rules for polite, cooperative interaction with the target.
Dynamism	The affinity seeker presents himself as an active, enthusiastic person.
Elicit Other's Disclosures	The affinity seeker encourages the target to talk by reinforcing the target's conversational contributions.
Facilitates Enjoyment	The affinity seeker tries to maximize the positiveness of relational encounters with the target.
Inclusion of Other	The affinity seeker includes the target in the affinity seeker's social groups

knowledge that a simulation should aim for. In particular, we envision the taxonomies of the preceding theories fitting together in a specific way, suggesting how they might be synthesized into a comprehensive framework. In our view, these theories belong to four categories -- interactional, motivational, strategic, and tactical -- ranged along a single scale representing wide scope with limited concreteness at the “top” and narrow scope with greater concreteness at the “bottom” [Table 3]. Regarding synthesis, we have discussed several studies that propose ways of connecting a theory in one category with a theory in an adjacent category.

Although it may be possible to fit together taxonomic approaches from several theories -- even from different disciplines such as we propose above -- nevertheless, any taxonomy

Table 3. A proposed way of envisioning relationships among theories of social behavior

Generalization Level	Components of Theories of Social Behavior			
<i>Interactional</i>	Dimensions of behavior and of relationships (Interpersonal Circumplex Theory; Svennevig)			
<i>Motivational</i>	Complementarity Principle (IPC)	Motivation to change the balance of dimensional loadings in a desired direction (Svennevig)	Motivation to avoid face threats (maintain status quo in relationship) (Politeness Theory)	Motivations to gain compliance, seek affiliation, etc
<i>Strategic</i>			The 4 strategies for redressing face threats	Strategy taxonomies for each motivation above
<i>Tactical</i>			Speech acts	

has limited appeal if there is no implied sense of how it and its members would actually be evaluated and used in ordinary interactions. Cody [1986] made a similar point when he pointed out that it is difficult to evaluate taxonomic approaches and that an information-processing, goal-directed framework seems more fruitful. We know of just one line of research in communication that adopts this tack, Kellermann's Conversational Constraint Theory (CCT) [Kellermann 2004, 2006; Kellermann & Park 2001; Kellermann & Lee 2001a].

The aim of CCT is to describe the decision process whereby communicative acts such as hinting, changing body posture, and making eye contact are weighed for their suitability for an overarching interpersonal goal and for the situation so that the goal is met in a socially acceptable way. CCT focuses on "compliance-gaining" interactions, where one person tries to persuade another to adopt a particular goal.

According to CCT, every potential act can be weighed in terms of its likely efficiency and its likely politeness for the goal and situation. To decide whether an act is worth using, one would compare its efficiency score against a minimum preferred level of efficiency, while comparing its politeness score against a minimum preferred level of politeness; only acts that score above *both* levels are *acceptable*. Furthermore, CCT claims that level values vary over a variety of situational, relational, personal, and interaction-specific factors (see Appendix B for the list, plus definitions).

Although the aim of CCT is to describe a decision process for selecting a communication act for a goal, it does not spell out an algorithm for doing so; yet, one can be inferred from its discussion of scores, situation factors, and thresholds. This process is illustrated in Figures 10 through 12. Figure 10 shows how the chosen goal results in a set of candidate acts, and Figure 11 shows how situation factors determine thresholds on the scores of candidate acts; CCT does not suggest whether these two subprocesses are done in parallel nor in a specific order, but neither depends on the other so it does not seem to matter. For the subprocess of determining thresholds, one checks each of the situation factors and adjusts one or both thresholds by a delta value; for example, if the person one is speaking to is in a higher social position, one would decrement the efficiency threshold and increment the appropriateness threshold, while if they were not in a higher position, one would do the opposite and increment efficiency and decrement appropriateness. Figure 12 shows how the thresholds from Figure 11 are overlaid on the candidate acts and their scores from Figure 10 to reveal the subset of candidates having scores passing

Choose a goal. Ex: Stop Annoying Habit

Note the scores that different speech acts have for this goal on both scales.

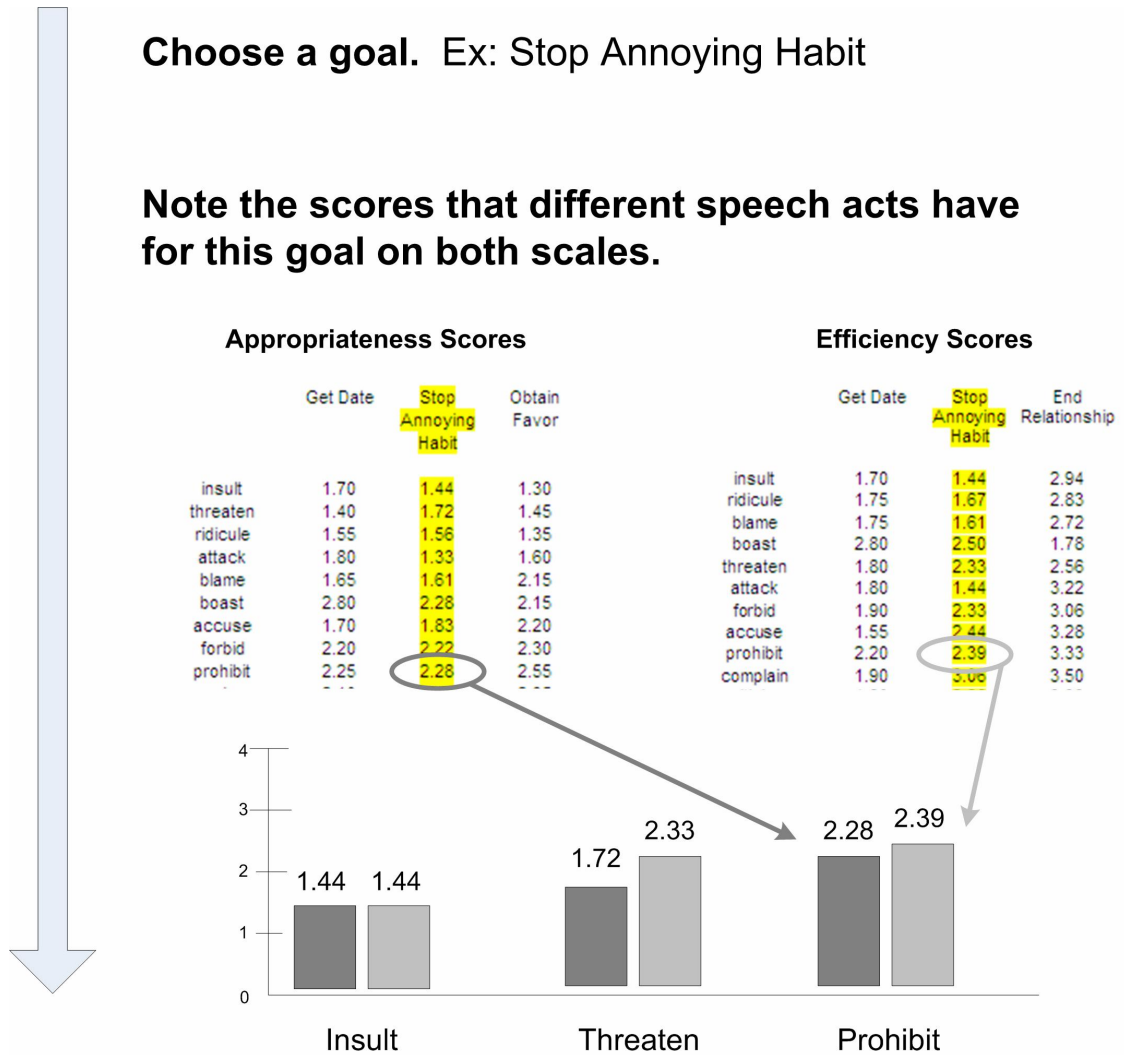


Figure 10. How CCT determines candidate acts from the chosen goal

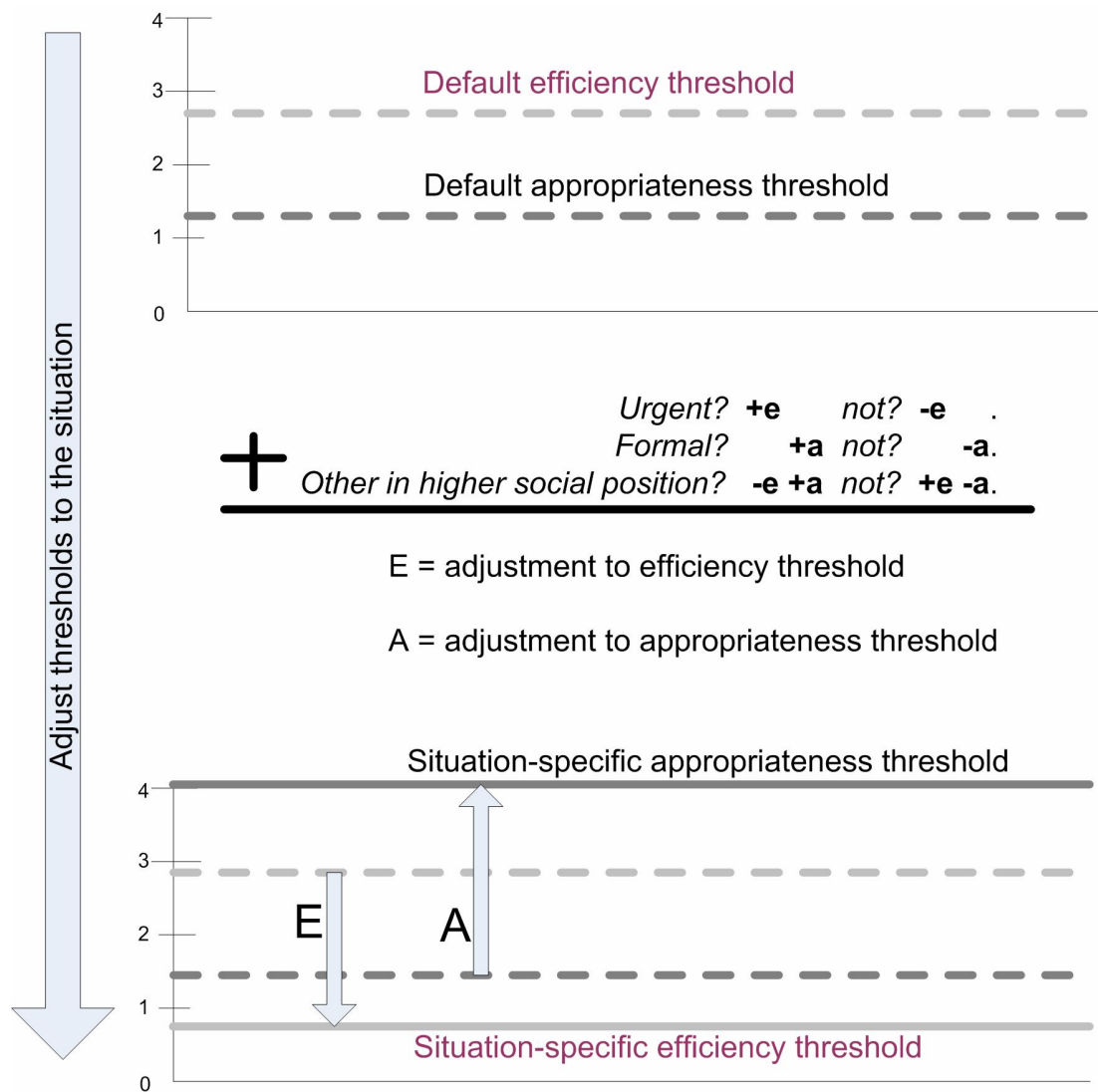


Figure 11. How CCT adjusts thresholds based on situation factors

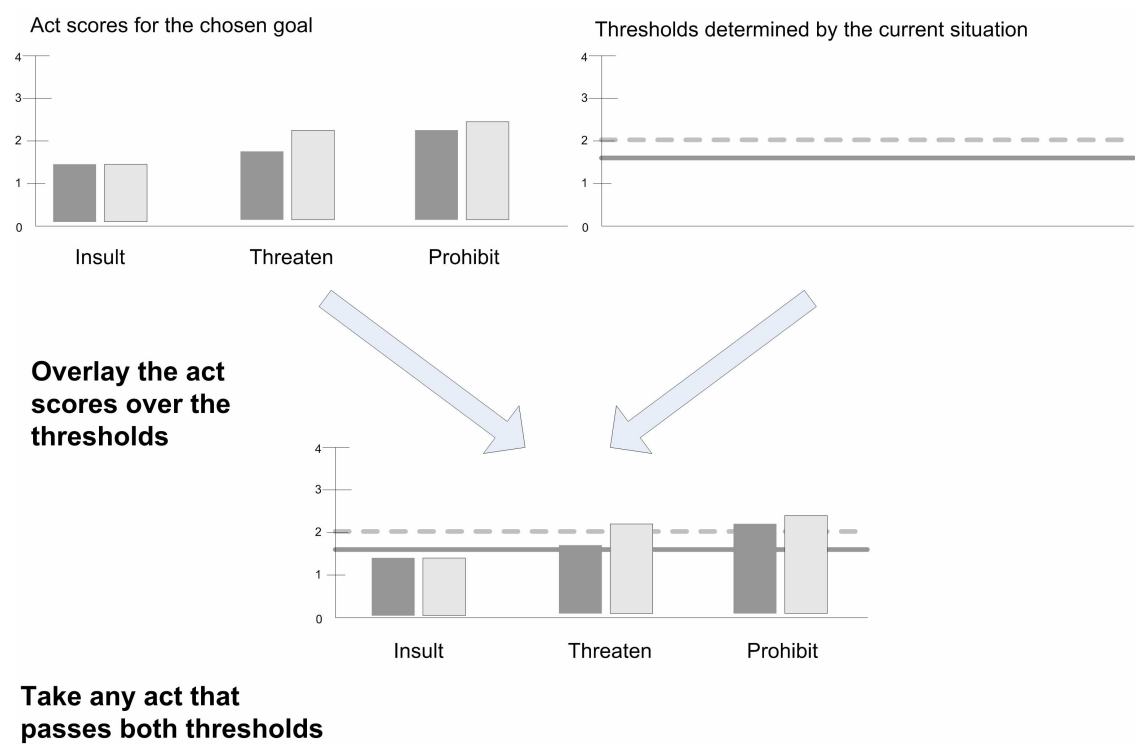


Figure 12. How CCT determines which acts are acceptable for the chosen goal in the current situation, by overlaying thresholds on the scores of candidate acts

both thresholds; this subset represents all the acts considered acceptable for use in the situation for the chosen goal.

Kellermann uses the term “politeness” interchangeably with “social appropriateness,” and in the questionnaire of one of her studies [Kellermann & Lee 2001b], she defined “politeness” this way for participants:

Here, we want you to tell us how polite you think different tactics are for resisting another’s relational escalation attempts. A polite tactic for resisting relational escalation is socially appropriate; it is pleasant, proper, considerate, and mannerly in resisting relational escalation. A polite tactic is nice. By contrast, an impolite tactic for resisting relational escalation is socially inappropriate; it is discourteous, ill-mannered, uncivil, and/or nasty. An impolite tactic is rude.

A Polite Tactic	An Impolite Tactic
Is pleasant and civil	Is nasty and uncivil
Is proper and considerate	Is improper and inconsiderate
Is courteous and appropriate	Is ill-mannered and inappropriate
Is nice	Is rude

In the same study questionnaire that was quoted above for its definition of “politeness,” the term “efficiency” is defined this way:

Here, we want you to tell us how efficient you think different tactics are for resisting another’s relational escalation attempts. An efficient tactic for resisting relational escalation is immediate and to the point; it does not waste time, energy, effort, or steps in resisting relational escalation. An efficient tactic is expedient. By contrast, an inefficient tactic for resisting relational escalation takes time, energy, and/or effort; it is an indirect and roundabout way of resisting relational escalation. An inefficient tactic is wasteful.

An Efficient Tactic	An Inefficient Tactic
Is expedient and doesn't take a lot of time	Is wasteful and takes time
Avoids extra steps and is direct	Employs extra steps and is indirect
Isn't too effortful and is immediate	Takes effort and is not immediate
Doesn't waste energy and is to the point	Wastes energy and is round-a-bout

The notion of efficiency used by Kellermann is very similar to the common notion, which might be operationalized as, “the extent to which one's action can be expected to be effective (i.e. reach the goal), factored by the effort expended through the action.”

To determine the set of goals and acts that would be governed by her decision process, Kellermann surveyed the communication literature on compliance-gaining strategies, finding 13 goal types to be representative [Table 4]. To create the set of communicative tactics, Kellermann compared tactics uncovered in her survey with entries in Wierzbicka's dictionary of English speech act verbs [Wierzbicka 1987]. Although there are several hundred entries in this dictionary, Wierzbicka grouped them into a few dozen

clusters, each with prototypical members. Kellermann's comparison revealed a high degree of overlap between tactics found in her survey and the prototypical members of Wierzbicka's groups. The comparison resulted in a list of 56 speech acts [Table 5] that are apparently not only representative of persuasion tactics but also of speech acts in general.

Table 4. Persuasion Goals Used in the 2004 CCT study

(Most items below begin implicitly with “Persuade to...”)

Get Date	Share Time Together	Fulfill Obligation
Stop Annoying Habit	Initiate Relationship	Obtain Permission
Obtain Favor	Get Advice	Provide Guidance
End Relationship	Move Relationship Forward	
Change Opinion	Obtain Information	

Table 5. Speech act types used in the 2004 CCT study

Accuse	Assure	Compliment	Explain	Joke	Prohibit	Ridicule
Acknowledge	Attack	Confess	Forbid	Justify	Promise	Suggest
Advise	Blame	Confirm	Forgive	Offer	Protest	Summarize
Apologize	Boast	Criticize	Give	Order	Question	Tell
Approve	Challenge	Demand	Hint	Permit	Remark	Thank
Argue	Claim	Disagree	Inform	Plead	Report	Threaten
Ask	Comment	Disclose	Insist	Point out	Reprimand	Vow
Assert	Complain	Excuse	Insult	Praise	Request	Warn

Obviously, there are many degrees of freedom in CCT that allow for nuanced ways of fitting the model to observed behavior; the question is whether any particular configuration of its parts can actually match human behavior, or should it be structured another way. Several studies were done to determine if CCT could be “tuned” to match human behavior [Kellermann 2004; Kellermann & Park 2001], and they provided fairly strong support.

One of CCT's innovations is that it does not assume that efficiency and politeness must necessarily be at odds with each other. The assumption that in order to be polite, one must accept less efficiency has been widely accepted in communication since the pioneering work on politeness done by Brown and Levinson [1987].

While CCT has largely proven its claims, and while its information-processing approach is very similar to that of AI, it does not specify mathematical formulae or processing logic that would allow it to be implemented as a computer simulation in its current form. Furthermore, CCT does not describe how multiple acts might be chosen as a coordinated plan for a goal, nor does it describe how multiple goals might be pursued simultaneously; hence, it is implicitly limited to single-goal, single-act scenarios.

CHAPTER 3

A COMPUTABLE MODEL OF SOCIAL COMMUNICATION

This chapter describes the strategy we followed to create a computable model of social communication, and the model itself. The key objective, in addition to making a simulation that works, is ensuring that the model can be evaluated empirically. We expect there will be an ancillary benefit of encouraging interdisciplinary collaboration as well.

Research Strategy

In order to make incremental progress on an AI model of interpersonal communication, a target too large to encompass in any single project, we have to find ways of splitting up the domain so that simulations of each subdomain can be easily understood and tested. There is a risk in such an approach that our “micro-world” will omit factors of fundamental importance. But that risk can be mitigated by inviting social scientists into a tighter working relationship with the AI effort, relying on their familiarity with the issues to guide the domain-simplification task. One way to encourage a tighter relationship is for AI to provide the service of implementing simulations of social science theories, using the computer metaphor to reveal vagueness, oversights, and unintended consequences of those theories, driving iterative improvement of the theories, and thereby providing value to social scientists.

The guiding motivation of such an approach is similar to that of unit-testing in software engineering: When reassembling components that have each been iteratively improved, the system as a whole should perform better, be more reliable, and be better understood.

And when testing the system as a whole, it should be easier to attribute causes for success or failure.

The trick in applying such an approach to AI models of interpersonal communication is finding social science models that are amenable to computation. At the time when LetterGen and Bickmore's model were being developed, there were hardly any such models -- all that we know of were in game theory. But since then, at least one has appeared, Kellermann's Conversational Constraint Theory (CCT) in communication, introduced in the previous chapter. By adapting and extending CCT, we can not only work toward a computable simulation but also, hopefully, encourage greater interdisciplinary collaboration. This dissertation describes a computer implementation of CCT and shows how that theory can be improved and describes the lessons the effort has for speech act planning.

Adapting CCT

We introduced CCT in the previous chapter, pointing out that it is unique in the discipline of communication, as far as we know, for proposing a *decision process* whereby considerations of efficiency and politeness can be weighed in order to determine a set of candidate acts that would be an acceptable way to pursue a particular interpersonal goal. While CCT does not propose specific mathematical formulae for computing such decisions, it does propose that decisions be made by comparing efficiency and politeness scores for candidate acts to a threshold for efficiency and a threshold for politeness, and that only acts whose scores satisfy both thresholds should be considered acceptable. And while CCT does not specify what the values of the thresholds should be, it does describe

how the threshold values should vary depending on the presence or absence of approximately ten situation factors. The mathematical and process-oriented formulation of CCT goes a long way toward specifying a computable simulation, even though it is not quite there.

What CCT needs in order to become a computable simulation is:

1. Efficiency and politeness scores for each candidate act, covering all goals and situation factors that CCT describes
2. A way of calculating threshold values, by assigning delta values to each situation factor affecting the thresholds.

We shall address the question of delta values for thresholds in the next section. But regarding the need for act scores, these were actually provided for 56 speech act types as a byproduct of a 2004 CCT study evaluating a hypothesis that compliance-gaining goals (e.g., obtain a favor) vary in the degree to which they constrain the politeness and efficiency of compliance-gaining behaviors (e.g., requesting versus hinting). The data was created by asking a large pool of undergraduates to rate each act's efficiency and politeness on a 7-point Likert scale. Kellermann explained her normalization method to us by email this way:

The tactical evaluations were examined for their distribution to make sure they “converged” on the same point. That is, I looked at the distribution of the ratings of each tactic, and checked to see if the scores were normally distributed, or the distribution was “taller” than normal (i.e., the scores converged on the mean with less than expected standard deviation). I found, for the most part, convergence of the ratings onto the mean, or normal distributions. Ratings converged. Z-scores are used to make this assessment. Because all tactics were judged on the same 7 point scale at the same time by participants, standardization is not needed to compare the tactics further, so no additional z-score coding across tactics was undertaken.

We could find no attempts to replicate the score data, but a subset of them can be mapped to findings from a 1994 study (by Kellermann and Tim Cole), so we did an “eyeball” assessment of their similarity. Appendix A shows a table listing all the score categories that we thought were comparable across the two studies. The 1994 study involved 64 communicative strategies, and for each strategy a participant was given 3 sample utterances and was asked to rate the efficiency and politeness of each utterance. Several of the strategy names are also names of speech acts in the 2004 study, and goals are inferable from the content of the utterances (so they can be compared with the goals in the 2004 study). Although some scores are similar across the two, there was no overall pattern of agreement. By itself, this comparison does not undermine the 2004 results, but it does point out the value of a formal replication attempt in future work. We use only scores from the 2004 study.

Other than proving that the scores of an act vary across goals, the scores data is largely incidental to the research questions of the 2004 CCT study, but it is a gold mine for speech act planning. Up to this point, the designer of a speech act planner might speculate about how to plan for goals such as “be polite” that seem to require meeting a (numeric) threshold, perhaps by doing a cost/benefit analysis using a set of (numeric) scores for each potential act. But there was no solid basis for choosing scores, so this kind of approach tended to be dropped in order to avoid the charge of using “magic numbers.” With the CCT study results, we have empirically-derived numbers, and furthermore, the use of thresholds in CCT shows that at least a few communication scholars agree with those in AI that think it is a valid way to simulate discourse planning.

The 2004 CCT study also helps to answer the questions we started with in this proposal about what goals and communicative tactics to include in our model; that is, what range of values would be representative for the independent variable (i.e., the chosen goals, plus relevant situation conditions), and what range for the dependent variable (i.e., communicative tactics). We adopt the 13 goal types, 10 situation factors, and 56 speech act types of that CCT study to define the domain and range of our simulation. These counts are significantly higher than those typical of speech act planners in AI.

Because the term “politeness” is often associated with automatic behaviors such as conventional greetings, we prefer the term “socially appropriate” because it suggests that the conversational participant is thoughtfully weighing how his options might affect others, and it is that type of decision that we aim to simulate. Kellermann uses these terms interchangeably, and we agree that as the term “politeness” is used in Politeness Theory, there is arguably no difference with the meaning of “socially appropriate.” Yet, one might argue that the term “socially appropriate” is broader or merely different than the term “politeness,” and one example proposed to me of this difference is that a rude joke shared with friends might be technically impolite (rude) but nevertheless socially

appropriate. Kellermann disagrees with this use of the term “appropriate”

[personal correspondence]:

It strikes me that the notion of politeness is being confused with the notion of acceptability. A joke has a level of politeness (like the “act” idea in Brown & Levinson’s politeness theory) but the issue is whether that level of politeness is acceptable with these people and in that situation. This is one of the two places I disagree with Brown and Levinson.

Politeness is the evaluation of the tactic (e.g., the joke) per se as “rude.” Acceptability is the evaluation of the use of a rude joke in that particular situation with those particular people. Acceptability is what is determined by the minimum preferred values for politeness and efficiency, and is defined by the preferred strategy space.

Basically, with friends, at home, doing social stuff, one doesn’t need to be particularly polite or particularly efficient. So, telling an off-color joke among friends says that this informal situation with co-equals in status in a private situation (at home) has a low preferred value for social appropriateness.... In other words, it is acceptable to tell the rude joke in that situation with those people at that time under those expectations.... The need to be polite is very low, and so the off-color joke is “acceptable.” In other words, the joke is impolite, but the joke is still acceptable.

The joke has a politeness level that is “low,” and the issue is what politeness level is needed in this situation, with these relational partners, etc.?

Politeness = social appropriateness. Acceptability is what is defined by the preferred strategy space, and so is determined by the minimum preferred levels of politeness and efficiency. Politeness and acceptability differ -- the idea behind CCT is that what is impolite may or may not be unacceptable, as per this example.

We think the distinction Kellermann makes between “appropriateness” and “acceptability” can be difficult to grasp because both are situation- and goal-dependent. We believe that “appropriateness” (and “politeness”) can be operationalized as, “the

extent to which one's action supports the positive or negative face of one's cointeractant,” and that “acceptability” can be operationalized as, “the extent to which one's action is as efficient and appropriate as the situation demands.” Therefore, when a situation demands that one provide face support, and one's action does that, then there is a high degree of overlap. But in the case of telling a rude joke to one's friends, there is little overlap: The situation does not require face support, and yet one's action can be very acceptable due to the *bonhomie* it creates. Our definition of appropriateness is the operationalization above, “the extent to which one's action supports the positive or negative face of one's cointeractant.”

Extending CCT

To render CCT's decision process as a computable simulation, the only necessary but missing piece would be to elaborate how the thresholds could be computed from situation factors, as already mentioned. But we believe that doing just that one task would both gloss over what we view as an oversight of CCT, as well as miss an opportunity to apply the theory to communication scenarios currently just outside its grasp. The oversight we mention is that CCT does not specify that the conditions of an act be checked before the act is chosen. And the potential opportunity we have in mind is to apply CCT not just to single-goal-single-act plans but multi-goal-multi-act ones as well.

To create our simulation, we have made extensions to CCT and confirmed with its author that we are interpreting its concepts as intended. Our model is called Computable Social Communication (CSC) and its elaborations are:

1. A method of pursuing *multiple* goals by simultaneously optimizing the efficiency and appropriateness of a plan containing *multiple* acts. (Like CCT, Bickmore's REA system plans only one act at a time as well. But obviously in long conversational turns, in e-mail, and so on, many acts must be coordinated simultaneously.)
2. A method of calculating threshold values from descriptions of the situation. (CCT is not clear about how much each threshold should be incremented or decremented for each situation factor.)
3. A method of filtering out candidate acts based on unmet conditions. (As already mentioned, CCT does not include a step where conditions would be checked. Such checking is important since the optimal act choice according to CCT might be to make a promise, but if one cannot fulfill the promise the ultimate result could be far from optimal.)

CSC could have been implemented in a variety of frameworks, and planners developed in the field of robotics initially seemed ideal, since they generate measured responses to a variety of numeric sensor inputs and have a large repertoire of moves to evaluate. The core of Bickmore's REA, in fact, is a set of mathematical formulae that are optimized using Maes' *Do The Right Thing* robotics architecture [Maes 1989], which uses a spreading activation network. There are also recent advances over classical AI planning

known as “contingency planning” that are able to take into account uncertainty about the starting state and the effect of each action. But these frameworks can be difficult to understand, and they often require domain-specific search heuristics. Contingency planning, in particular, is useful only when plan operators in the domain typically have conditions satisfied by the effects of other operators, which leads to tree-shaped plans, because contingency planning focuses on pruning the search space for the tree [Huang 2006]. Plans in CSC are flat because the conditions of its speech acts (e.g., be able to do what one promises) are typically not satisfiable by the effects of any other speech act. In fact, we found that the kind of optimization we wanted to do would easily fit a linear programming framework, a framework that is fairly easy to grasp and which allows us to keep the core planning concepts at center stage. As an example, consider the following formulae:

Minimize

num_acts: vow + insist + assert + offer + argue

Such that

Get_Date.effMin: $-0.3 \text{ vow} - 1.15 \text{ insist} + 0.85 \text{ assert} + 1.3 \text{ offer} - 1.65 \text{ argue} \geq 0.7$

Get_Date.appMin: $0.45 \text{ vow} - 1.2 \text{ insist} + 1.1 \text{ assert} + 1.6 \text{ offer} - 1.55 \text{ argue} \geq 1.2$

Binaries

vow
insist
assert
offer
argue

This is a standard integer-linear-programming (ILP) file format that says 'vow', etc are variables that take values in the set $\{0,1\}$, and that of all the possible solutions to the two inequalities in the 'Such that' section, one should prefer those that minimize the

'num_acts' sum. In other words, find the smallest set of acts that fit the thresholds. In any such solution, the variables assigned a value of 1 are those that should be included in the plan, and the variables assigned a value of 0 should be left out of the plan. The coefficients of the variables are the efficiency and appropriateness scores found in Kellermann's study results. To solve for multiple goals, a pair of inequalities is added for each goal; the coefficients will be specific to each goal but the thresholds remain the same. Adding inequalities in this way fits with the common perception that each additional goal further constrains the space of possible solutions.

In order to formulate the planning problem in this manner, we made two assumptions:

1. CCT does not say how the scores of several candidate acts might be combined before comparing them with the thresholds. We assume that addition is a better choice than more complicated n-ary functions.
2. CCT says that act scores must fit within thresholds but does not say how one might select among all the possible solutions. We assume that minimizing the number of acts is best, because it follows Grice's maxim of brevity.

Verifying both of these assumptions is a worthy goal for future work.

CSC solves its ILP problems using the GNU Linear Programming Kit, “GLPK” [FSF 06]. Although ILP is NP-hard, and we are solving systems with 56 variables, it typically takes less than 100 milliseconds on a standard consumer computer to find a solution. ILP problems with only a few variables can take hours using the same hardware, but by defining our minimizing function as addition with all coefficients as +1, GLPK is able to use a special “branch-and-bound” algorithm and be extremely efficient.

Because Kellermann compiled data on 56 speech act types, one can generate plans with a greater variety of acts than could LetterGen (which has 40 types), and with far greater variety than can the typical speech act planner (which has fewer than ten types). Bickmore's REA avatar system had only a handful of types, but it did not generate plans from these; instead, it had a library of act instantiations such as “How about this weather?” and evaluated those. The REA library is split between efficiency-only oriented utterances and politeness-only oriented ones (“small talk”). None of REA's utterances are meant to achieve efficiency and politeness simultaneously. But the fact that REA planned with utterances raises an important point: CCT does not specify a method for instantiating the plans it generates; it was assumed that some external natural language generator capable of socially-appropriate styling like PAULINE [Hovy 1988] would be capable of continuing the decision process from that point [Kellermann, personal communication]. We shall return to this assumption of CCT in our concluding chapter.

The next contribution of CSC, after providing a way of optimizing both efficiency and appropriateness through multiple acts, is providing a way to calculate the threshold values from situation descriptions. CCT outlines a set of principles describing what aspects of a conversational situation would affect the speaker's minimum preferred level of efficiency, appropriateness, or both. For example, the Bond Principle is,

The greater a person's concern about the current status of his/her relational bond with a conversational partner, the higher the person's minimum preferred level of politeness. The less concern about the current status of the relational bond, the lower the minimum preferred level of politeness. [Kellermann 2006]

As an example of such a scenario, imagine meeting and talking with a stranger. If one had little interest in interacting with the stranger again, relational concern would be low, and as the principle asserts, one would not feel compelled to be especially polite. Otherwise, if one wanted the meeting to extend into a cordial relationship, relational concern would be higher and, correspondingly, one would expect to be more polite. Similarly, if the speaker were already in a cordial relationship with the hearer, and felt secure in or perhaps ambivalent about the relationship, politeness would be low, while if the speaker felt anxious to maintain or escalate the relationship, politeness would be higher. CCT identifies ten such principles (see Appendix C).

The mapping from situation factors to changes in efficiency and appropriateness levels is qualitative rather than quantitative, which makes it difficult to incorporate into our simulation. To incorporate the mapping, the simulation would need a pair of values representing the default minimum preferred levels of efficiency and politeness for the typical situation, plus a pair of delta values that indicate how much to increment or decrement the defaults based on the strength of each factor present in the situation. We know of no empirical study, nor does CCT's author, that could provide such numbers. We made an exploratory attempt to approximate defaults and deltas by seeing how fast the number of acts in a plan increases as thresholds are increased, but we were not satisfied with the results, so we rely on guessed values (i.e., 0.5 for all origin and delta values) for the sample sessions presented in our later section. Verifying our guessed values would be a worthwhile project for future work.

The third and last contribution of CSC is providing a method of filtering out candidate acts based on unmet conditions. As mentioned above, CCT does not mention act conditions at all, despite borrowing its set of acts from speech act theory, but their importance is clear from simple examples such as promising to do something that one is incapable of achieving.

For the CSC simulation, we need only the conditions of each act type (although we will return to dictums and illocutionary purposes in our Conclusion chapter). We have rephrased Wierzbicka's wording to fit our questionnaire style. The conditions need further formalization because they are not yet as relevant to the goal and situation as they should be; for example, a threat should match the level of seriousness of whatever it is intended to discourage. That is, one threatens potential criminals with prison, but one threatens naughty children only with a loss of television privileges. Any formalization of speech act conditions found in linguistics is likely to have such a limitation because linguists try to avoid committing to psychological effects on beliefs, goals, and so on. Such formalization is beyond the scope of this project as well, but is a high priority for future work.

We have created a process loop where we check the conditions of acts after generating a plan, and if any conditions fail, we re-generate a plan using any acts whose conditions were satisfied plus all the acts whose conditions have not yet been checked. Checking conditions only after a plan has been generated is well-suited to an interactive application like CSC, although if there were a computable representation of the entire situation, and

if the act conditions were rendered in a computable pattern-recognition format, one could eliminate all nonconforming acts at the start and generate a plan only once.

At the end of a CSC simulation session, either some set of acts recommended by the system has proven to be workable (because the user indicated that all of the act conditions were satisfied), or there is no solution because the problem is overconstrained. But in practice, there is almost always a solution because the lack of *maximum* thresholds means that the system can usually overcome any obstacle by adding more acts that have a positive effect on the overall efficiency or politeness score. Unsolvability occurs only when every act that has such positive effects has been rejected because the conditions were not satisfied, or because the threshold is so high that even the sum of all positive workable-act scores is not enough to reach it.

In the CSC simulation process we have described, there are three major tasks to be coordinated: determining how the thresholds should be adjusted for the situation, formulating the planning situation as an integer-linear-programming problem and solving it, and then checking the conditions of the candidate acts in the solution. Our method of weaving together those tasks is summarized by this algorithm:

- 1:** Determine which subset $\{G_1..G_n\}$ of the 13 goals is to be solved for.
- 2:** Set the efficiency and appropriateness thresholds to their initial default values.
- 3:** For every threshold-affecting situation factor F_i ,
 - 3.1:** Prompt the user to rate the factor on a scale;
 - 3.2:** Normalize the user's response to a negative-zero-positive scale, apply the

multiplier, and add to the appropriate threshold (perhaps both).

4: For every act A_j ,

4.1: Label the act as “unknown acceptability.”

5: While there is still an act not marked “unacceptable,”

5.1: For every goal G_k to be solved,

5.1.1: Insert into the ILP problem file a pair of inequalities, where the variables of both correspond to the acts not marked “unacceptable”;

5.1.2: One of the inequalities will represent efficiency; its coefficients are found through table lookup based on goal G_k , and its threshold was calculated in step 3;

5.1.3: The other inequality represents appropriateness. Its coefficients are also found through table lookup based on goal G_k , and its threshold was also calculated in step 3.

5.2 Run the ILP solver over all pairs of inequalities simultaneously, requiring all variables to take values in $\{0,1\}$, and minimizing over the sum of the variables. Let the resulting plan be said to have candidate acts $\{C_1 \dots C_m\}$.

5.3: Tentatively set a flag indicating that the problem is “solved.”

5.4: For every candidate act C_p ,

5.4.1: If the act has already been marked “acceptable,” then skip to the next act;

5.4.2: Otherwise, determine whether the conditions of the act are satisfied as a whole or not.

5.4.3: If the conditions are not satisfied, mark the act “unacceptable” and change the problem flag to “not solved.”

5.5: If the problem flag is still set to “solved,” then exit the loop and declare success.

Two aspects of speech act planning that CSC does not provide answers for are:

- How should the planned acts be ordered relative to each other when realized?
- How might the planned acts interfere with or promote each other? For example, a plan using the acts of begging and threatening might score well, but might these two undermine each other?

These are open questions in speech act planning and there is no general consensus on how they might be solved.

CHAPTER 4

EVALUATION 1: SUBJECTIVE PLAUSIBILITY

Transcripts of three sample CSC sessions are provided in Appendices D, E, and F. The first two transcripts show CSC planning for a single goal, and the third shows it planning for two goals simultaneously. The first and second transcripts show how CSC is sensitive to small changes in the situation. We present these sample sessions as a kind of evaluation, complementary to our empirical evaluation in the next chapter, but in this case depending on the reader's subjective sense of whether the simulation's generated plans are plausible. We assume that the reader will share our sense that the plans *are* plausible.

The scenario of the first two sessions is: You are an in-car automated driving advisor, your driver is a male engineer who dislikes being told what to do, and you want to persuade him to use his turn signal when changing lanes. We assume the goal here is an instance of persuading the driver to Fulfill an Obligation.

Sample One: The Reckless Driver

In the first sample session, where the simulator/planner plays the role of an in-car driving advisor, the planner's first recommendation was the acts {thank, apologize}. The “thank” act has the following conditions:

- There is some act that your audience has already done;
- The event benefited you;
- You feel something good about your audience because of what they did;
- Your audience will feel good knowing what you think about them.

If it were the responsibility of the program to find something to satisfy these conditions, it might consider many possibilities such as thanking the driver for breathing, or for brushing his teeth that morning, or for holding tightly to the steering wheel. But all these possibilities would fail the last condition because they would not make the recipient feel good, not to mention the risk of sounding sarcastic that they would introduce. There are other possibilities that might pass all the conditions but should still be rejected; for example, imagine that the driver is well-known for making large donations to charity. The problem with such possibilities is that they are not goal-relevant. As mentioned in chapter 3, CSC's conditions do not handle such cases explicitly, and it is an open problem for speech act planning in general, but it is an important area for our future work.

We rejected both “thank” and “apologize” in this sample session. We could not think of anything in the scenario that would satisfy all of the conditions of either act and which was also goal-relevant.

The planner's second recommendation was the acts {compliment, praise}. Complimenting only required finding something good about the driver to remark on, and we saw no harm in saying he is a “good” driver. But giving praise required that the driver have done something worth rewarding, and we could not think of a single thing, so we rejected that act.

The planner's third recommendation was {compliment, suggest}, where it re-used “compliment” because it had already been accepted as a possibility. The other act, “suggest,” required that there be some action that would be to the driver's benefit that we could bring to his attention. Using one's turn signal has the potential benefit of avoiding being hit by less attentive drivers, so that act was also accepted. At this point, the session ended successfully with a plan comprised of “compliment” and “suggest.”

Sample Two: The Reckless Driver and the Impolite Interface

This example demonstrates the sensitivity of CSC to situational factors. To make that point, we have reused the “Reckless Driver” scenario but have changed our answer to the question, “How polite does your audience expect you to be in this conversation?” from 5 to 3. That is, the only difference between the two scenarios is that in this case we assert that the driver expects the interface to be less polite. By changing our answer to the questionnaire in this way, the final value for minimum preferred efficiency remains at 2, but the final value for minimum preferred level of appropriateness goes down from 4.5 in the previous scenario to 3.5 (highlighted in the transcript in Appendix D).

With this slight change of the appropriateness threshold, the planner's first suggested plan has changed from {thank, apologize} to {thank, question}. In the planner session for the previous scenario, and for this one, the planner came across {thank, question} in the search space before coming across {thank, apologize}, but in the previous session, the appropriateness threshold was too high to be met by {thank, question}, so it kept searching. Since that threshold is low enough in the current scenario to be met by {thank, question}, search did not have to go further. However, there is no reason why {thank, question} should be found first, given the way we have set up the search space, so this difference in plans is just a fortuitous coincidence that happens to show the influence of situation factors.

We rejected the “question” act because its conditions could not plausibly be met, based on our understanding of the scenario, and the next recommended plan was {disclose, compliment}. This second recommendation is similar to the second recommendation of the previous scenario, {praise, compliment}. It is worth noting that the difference in act choices across both recommendations is entirely explainable by the appropriateness scores of the acts for the Fulfill Obligation goal – “apologize” scores at 2.05 versus 1.10 for “question,” and “praise” scores at 1.95 versus 0.90 for “disclose.”

We also rejected “disclose” because its conditions could not be met, but we found the next recommendation, {compliment, tell}, acceptable. This third recommendation is similar to third recommendation of the previous scenario, {compliment, suggest}, and again the difference can be explained by the difference in act appropriateness scores -- 1.95 for “suggest” versus 0.90 for “tell.” We imagined that the conditions of “tell” could

be satisfied by asserting straightforwardly that, if the driver kept changing lanes without signaling, he would eventually have a collision. By comparison, the “suggest” act of the previous scenario gave similar information but highlighted the benefit to the driver of complying. We assume the “compliment” act of both plans would be instantiated the same way in both plans. When interpreted in this way, the {compliment, tell} plan of the current scenario seems less polite than the {compliment, suggest} plan of the previous scenario, which fits what one would expect from lowering the appropriateness threshold in this scenario.

If one changed the scenario even more, one could get recommended plans that have fewer acts or more acts, not just a change in acts. For example, a radical increase in the appropriateness threshold would require adding acts that could raise the appropriateness of the plan while not lowering its efficiency below the efficiency threshold.

Sample Three: The Embarrassing Husband

In the third session, the scenario is: You and your husband are a middle-aged couple invited to a fancy dinner party. Your husband is slurping his soup and getting some in his lap. We assume stopping him from slurping his soup is an instance of `Stop_Annoying_Habit`, and that helping him avoid staining his trousers is an instance of persuading by `Providing_Guidance`. That is, the scenario requires the planner to solve two social goals simultaneously. Since Kellermann's data indicates that a single act can contribute different amounts of appropriateness depending on the goal being pursued, and different amounts of efficiency as well, it is nontrivial to find a minimal set of acts that satisfies more than one goal at a time.

For this scenario, the planner suggests three acts, {give, ask, inform}, which are plausible when instantiated like this:

Give:	Unfold your husband's napkin and hand it to him.
Ask:	Ask your husband if he is aware that his trousers are being stained.
Inform:	Tell your husband that he is slurping loudly and embarrassing you.

This example is interesting because the thresholds indicated by the situation were relatively low (i.e., 4.5 for appropriateness and only 0.5 for efficiency), and yet more acts were needed than in our previous scenarios (i.e., three acts instead of two). The disparity between the low thresholds and the need for more acts indicates that the most influential constraint on the search was the need to solve the goals simultaneously.

One might also imagine non-verbal methods the wife might use to the same effect, such as a solemn stare. Non-verbal acts would be an interesting extension of CSC, but we are not aware of any taxonomy nor any set of scores that would allow us to include them at this point, so we leave it as an area for future work.

Having a simulator makes it easy to imagine many scenarios and see whether the generated plans seem plausible (and may even suggest entertaining ways to extend the scenario). We invite the reader to download the system (see Appendix B), play with scenarios, and see how plausible its plans are for oneself.

CHAPTER 5

EVALUATION 2: PREDICTIVE POWER

To determine the predictive power of our CSC model, we used a survey to compare the model's judgments of the acceptability of a set of communicative plans with human judgments of the acceptability of the same plans. There is a supportive trend for the bulk of the model's predictions.

Survey Design

As mentioned earlier, the claimed contribution of CSC is: *CSC can generate plans for multiple goals using multiple acts in a way that is sensitive to a wide variety of specific situation conditions*. To evaluate the claim, we designed our survey to gather human judgments on a set of plans generated by CSC. The generated plans all address the same pair of goals and the same situation conditions, but the plans differ by whether CSC expects each plan to be acceptable for both goals, for only one goal, or for neither goal. Generating plans expected to fail is not useful by itself, of course, but is intended to demonstrate that CSC's predictions are reliable across the spectrum of plan quality, not just at the extreme of plans widely agreed to be good ones. Similarly, in the survey, participants also indicate for each goal whether each plan would be acceptable (i.e., for one specific goal, for both goals, or for neither). By comparing simulator and participant judgments at this level, we can not only determine the overall accuracy of the simulator, but we also have enough data to identify possible sources of weakness in the simulator if it is found to be inaccurate in some or all cases. For example, if we find that the simulator

is inaccurate for only one of the goals and only in plans that share a certain speech act, then we can identify the score of that speech act for that goal as being suspect rather than the simulator itself.

To ensure that the simulator and the participants would have the same “inputs,” we created an everyday scenario of someone pursuing goals of Getting Advice and Obtaining a Favor and then made sure that the scenario was fleshed out enough to provide values for all the situation factors that the simulator depends upon. (See the pilot and survey questionnaires in Appendices G and H, respectively.) In the scenario, the participant and the simulator play the role of a young apartment dweller who is preparing for a short trip by looking for someone who can feed their pet (i.e., a case of obtaining a favor) and who also knows a good route to the airport (i.e., a case of getting advice). The parent of the apartment dweller has recommended a family friend, and it is the job of the apartment dweller to meet with this family friend and persuade them to adopt both goals. The family friend is described as a bit old-fashioned, formal, and opinionated, and these characteristics are further broken down into the situation factors of the model as particular numeric values on a scale for each factor. For example, the formality of the situation is described as 6 on a scale of 1 to 7. The urgency of the situation is similarly high, and so on. To ensure the “inputs” are the same for the simulator and participants, we had to make a simplifying assumption that all conditions of all candidate acts were satisfied in the scenario – without this assumption, we felt the scenario description (already two pages long) would have become unwieldy for participants.

The scenario description forms the first half of the survey questionnaire; the second half is comprised of plans generated by the simulator for the scenario, plus the survey questions. Each plan is an unordered set of three to four speech act types with a short sample English sentence for each; for example, consider this example:

Advise – This is the best way to do it.

Explain -- ...and that's why I need your help.

Question – Are you sure that's right?

Is this set an acceptable way to Get Advice? Yes___ No___ Not sure___

Is this set an acceptable way to Obtain a Favor? Yes___ No___ Not sure___

The English samples are designed to be generic and not refer to the scenario, because we wanted to avoid introducing bias if one sample were to appear more topically relevant than others.

When presented with a plan using generic English samples like those above, human subjects might complain that it is difficult to evaluate such plans, and they might request more concrete sample utterances. Although we sympathize, we avoided such concreteness, and we have principled reasons for doing so. Most importantly, using concrete utterances would make it much harder to evaluate CSC because it would be much harder to assign credit or blame for the (im)plausibility of its plans, because there would be more possible causes for success or failure. For example, creating concrete utterances for CSC's plans would mean that either an automated language “realizer” would be added to the program, or a human would have to assign canned text to each act. The text created by realizers is often stilted and could bias the results, and human-generated text might use a style that is noticeably human and bias the results in the

opposite direction. This problem of evaluating core functionality based on the outward behavior of a large, complicated architecture is the same problem one faces when evaluating Bickmore's REA agent, where the planner is nested inside a 3D animated character that speaks and responds to speech, gesture, and gaze. On the other hand, we know from Kellermann's studies with the same abstract act types that one can get internally-consistent and plausible results without using utterances in the questionnaire.

There are eight plans to be judged in the questionnaire, two each for the following four conditions:

- The plan is predicted to be acceptable for both goals
- The plan is predicted to be acceptable only for the Get Advice goal
- The plan is predicted to be acceptable only for the Obtain a Favor goal
- The plan is predicted to be unacceptable for either goal

There are two plans for each case so that, if a prediction fails for one plan, we can look at responses for the other plan of the same case to see if the failure of the prediction might be a fluke.

Figure 13 illustrates the survey design as a bar chart, showing four bars for each of the eight plans, plus two horizontal threshold lines. The threshold lines represent efficiency and appropriateness thresholds, and since the values of the thresholds are determined by situation factors as described in the scenario, the same threshold values apply to *all* candidate plans. The bars plotted vertically against the thresholds indicate the efficiency and appropriateness scores of each plan, and since these scores differ by goal, there are four bars for each plan (i.e., two kinds of score times two kinds of goal). In the chart, a

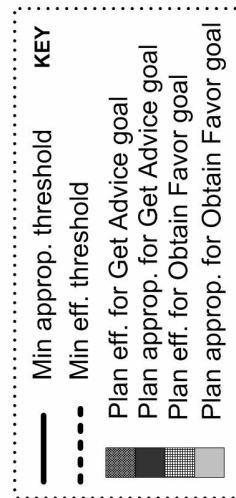
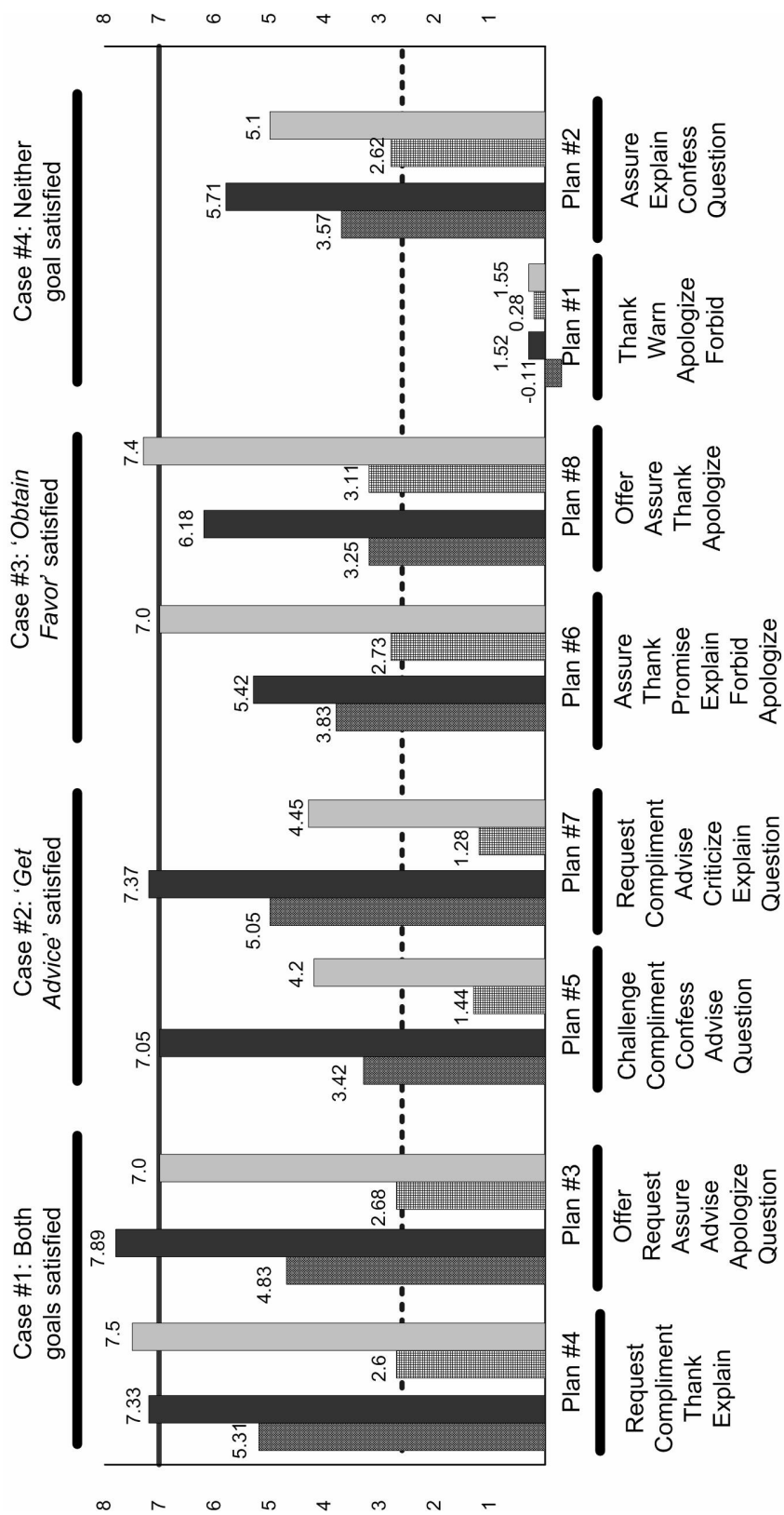


Figure 13.
Four cases guiding
the design of the pilot

plan is acceptable for a goal if its efficiency score meets or exceeds the efficiency threshold and its appropriateness score meets or exceeds the appropriateness threshold. For example, plan 3 satisfies both goals because all four of its scores meet or exceed the corresponding thresholds. But plan 5 is acceptable only for the Get Advice goal because only the bars for that goal meet or exceed their thresholds.

Each pair of bars in Figure 13 represents a hypothesis of the study – a prediction of whether a certain plan is acceptable for a certain goal. The questions of the survey correspond one-to-one with these hypotheses. For example, one question asks if plan 4 is an acceptable way to pursue the Get Advice goal. (This question corresponds to the leftmost pair of bars in the diagram.) The aim of our analysis will be to take the number of participants who agree with our prediction for each question and determine the significance of that level of agreement. Since each prediction is binary-valued (plus allowing for a “not sure” option), we use a binomial distribution calculation to determine significance [Abdi 2007]. We review the results for a pilot and our survey in later sections.

The job of the simulator in creating the survey was to find plans that met the threshold criteria. The simulator is designed to search only for plans that satisfy all goals, so for lesser cases of mixed success or total failure, we calculated plans by hand.

The survey is cross-sectional, capturing the judgments of all participants on one day. The questionnaires were distributed in the early morning and afternoon to three introductory computer science classes at a community college in Hawaii. We aimed to recruit an

approximately equal number of men and women all between 18 and 23 years old, to match roughly the demographic of Kellermann's original CCT studies – undergraduates at Californian and Midwestern universities. Participants were offered \$5 if they would complete all items of the questionnaire.

Lessons from the Pilot

To validate the questionnaire, we ran a pilot using a website dedicated to online surveys and experiments hosted by the Department of Psychology at the University of Zürich that is open to noncommercial researchers for a monthly listing fee (i.e., <http://genpsylab-wexlist.unizh.ch/>). The online survey was open continuously for two months to anyone interested in participating, was anonymous, and we did not recruit participants. We closed the survey once we reached 80 participants, which we winnowed down to 53 after removing outliers who had answered only two or fewer of the sixteen questions.

At the time of the pilot, we wanted the data to be as rich as possible, so instead of offering yes/no options for each question, we prompted with a 7-point Likert scale of plan acceptability ranging from “not at all” to “absolutely.” We expected that the greater granularity of such options would help to suggest reasons why some predictions might not have been supported, although we did not use this granularity for the analysis in the end. However, we did use the granularity of responses to further winnow outliers: We calculated the difference between each response and the average response for that question, summed the differences for each participant, and removed anyone whose sum

was unusually high. There was a cluster of eight participants whose sums were especially high and whom were removed, leaving 45 participants for the final pilot analysis.

Our choice to prompt with a Likert scale instead of a binary choice made analysis more difficult than expected, because our predictions are binary and we needed to map the scaled responses to the binary predictions. There is an obvious way to do such a mapping; namely, by assigning all values below the median to “no” and all those above to “yes”; however, we were aware that scaled options are open to biases such as participants who have a negative view of the world and rate all options low. Therefore, we decided to compare each response to a normed response for that participant across all his responses. Each participant’s normed response was calculated as half the difference between his highest-valued response and his lowest-valued, added to his lowest value. To do the mapping to our predictions, any response higher than the participant’s norm was counted as a “yes” and any lower were counted as a “no”; any response exactly equal to the norm was considered a non-response and ignored like all actual non-responses were. Once the mapping was done, we had a vector for each participant indicating a “yes,” “no,” or non-response for each question. For the actual analysis, we counted the number of responses that agreed with the prediction for each question, versus the total number of responses for the question, and calculated a binomial test of statistical significance.

The null hypothesis corresponding to each of our predictions is that it is equally likely that someone would agree with our prediction as that they would not agree. In actual results, the number of agreements and disagreements is unlikely to be exactly equal, and

the question is how significant that imbalance (and difference from the null hypothesis) is. A similar real-world test would be to flip a coin, noting the number of heads and tails, and then, assuming the counts are not exactly equal, calculate whether the inequality is due to mere chance or whether the inequality is so large as to suggest that it is an unbalanced, trick coin. To have an interesting survey result, one hopes that the number of agreements is so much higher than the number of disagreements that it just would not be plausible to accept the null hypothesis; that is, hopefully, the results compel one to drop the null hypothesis and accept instead the proposed hypothesis. Viewed another way, our proposed hypothesis is that there is a “trick coin” underlying judgments of social acceptability, a strong tendency shared across the survey population to judge certain communicative plans the same way. The binomial test is a widely accepted method of determining the probability that such imbalances are not due to mere chance.

Figure 14 shows the same survey design elements as Figure 13 but adds a table at the bottom showing the binomial test result for each of the sixteen predictions. We consider probabilities of .85 to be weakly supportive and those .95 and above to be strongly supportive. Nine of the sixteen hypotheses received strong support, and one, weak support. This result is not bad, but we knew that that did not have to be the end of the story since our predictions were based on working assumptions for the threshold values. There is an opportunity with the hypotheses showing strong disagreement to turn them into strong agreement, if the thresholds can be adjusted to change those predictions without changing the predictions that already have strong support. For example, the

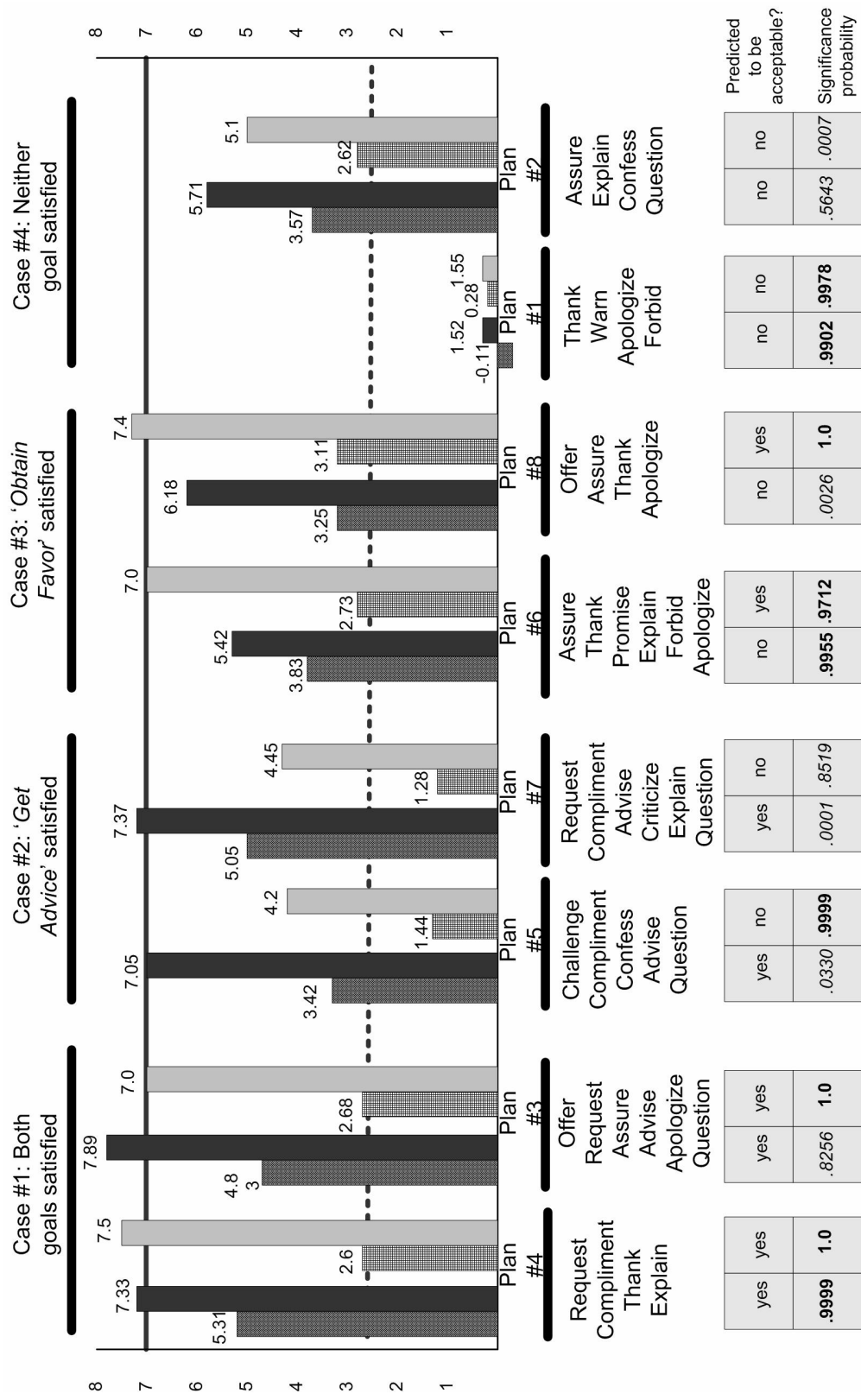


Figure 14. Pilot results shown below the pilot design

Obtain Favor prediction for plan 2 received support of only .0007, but if the appropriateness threshold had been 5.1 instead of 7.0, the level of support would flip to .9993. There are four such opportunities:

1. If the appropriateness threshold were lowered to 6.18, the Get Advice prediction of plan 8 would receive support of .9974;
2. (As already mentioned) If the appropriateness threshold were lowered to 5.1, the Obtain Favor prediction of plan 2 would have support of .9993 (and the Get Advice prediction of plan 8 would receive support of .9974, but we would lose support for the Get Advice prediction of plan 6);
3. If the appropriateness threshold were raised *just above* 7.05, or the efficiency threshold raised *just above* 3.42 (or both), the Get Advice prediction of plan 5 would receive support of .9670;
4. If the appropriateness threshold were raised *just above* 7.37, or the efficiency threshold raised *just above* 5.05 (or both), the Get Advice prediction of plan 7 would receive support of .9999.

(The remaining predictions without strong support, namely the Get Advice predictions of plans 2 and 3, are not opportunities since their probabilities are near 0.5.)

We cannot take advantage of all of these opportunities, since they are inconsistent. But we tried each one and found that only the first two increase the number of supported predictions -- and then by a net gain of only one -- to ten with strong support.

Figure 15 shows the effect of adopting opportunity #1 above, and Figure 16 shows the effects of adopting opportunity #2. In places where changing the threshold causes a

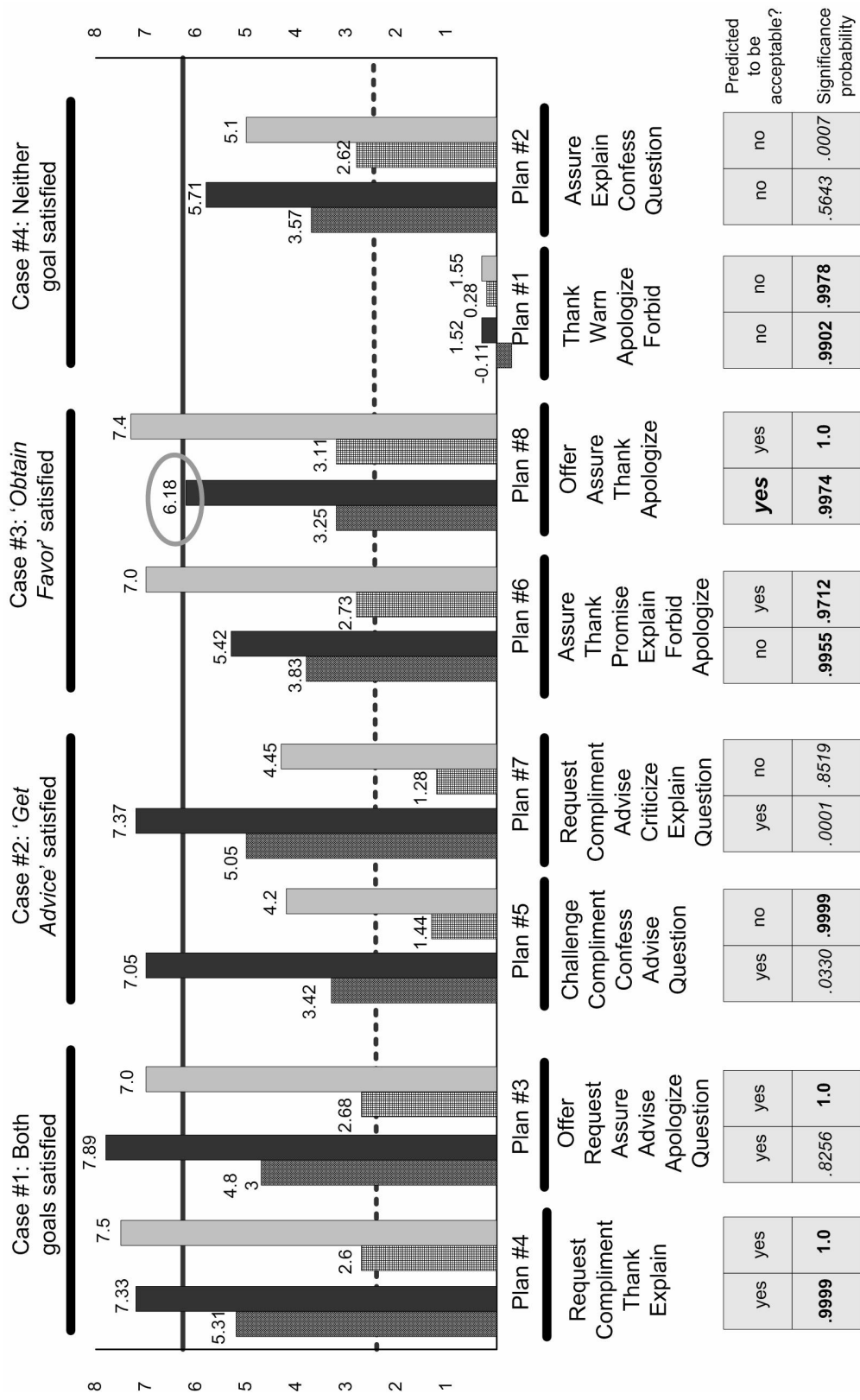


Figure 15. Pilot results after adjusting the appropriateness threshold down to 6.18

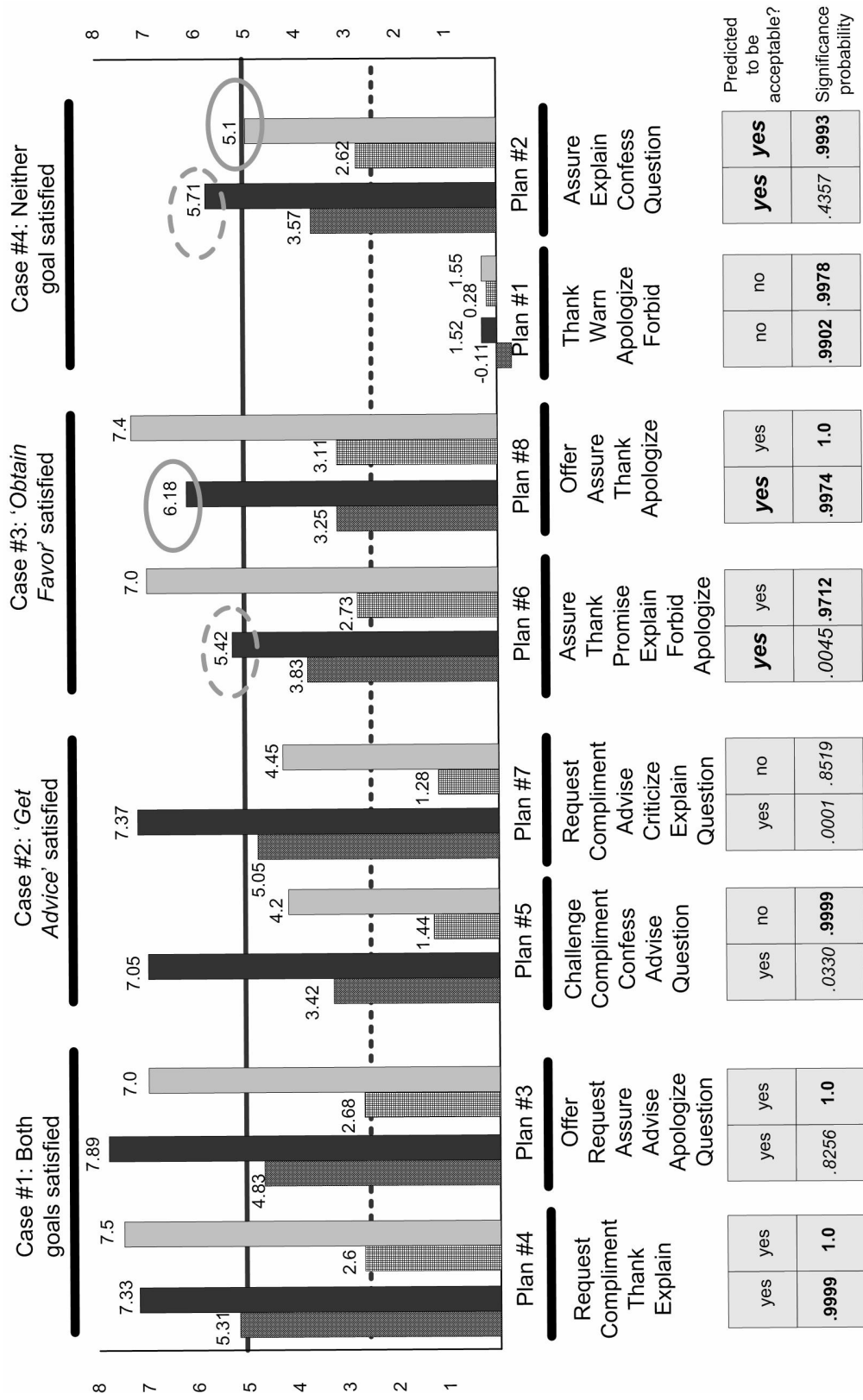


Figure 16. Pilot results after adjusting the appropriateness threshold down to 5.1

change from disagreement to agreement, we use a solid circle around the crossing point, and in places where changing the threshold causes a change from agreement to disagreement, we used dotted circles. For example, in Figure 17 there are two new agreements and two new disagreements (but one of the disagreements is weak, for a net gain of one strong agreement).

We took this finding as a lesson that the appropriateness threshold should be lowered from 7.0 to at least 6.18, and planned to feed this change into a revised set of plans to be used in the survey described later.

After adjusting a threshold in one of the two suggested ways, there are six predictions whose lack of support calls for explanation. But before getting too concerned with the lack of support, we want to take advantage of the redundancy we designed into the questionnaire (i.e., using two different plans to test each of the four cases in which we are interested) to see whether the lack of support is present in the redundant instances as well. It turns out that in two of the three cases where participants showed little or no consensus (before adjusting thresholds), the redundant prediction found strong agreement [Table 6].

It also turns out that in one of the three cases where there is strong disagreement even after changing the threshold, that the redundant plan actually shows strong agreement [Table 7].

Table 6. Comparing support for redundant predictions, where one prediction found little or no consensus

Prediction with little or no consensus	Redundant prediction
“Plan 2 is not acceptable for the Get Advice goal” : .5643	“Plan 1 is not acceptable for the Get Advice goal” : .9902
“Plan 3 is acceptable for the Get Advice goal” : .8256	“Plan 4 is acceptable for the Get Advice goal” : .9999
“Plan 7 is acceptable for the Obtain Favor goal” : .8519	“Plan 7 is acceptable for the Get Advice goal” : .0001

Table 7. Comparing support for redundant predictions, where one prediction found strong disagreement

Adjusted appropriateness threshold	Prediction with strong disagreement	Redundant prediction
5.1	“Plan 6 is not acceptable for the Get Advice goal” : .0045	“Plan 8 is not acceptable for the Get Advice goal” : .9974
6.18	“Plan 2 is not acceptable for the Obtain Favor goal” : .0007	“Plan 1 is not acceptable for the Obtain Favor goal” : .9978

The fact that testing the same prediction with different plans results in vastly different levels of support suggests either that the plan of the failed prediction had some flaw (such as a participant bias about one of its speech acts) or that the model itself is not reliable for such predictions.

In the remaining two of three cases where there is strong disagreement (i.e., plans 5 and 7 should be acceptable for the Get Advice goal), the two cases are redundant. Such a double failure casts doubt on the reliability of the model for such predictions. Note, however, that almost all of these troublesome cases involve predictions where plans are

expected to satisfy only one of the two goals, or neither goal; the predictions involving plans expected to solve both goals received the most consistent support, with the lowest probability being .8256.

In our quest to explain our failed predictions, we have already revisited our assumed threshold values, and we have compared results for our redundant predictions. Two other approaches to try are:

- Considering whether faulty score data might have led us to predict the wrong outcomes;
- Considering whether some of the sample sentences given for speech act types in the questionnaire was biased by being more goal-relevant than other samples.

Regarding the possibility of faulty score data, recall that we had compared the 2004 score data we are using with score data from a 1994 study and had found no overall pattern of agreement in an “eyeball assessment.” That lack of agreement might indicate a problem in the 2004 scores. However, since the 2004 study has 1456 scores and only 34 of those are comparable to the 1994 scores, we must be careful not to assume too much from this lack of agreement. We looked at which of the speech acts used in the pilot might be in the common set of 34 and found that only four are. And of those four, only two are involved in failed predictions. Thus, there really is no basis for assuming that there are faulty scores and that they are responsible for the failure of any predictions.

Regarding the possibility that some sample sentences might have introduced bias, consider that some samples mention “helping” and “favors” and thus might make their associated speech act types appear more relevant to the Get Advice and Obtain Favor

goals. In contrast, the sample for speech act type “forbid” is “don’t mess with that,” which is less obviously relevant to the two goals. We made a subjective assessment of which samples were more relevant to helping and favors, and then we looked to see if the associated acts tended to occur in failed predictions of “yes” responses, but found no such pattern. We did the same assessment for acts in failed predictions of “no” responses, and again found no pattern. These efforts exhausted our ideas for how we might explain the failures.

Nevertheless, in the pilot nine of our sixteen predictions found strong support; eleven, if thresholds were adjusted. And in three of the remaining five cases we can find reason to believe some flaw in the questionnaire might be the cause of the failure, rather than the model itself, since redundant predictions found strong agreement.

We learned from the pilot that the questionnaire should offer binary choices (plus a “not sure” option) rather than a Likert scale, since scales added unnecessary complexity. And we also learned that the appropriateness threshold value we had adopted as a working assumption should be lowered. Finally, we observed that the predictions involving plans expected to satisfy both goals received strong support more consistently than the other three cases of mixed plan success or total failure, and that we should monitor the survey results for the same pattern.

Survey Results

Our survey design was the same as our pilot design, including the built-in redundancies, except that the survey required a new set of plans due to the adjusted appropriateness

threshold value. We happened to make a calculation error when generating the new set of plans, using a threshold value of 4.45 instead of 5.1 or 6.18 for appropriateness as suggested by the pilot results, but this error had no apparent effect on the survey results, as we shall demonstrate. Figure 17 illustrates the survey design and results.

Although we aimed for gender balance in our participant pool, only five of the 23 participants were female. We might have avoided the imbalance by not using a computer science course in which to distribute the survey, but this was a convenience sample. None of the responses displayed any unusual pattern of answers (based on an “eyeball” assessment), and only one participant was marked as an outlier, due to leaving one question blank.

Eight of the sixteen predictions received strong support, and two were weakly supported, similar to the pilot results where nine were strongly supported and one, weakly. And the

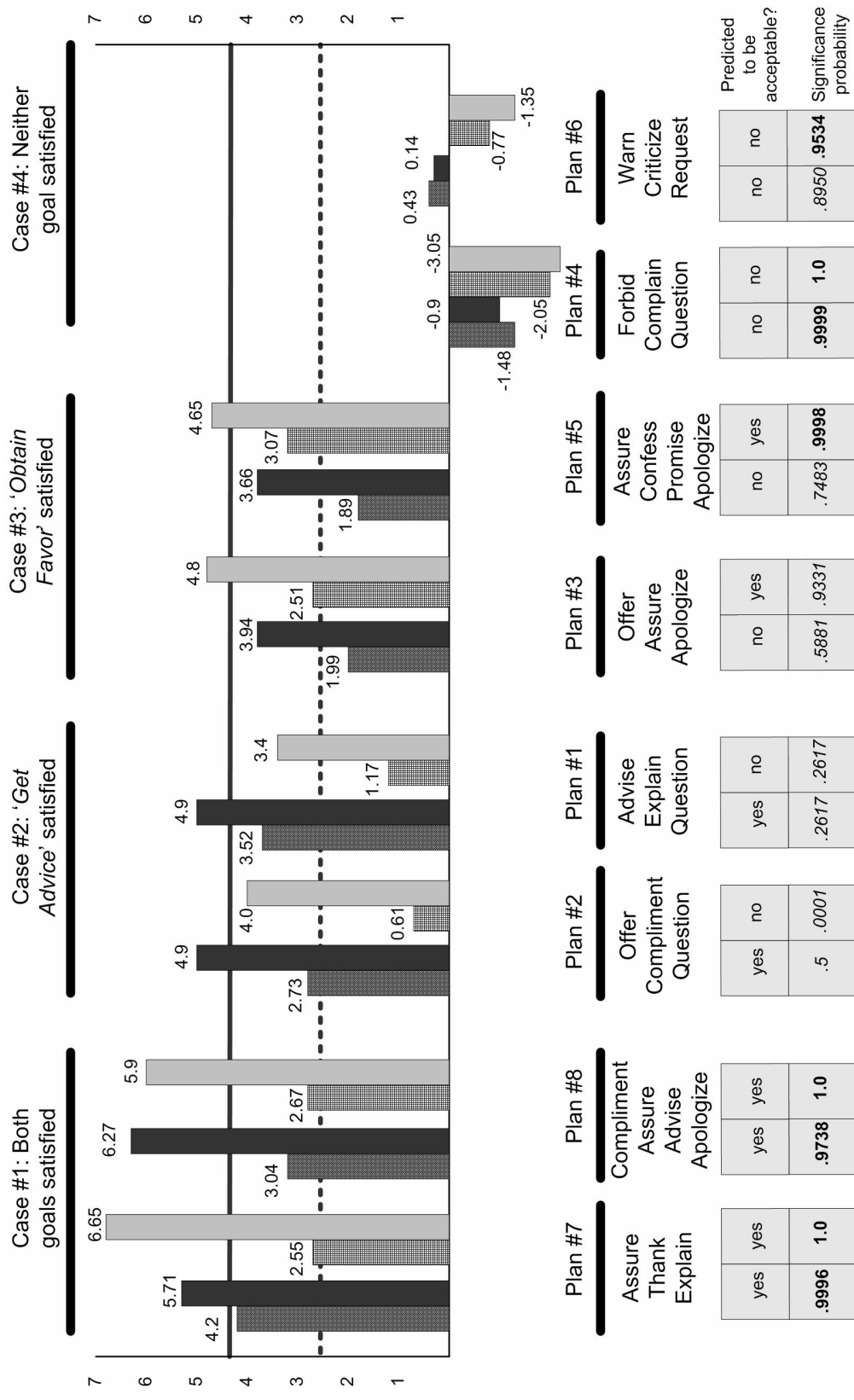


Figure 17. Four cases guiding the design of the survey, with an appropriateness threshold adjusted based on pilot results

cases of strong disagreement or weak support all occurred among the predictions having to do with plans expected to satisfy only one of the goals, just as in the pilot. That is, in both the pilot and the study, the model's ability to generate plans satisfying multiple goals through multiple acts received strong support.

We looked for opportunities for getting higher levels of support by adjusting threshold values, just as we did in the pilot. We found four such opportunities:

1. If the appropriateness threshold were lowered to 4.0 and the efficiency threshold lowered to 0.61, then the Obtain Favor prediction of plan 2 would receive support of .9999;
2. If the appropriateness threshold were raised to *just above* 4.9, or if the efficiency threshold were raised to *just above* 3.52, or both, the Get Advice prediction of plan 1 would receive support of .7383;
3. If the appropriateness threshold were lowered to 3.4 and the efficiency threshold were lowered to 1.17, then the Obtain Favor prediction of plan 1 would receive support of .7383;
4. If opportunities 1 and 3 were combined by taking the lower bounds of the proposed values (i.e., 3.4 for appropriateness and 0.61 for efficiency), then we might get the benefit of both opportunities.

Of these four opportunities, the last turned out the best, yielding strong support for nine predictions (by flipping the support for plan 2's Obtain Favor prediction), and weak support for two predictions (by flipping the support for plan 1's Obtain Favor prediction) [Figure 18]. None of the predictions received strong disagreement after adopting the fourth opportunity.

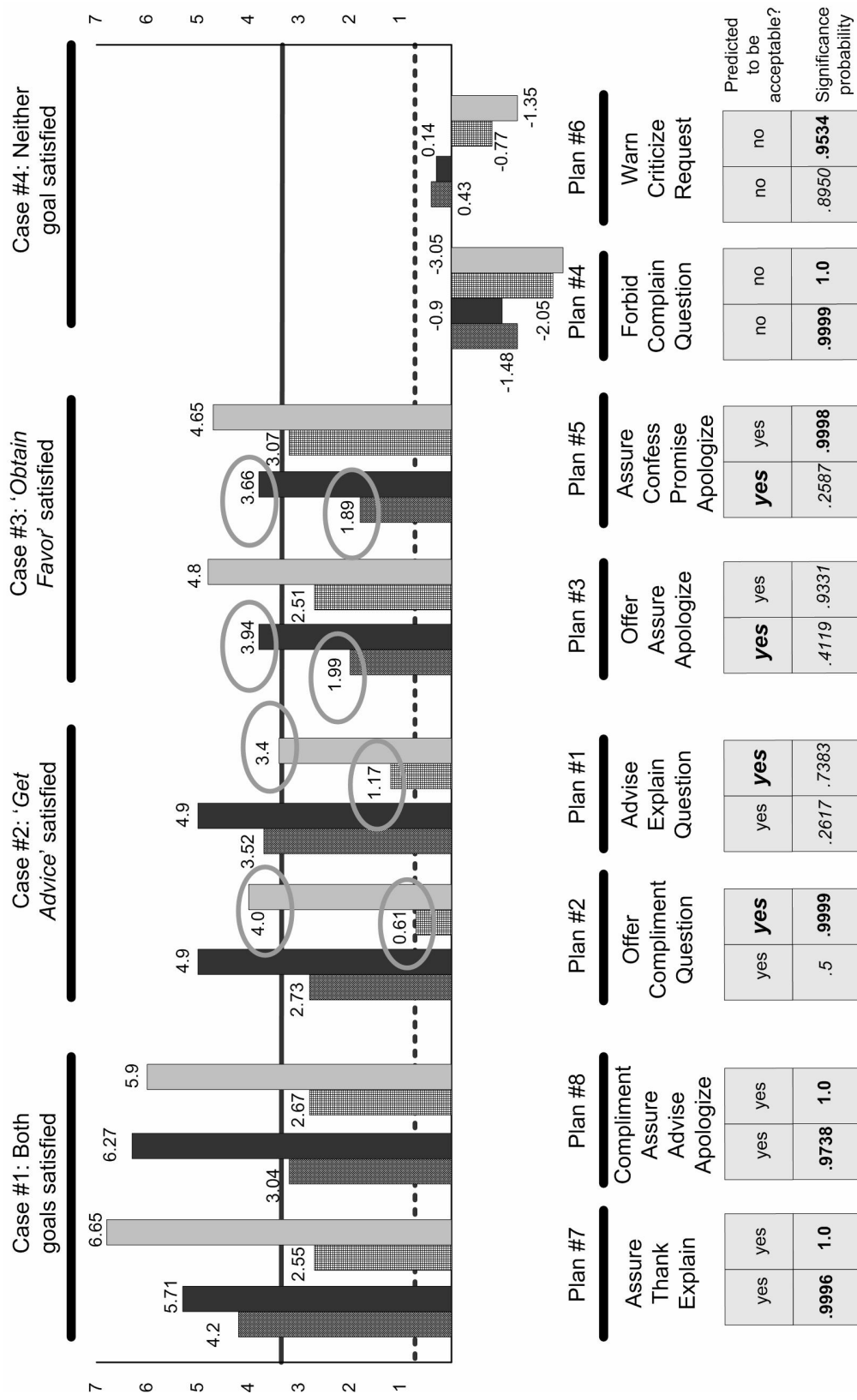


Figure 18. Survey results after adjusting both thresholds

Adopting the fourth opportunity also demonstrates that our miscalculation in using an appropriateness threshold of 4.45 in generating the plans of the survey made no difference, since by adopting the opportunity we are using a value of 3.4 for that threshold, which is even further from the value of 5.1 suggested by the pilot than our miscalculated value is.

In summary, the pilot and survey both revealed a supportive trend for the majority of their predictions (for twelve predictions and eleven predictions, respectively, out of sixteen). In particular, plans that were generated to satisfy multiple goals using multiple acts received strong support in all cases but one, where there was moderately strong support. These plans represent the claimed contribution that our tests were meant to verify, so the CSC model has been shown to be well-supported empirically.

CHAPTER 6

CONCLUSION

The Computable Social Communication (CSC) simulator is able to generate plausible plans using a wider range of social goals, speech acts, and situation conditions than those of any previous speech act planner. And its planning method is directly based on empirical findings about human communicative behavior. Furthermore, it builds on the theory behind the empirical studies, Conversational Constraint Theory, in two ways. First, CCT assumes that only one goal is being pursued and that the solution should have just one act, while CSC demonstrates how this can be extended so that multiple goals can be solved simultaneously with multiple acts. Solving for multiple goals, with simultaneous constraints on both efficiency and appropriateness, is also an advance over recent speech act planning (SAP) systems. Second, CSC ensures that its plans are more plausible than CCT's by checking the conditions of each candidate act.

CSC has limitations in common with recent SAP work; for example, it does not suggest an ordering for the acts of its plans, nor does it check for potential interference among acts. As an example of potential interference among acts, consider that acts of begging and threatening might both be useful for obtaining a favor; yet, performing these acts in quick succession might undermine the effort as a whole. Even more importantly, CSC stops planning at the act level and does not go all the way to the utterance level. Not having utterances to judge from makes evaluating the plausibility of the plans open to a fair amount of interpretive leeway, which is an impediment to objective evaluation.

Our claim that CSC's plans are plausible rests to a large extent on the data that Kellermann collected and the theory she supported through analysis of that data. But our extensions to her work need their own support. We conducted a human-subjects study (with a pilot) in pursuit of that support. In summary, the pilot and survey both showed a supportive trend for the majority of their predictions (for twelve predictions and eleven predictions, respectively, out of sixteen). In particular, plans that were generated to satisfy multiple goals using multiple acts received strong support in all cases but one, where there was moderately strong support. These plans represent the claimed contribution that our tests were meant to verify, so the CSC model has been shown to be well-supported empirically.

Future directions

As already mentioned, CSC shares some of the limitations of speech act planning work in general. Prime areas for future work would be to overcome these limitations, particularly our reliance on speech act types as the final output of planning rather than fully realized utterances. We adopted this limitation knowingly to avoid introducing bias during evaluation, since realizer output is often stilted. But some reviewers have suggested that limiting the scope of the planner to a particular domain of discourse and using utterances from a large corpus in that domain as the units of planning might help avoid the problem. Still, there would be significant challenges. In particular, it would require a significant amount of work to adapt CSC to provide enough information to allow the planner/realizer to do its task. The work would entail re-representing the goals, acts, and conditions of

CSC from their current atomic values into a propositional form, and then defining the role bindings among goals, act effects, and act conditions. Such representation work is one of the defining characteristics of SAP research, but it has been applied to a relatively narrow set of acts and conditions in that field so far. As an example of a typical SAP speech act representation, consider this “plan operator” format for the act of “requesting,”

Name: Requesting(Speaker, Hearer, ActionRequested)
Effects: intendsToDo(Hearer, ActionRequested)
Conditions: hasGoal(Speaker, intendsToDo(Hearer, ActionRequested)) and
 believes(Speaker, isCapableOfDoing(Hearer, ActionRequested))

While the analysis of “requesting” represented by this plan operator does not list all of the conditions that Wierzbicka's dictionary does for the same act, it does make each condition more precise and clarifies how role bindings made while pattern-matching against one propositional condition would add meaningful constraints on the other propositions. For example, when a planner is able to match a goal to the effect of this operator, the role bindings made from that matching attempt would propagate to the conditions, so that when the planner checks the conditions against its knowledge of the situation, and asks specifically, “Is the hearer capable of doing the requested action?,” it will know what the action is. Compare that process to the way acts are chosen in CCT and CSC, where there is no role binding, and consequently the conditions are not instantiated with the specifics of the goal, and the planned acts are not instantiated with the specifics of either the goal or the satisfied conditions. Clearly, if the goal type “Persuade to Fulfill Obligation” were matched to the specific goal, “persuade driver to

use his turn signal,” for example, the bindings from such a match would be very helpful in determining if the conditions of the planned acts, “complimenting” and “suggesting,” could be met. (That is, the bindings may overconstrain the conditions and prevent the acts from being used, or may guide the planner to facts that would be especially suitable for complimenting about and suggesting about.)

Such plan operators could be improved in several ways. First, once one notices that there is a strong resemblance between Wierzbicka's regular conditions and the “Conditions” field shown here, and that Wierzbicka's illocutionary purpose condition corresponds to the “Effects” field, it seems obvious to add her third and last condition type, the dictum, which describes how the act should be expressed. Second, recent SAP work by Mark Lee argues convincingly that,

Precomputing all the effects of any speech act is unrealistic, since any speech act has a set of perlocutionary effects depending on the context or belief states of both the speaker and hearer, which are in addition to [the] conventional illocutionary effects [of the act].
[Lee 1998, p. 155]

In other words, one cannot expect to be able to list all of the effects of an act, in advance, in a plan operator. Instead, Lee proposes that in place of an effects list, planners instead apply an inference rule so that every condition is ascribed as a belief the hearer will have about what the speaker believes. Such a strategy is virtually identical to the way Wierzbicka's illocutionary purposes relate to her regular conditions, since she formulates all purposes as “I say this, in this way, in order to make the hearer aware of X,” where X is always one of her regular conditions. The intuition that Lee and Wierzbicka are

appealing to is: when a hearer realizes that a speaker has performed a certain act, the hearer is likely to assume that all of the act's conditions must have been satisfied, unless the hearer has an explicit belief to the contrary.

The final step in Lee's proposal is that there be a set of inference rules that simulate all of the conclusions a hearer would draw from his new beliefs triggered by what the speaker has said. For example, these rules might start from the belief that the speaker wants the hearer to intend to do something, and the inference chain might yield an actual intent on the part of the hearer to do the act the speaker mentioned wanting. Therefore, one might re-represent "requesting" in this way,

Name: Requesting(Speaker, Hearer, ActionRequested)
Effects: *...after inferencing:* intendsToDo(Hearer, ActionRequested)
Conditions: hasGoal(Speaker, intendsToDo(Hearer, ActionRequested)) and
believes(Speaker, isCapableOfDoing(Hearer, ActionRequested))
Dictum: say(Speaker, hasGoal(Speaker, intendsToDo(Hearer, ActionRequested)))

In such an approach to modeling strategic interpersonal communication, the inference rules are the critical piece. They are the reason we have gone into a deep exploration of this potential area for future work, because we want to explain why an AI approach like CSC cannot go further without a new kind of empirical data to guide its processing. The problem is that there is no empirical data to tell us what the rules should be, and without it we are left with our introspections, which are difficult to evaluate objectively. To make clear what kind of inference chains we have in mind, consider the difference in propositional meaning between CCT's 13 goal types and the act effects of its 56 acts --

none of the acts has a direct effect that matches any of those goals, so there must be some chain of inference that connects them, so that the particulars of the goal can be propagated to the act conditions and the dicta. Furthermore, although we have *empirically*-determined values for the relevance of each act to each goal (in terms of efficiency and appropriateness), it would be more satisfying from a theory-building standpoint to have *analytic* reasons to explain each relevance value.

The degree of fit between each rule in an inference chain might provide analytic support for act scores. Our earlier work on strategic communication [Pautler 1997; Pautler & Quilici 1998] was focused on the exact problem of identifying such inference chains, and it attempted to do so by synthesizing many lines of research from social psychology, communication, and pragmatics. But there were so many parts in the synthesis, and so many gaps filled through introspection, that it was difficult to evaluate objectively. We propose, now that we have identified the lack of well-grounded accounts of inference chains as a fundamental limitation on generating utterances from abstract plans, that empirical studies be done to uncover such chains, and that communication and social psychology are better-suited to undertake such studies than SAP or AI is. A partnership where the social discipline provided empirical data and theory, and the computer discipline provided representation and simulation, would help both disciplines advance faster than the current situation of occasional meetings in the literature.

There are also minor ways that CSC could be improved. As already mentioned, our attempt to approximate a default set of origin coordinates plus deltas for calculating the

effect of various situation factors on thresholds did not pan out. However, CCT's author indicates that this is an area planned for future work on her side [Kellermann 2006, personal communication]. So, the answer may yet come from communication, although the threshold values indicated by our survey analysis provide good working assumptions in the meantime.

Another of our assumptions worth verifying is the use of addition to combine scores before comparing to the thresholds; perhaps other n-ary functions would match human judgments even better. And an assumption implicit in the use of thresholds is that plans can only be acceptable or not acceptable with no shading in between. There should be a human participant study to determine if plans that fall just short of the thresholds are judged to be more acceptable than those that fall well short.

Another limitation is that every act type in CSC can appear at most once in a plan, since each act is represented by a variable whose value must be in $\{0,1\}$ in our inequalities. The range of allowed values could be liberalized to any nonnegative integer, but the fact that plans could then contain, say, multiple “excuse” acts highlights the problem that there is no guidance about what each “excuse” should be about. (That is, to answer the “aboutness” question, one would need an account of inference chains as discussed above.)

A final limitation is that there is no account in CSC (nor in CCT) for the effect that factors such as act order, language style, or discourse history can have on interpersonal interactions. One such factor that we find especially fruitful for thought is the variety of effects that not speaking at all can have. In this case, the “act” involves no performance

effort, has no propositional content, and has no styling (although there may be associated body language). But this act has a variety of effects, which is evident from the variety of names given to it in everyday talk, such as “pregnant pauses,” “stonewalling,” “giving the silent treatment,” and so on. Which of the effects follows from the act seems to depend solely on the conditions in which it is used, such as the discourse history. We believe that the act of saying nothing, as strange as it is, would not be a problem for CSC and fits the scheme well, just as 56 acts and 13 goal types already do. And perhaps discourse history could be fit into the scheme of situation factors without too much violence.

In short, we believe that CSC has the potential to assimilate and clarify many issues in the modeling of strategic social communication. We have shown how an innovative theory from communication fits well with speech act planning, and how further integration could be done. Finally, by implementing the theoretical work of social disciplines in computer simulations, we enable rapid testing of thought experiments and demonstrate the value of deeper collaboration across disciplines.

APPENDIX A

A COMPARISON OF SCORES ACROSS TWO STUDIES

1994 strategy name	Sample utterance	Effi- ciency	Appro- priate- ness	2004 similar act	Goal inferred from utterance	Effi- ciency	Appro- priate- ness
Assertion	She said: "Play with me now!"	1.23	6.72	Order	Share Time Together	3.35	2.90
	I just gave him a command and he did it.	2.16	6.63		?	?	?
	I said: "Don't ask me any questions. Just do it and do it now!"	1.26	6.11		?	?	?
Benefit target	I got Tony to start running five days a week by pointing out that it was for his own good.	5.27	4.16	Advise	Provide Guidance	5.84	5.95
	You point out to your friend that if he accompanies you to the Midwest then he will have the opportunity to explore a part of the country with which he is unfamiliar.	5.65	3.19				
	To get Ben to comply I said: "You will benefit if you do."	4.04	4.44				

Challenge	I said: "Why don't you go on and do it? Huh? Huh?"	1.57	4.72	Challenge	?	?	?
	To get Larry to paint the house, I bet him that he couldn't.	3.06	3.84		Obtain Favor	3.83	4.35
	I dared Mark to ask Susan out for a date.	3.04	4.00		?	?	?
Compliment	I said: "I would like your advice on making up with Tanya because I think you handle these situations well."	6.63	4.79	Compliment	Get Advice	5.05	6.14
	I compliment him so he'll do it.	5.55	2.91		?	?	?
	I would like your help because you are so capable.	5.91	5.22		Obtain Favor	4.44	5.90
Criticize	He said: "You're so lazy, you never want to do anything. Why don't you go out for a change?"	1.83	4.20	Criticize	Stop Annoying Habit	2.28	2.06
	Aaron started showing up on time because I criticized him for not doing it.	2.52	4.09		Fulfill Obligation	2.80	3.45
	John told me I was such a fat cow that I really ought to eat less.	1.22	5.33		Stop Annoying Habit	2.28	2.06
Direct request	Pete asked Bob: "Could I borrow your notes?"	5.40	6.60	Request	Obtain Permission	5.00	5.75
	I would just ask the Smiths not to cut down the tree.	4.58	5.50		Obtain Favor	4.23	4.90
	I simply tell Bob what I want.	3.57	6.56		?	?	?

Hinting	Rather than directly asking Alicia to open the window, I off-handedly mentioned that it was hot in the room.	5.23	2.48	Hint	Obtain Favor	4.28	4.70
	To get Mark from letting his friend move in I would drop subtle hints about how little space we have.	4.32	1.44				
	Instead of asking her directly if she wanted to have sex, I would turn the lights down, put some music on, offer a drink and see how things went.	4.52	2.11				
Pre-giving	I helped my father around the house and then I asked to borrow \$50.	4.67	2.56	Give	Obtain Favor	4.33	5.45
	I give him a small gift or card before I ask him to do what I want.	5.81	2.69		?	?	?
	I sent my friends who lived on the beach a Christmas present and then called up awhile later and asked if I could visit them over the holidays.	4.09	3.28		Obtain Favor	4.33	5.45
Promise	I said: "I will give you a reward if you do what I want."	4.43	5.60	Promise	?	?	?
	I told my friends I would buy pizza and beer if they helped me move.	5.71	4.56		Obtain Favor	4.06	4.90
	I'll buy you a car if you get all A's this year.	4.22	5.06		Provide Guidance	4.63	4.84
Suggest	She said: "You know, one alternative is to take the Airbus to L.A."	5.20	4.92	Suggest	Provide Guidance	5.58	6.16
	She said: "Why don't you think about joining a softball team?"	5.13	3.97				
	If I were you, I'd probably call accounting before I got started on the project.	4.74	3.94				

Threat	I said: "John, if you don't do what I want, I will punish you."	1.47	5.28	Threaten	?	?	?
	Paul threatens to reveal some deep dark embarrassing secret about Tim to someone special if he does not comply.	1.55	4.00		?	?	?
	Unless you stop being late for work, your future with this company will be in serious jeopardy.	4.00	5.72		Fulfill Obligation	2.65	1.80
Warning	Jerry said: "Be careful, she's out to get you."	3.57	5.80	Warn	Provide Guidance	4.58	4.74
	I warned him against it so he didn't do it.	4.77	5.09		?	?	?
	You might be risking your job if you get pregnant.	3.57	4.72		Provide Guidance	4.58	4.74

APPENDIX B

DOWNLOAD THE SIMULATOR

Installation instructions are included in the zip:

<http://wooden-robot.net/ComputableSocialCommunication.tar.gz>

APPENDIX C

CCT'S FACTORS DETERMINING MINIMUM PREFERRED LEVELS

These factors are presented verbatim from an unpublished manuscript by Kathy Kellermann [Kellermann 2006].

Situational Determinants

Urgency Principle: The greater the situational urgency for achieving a goal, the higher a person's minimum preferred level of efficiency. The less the situational urgency, the lower the minimum preferred level of efficiency.

Formality Principle: The more formal a situation, the higher a person's minimum preferred level of politeness. The less formal a situation, the lower the minimum preferred level of politeness.

Privacy Principle: The more public a situation, the higher a person's minimum preferred level of politeness. The more private the situation, the lower the minimum preferred level of politeness.

Relational Determinants

Bond Principle: The greater a person's concern about the current status of his/her relational bond with a conversational partner, the higher the person's minimum preferred level of politeness. The less concern about the current status of the relational bond, the lower the minimum preferred level of politeness.

Position Principle: The lower a person's relational position relative to a conversational partner, the higher the person's minimum preferred level of politeness and the lower the person's minimum preferred level of efficiency. The higher the relational position, the lower the minimum preferred level of politeness and the higher the minimum preferred level of efficiency.

Personal Determinants

Social Orientation Principle: The more a person systematically focuses on social cues, the higher the person's minimum preferred level of politeness. The less focus on social cues, the lower the minimum preferred level of politeness.

Goal Orientation Principle: The more a person systematically focuses on goal achievement, the higher the person's minimum preferred level of efficiency. The less focus on goal achievement, the lower the preferred level of efficiency.

Interactive Determinants

Matching Principle: The more polite or more efficient a conversational partner, the higher a person's minimum preferred level of politeness or efficiency, respectively. The less polite or efficient a partner, the lower the minimum preferred level of politeness or efficiency, respectively.

Self Expectation Principle: The more polite or efficient a person thinks he/she should be in the conversation, the higher the person's minimum preferred level of politeness or efficiency, respectively. The less polite or efficient a person thinks he/she should be, the lower the person's minimum preferred level of politeness or efficiency, respectively.

Other Expectation Principle: The more politeness or efficiency a person thinks the conversational partner expects of him/her, the higher the person's minimum preferred level of politeness or efficiency, respectively. The less politeness or efficiency a person thinks the partner expects, the lower the person's minimum preferred level of politeness or efficiency, respectively.

APPENDIX D

TRANSCRIPT FOR “THE RECKLESS DRIVER”

SCENARIO: You are an in-car automated driving advisor, your driver is a male engineer who dislikes being told what to do, and you want to persuade him to use his turn signal when changing lanes. We assume the goal here is an instance of persuading the driver to Fulfill an Obligation.

The planner's first recommendation was the acts {thank apologize}. We could not think of anything in the scenario to support the conditions of these acts, so we rejected both.

The planner's second recommendation was the acts {compliment praise}. Complimenting only required finding something good about the driver to remark on, and we saw no harm in saying he is a “good” driver. But giving praise required that the driver have done something worth rewarding, and we could not think of a single thing, so we rejected that act.

The planner's third recommendation was {compliment suggest}, where it re-used Complimenting because it had already been accepted as a possibility. The other act, Suggesting, required that there be some action that would be to the driver's benefit that we could bring to his attention. Using one's turn signal has the potential benefit of avoiding being hit by less attentive drivers, so that act was also accepted.

In the transcript below, the simulator's prompts are shown in plain text, our responses are shown in **bold**, and output from the GLPK solver is *italicized*.

```
distrib> perl cscPlanner.pl
```

```
-----  
These are the goal types for which I can generate speech act plans:
```

- 1 : Get_Advice
- 2 : Move_Relationship_Forward
- 3 : Fulfill_Obligation
- 4 : Initiate_Relationship
- 5 : Provide_Guidance
- 6 : Get_Date
- 7 : Obtain_Permission
- 8 : Obtain_Favor
- 9 : Share_Time_Together
- 10 : End_Relationship
- 11 : Obtain_Information
- 12 : Stop_Annoying_Habit
- 13 : Change_Opinion

```
Enter goal #s: 3
```

```
-----  
Do you know your min preferred levels of efficiency and politeness for each goal in this situation?  
[yn]: n
```

```
Defaults: min preferred efficiency = 0.5; min preferred politeness = 0.5
```

```
-----  
How goal-oriented are you in general? (not at all) [1...7] (very)  
Enter #: 4
```

```
Update: min preferred efficiency = 0.5; min preferred politeness = 0.5
```

```
-----  
How efficient is your audience? (not at all) [1...7] (very)  
Enter #: 7
```

```
Update: min preferred efficiency = 2; min preferred politeness = 0.5
```

```
-----  
How formal are the circumstances in which you are speaking? (not at all) [1...7] (very)  
Enter #: 1
```

```
Update: min preferred efficiency = 2; min preferred politeness = -1
```


How urgent is your goal? (not at all) [1...7] (very)
Enter #: 6

Update: min preferred efficiency = 3; min preferred politeness = -1

How polite does your audience expect you to be in this conversation? (not at all) [1...7] (very)
Enter #: 5

Update: min preferred efficiency = 3; min preferred politeness = -0.5

How polite is your audience? (not at all) [1...7] (very)
Enter #: 4

Update: min preferred efficiency = 3; min preferred politeness = -0.5

How efficient does your audience expect you to be in this conversation? (not at all) [1...7] (very)
Enter #: 4

Update: min preferred efficiency = 3; min preferred politeness = -0.5

How efficient do you think you should be in this conversation? (not at all) [1...7] (very)
Enter #: 5

Update: min preferred efficiency = 3.5; min preferred politeness = -0.5

How attentive are you to social cues in general? (not at all) [1...7] (very)
Enter #: 6

Update: min preferred efficiency = 3.5; min preferred politeness = 0.5

What is your social position relative to your audience? (much lower) [1...7] (much higher)
Enter #: 1

Update: min preferred efficiency = 2; min preferred politeness = 2

How concerned are you for the status of your relational bond with your audience? (not at all) [1...7] (very)
Enter #: 5

Update: min preferred efficiency = 2; min preferred politeness = 2.5

How private is the situation you are speaking about? (not at all) [1...7] (very)
Enter #: 5

Update: min preferred efficiency = 2; min preferred politeness = 3

How polite do you think you should be in this conversation? (not at all) [1...7] (very)
Enter #: 7

Update: min preferred efficiency = 2; min preferred politeness = 4.5

```
-----
lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Fulfill_Obligation-2e-4.5p-v1.lp'...
lpx_read_cpxlp: 2 rows, 56 columns, 112 non-zeros
lpx_read_cpxlp: 56 integer columns, all of which are binary
lpx_read_cpxlp: 66 lines were read
lpx_simplex: original LP has 2 rows, 56 columns, 112 non-zeros
lpx_simplex: presolved LP has 2 rows, 56 columns, 112 non-zeros
lpx_adv_basis: size of triangular part = 2
  0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)
  5: objval = 4.073170732e+00 infeas = 0.000000000e+00 (0)
* 5: objval = 4.073170732e+00 infeas = 0.000000000e+00 (0)
* 11: objval = 1.760000000e+00 infeas = 0.000000000e+00 (0)
OPTIMAL SOLUTION FOUND
Integer optimization begins...
Objective function is integral
+ 11: mip = not found yet >= -inf (1; 0)
+ 20: mip = 2.000000000e+00 >= 2.000000000e+00 0.0% (10; 0)
+ 20: mip = 2.000000000e+00 >= tree is empty 0.0% (0; 19)
INTEGER OPTIMAL SOLUTION FOUND
Time used: 0.0 secs
Memory used: 0.1M (96056 bytes)
lpx_print_mip: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...
-----
```

There is an optimal solution containing these 2 act(s): thank apologize
For goal Fulfill_Obligation,
achieved efficiency: 3.25
achieved politeness: 4.55

But first we have to verify that the conditions of these acts are all satisfied...

[thank] has the following preconditions:

- * There is some act that your audience has already done
- * The event benefitted you
- * You feel something good about your audience because of what they did
- * Your audience will feel good knowing what you think about them

Are all of these conditions satisfied in the current situation? [ynq]: n

[apologize] has the following preconditions:

- * There is some act that you have already done
- * You feel bad about it
- * Your audience might feel something bad about you because of it

Are all of these conditions satisfied in the current situation? [ynq]: n

Searching for a new plan...

lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Fulfill_Obligation-2e-4.5p-v2.lp'...

lpx_read_cpxlp: 2 rows, 54 columns, 108 non-zeros

lpx_read_cpxlp: 54 integer columns, all of which are binary

lpx_read_cpxlp: 64 lines were read

lpx_simplex: original LP has 2 rows, 54 columns, 108 non-zeros

lpx_simplex: presolved LP has 2 rows, 54 columns, 108 non-zeros

lpx_adv_basis: size of triangular part = 2

0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)

4: objval = 2.550000000e+00 infeas = 0.000000000e+00 (0)

* 4: objval = 2.550000000e+00 infeas = 0.000000000e+00 (0)

* 9: objval = 1.844444444e+00 infeas = 0.000000000e+00 (0)

OPTIMAL SOLUTION FOUND

Integer optimization begins...

Objective function is integral

+ 9: mip = not found yet >= -inf (1; 0)

+ 46: mip = 3.000000000e+00 >= 2.000000000e+00 33.3% (37; 0)

+ 48: mip = 2.000000000e+00 >= 2.000000000e+00 0.0% (12; 51)

+ 48: mip = 2.000000000e+00 >= tree is empty 0.0% (0; 75)

INTEGER OPTIMAL SOLUTION FOUND

Time used: 0.0 secs

Memory used: 0.1M (96216 bytes)

lpx_print_mip: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...

There is an optimal solution containing these 2 act(s): compliment praise

For goal Fulfill_Obligation,

achieved efficiency: 3.1

achieved politeness: 4.55

But first we have to verify that the conditions of these acts are all satisfied...

[compliment] has the following preconditions:

* There is something good about your audience

* Your audience will feel good knowing what you think about them

Are all of these conditions satisfied in the current situation? [ynq]: **y**

[praise] has the following preconditions:

* There is something good about your audience

* Your audience should be rewarded

Are all of these conditions satisfied in the current situation? [ynq]: **n**

Searching for a new plan...

lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Fulfill_Obligation-2e-4.5p-v3.lp'...

lpx_read_cpxlp: 2 rows, 53 columns, 106 non-zeros

lpx_read_cpxlp: 53 integer columns, all of which are binary

lpx_read_cpxlp: 63 lines were read

lpx_simplex: original LP has 2 rows, 53 columns, 106 non-zeros

lpx_simplex: presolved LP has 2 rows, 53 columns, 106 non-zeros

lpx_adv_basis: size of triangular part = 2

0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)

4: objval = 2.550000000e+00 infeas = 0.000000000e+00 (0)

* 4: objval = 2.550000000e+00 infeas = 0.000000000e+00 (0)

* 9: objval = 1.844444444e+00 infeas = 0.000000000e+00 (0)

OPTIMAL SOLUTION FOUND

Integer optimization begins...

Objective function is integral

```
+ 9: mip = not found yet >= -inf (1; 0)
+ 45: mip = 3.000000000e+00 >= 2.000000000e+00 33.3% (36; 0)
+ 47: mip = 2.000000000e+00 >= 2.000000000e+00 0.0% (11; 51)
+ 47: mip = 2.000000000e+00 >= tree is empty 0.0% (0; 73)
```

INTEGER OPTIMAL SOLUTION FOUND

Time used: 0.0 secs

Memory used: 0.1M (96056 bytes)

lpx_print_mip: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...

There is an optimal solution containing these 2 act(s): compliment suggest
For goal Fulfill_Obligation,
achieved efficiency: 2.95
achieved politeness: 4.55

But first we have to verify that the conditions of these acts are all satisfied...

You have already told us that all of the conditions of [compliment] are satisfied.

[suggest] has the following preconditions:

- * There is a certain event that would be good for your audience
- * Your audience might not be aware of the opportunity
- * Your audience could make it happen, and has no resistance, but currently isn't intending to

Are all of these conditions satisfied in the current situation? [ynq]: y

The final solution contains these 2 act(s): compliment suggest

APPENDIX E

TRANSCRIPT FOR “THE RECKLESS DRIVER AND THE IMPOLITE INTERFACE”

This example demonstrates the sensitivity of CSC to situational factors. To make that point, we have reused the “Reckless Driver” scenario but have changed our answer to the question, “How polite does your audience expect you to be in this conversation?” from 5 to 3. That is, the only difference between the two scenarios is that in this case we assert that the driver expects the interface to be less polite. By changing our answer to the questionnaire in this way, the final value for minimum preferred efficiency remains at 2, but the final value for minimum preferred level of politeness goes down from 4.5 in the previous scenario to 3.5 (highlighted in **bold underlining** below).

With this slight change of the politeness threshold, the planner's first suggested plan has changed from {thank apologize} to {thank question}. In the planner session for the previous scenario, and for this one, the planner came across {thank question} in the search space before coming across {thank apologize}, but in the previous session, the politeness threshold was too high to be met by {thank question}, so it kept searching. Since that threshold is low enough in the current scenario to be met by {thank question}, search did not have to go further. However, there is no reason why {thank question} should be found first, given the way we have set up the search space, so this difference in plans is just a fortuitous coincidence that happens to show the influence of situation factors.

We rejected the “question” act because its conditions could not plausibly be met, based on our understanding of the scenario, and the next recommended plan was {disclose compliment}. This second recommendation is similar to the second recommendation of the previous scenario, {praise compliment}. It is worth noting that the difference in act choices across both recommendations is entirely explainable by the politeness scores of the acts for the Fulfill Obligation goal: “apologize” scores at 2.05 versus 1.10 for “question,” and “praise” scores at 1.95 versus 0.90 for “disclose.”

We also rejected “disclose” because its conditions could not be met, but we found the next recommendation, {compliment tell}, acceptable. This third recommendation is similar to third recommendation of the previous scenario, {compliment suggest}, and again the difference can be explained by the difference in act politeness scores -- 1.95 for “suggest” versus 0.90 for “tell.” We imagined that the conditions of “tell” could be satisfied by asserting straightforwardly that, if the driver kept changing lanes without signaling, he would eventually have a collision. By comparison, the “suggest” act of the previous scenario gave similar information but highlighted the benefit to the driver of complying. We assume the “compliment” act of both plans would be instantiated the same way in both plans. When interpreted in this way, the {compliment tell} plan of the current scenario seems less polite than the {compliment suggest} plan of the previous scenario, which fits what one would expect from lowering the politeness threshold in this scenario.

If one changed the scenario even more, one could get recommended plans that have fewer acts or more acts, not just a change in acts. For example, a radical increase in the politeness threshold would require adding acts that could raise the politeness of the plan while not lowering its efficiency below the efficiency threshold.

```
distrib> perl cscPlanner.pl
```

```
-----  
These are the goal types for which I can generate  
speech act plans:
```

```
1 : Get_Advice  
2 : Move_Relationship_Forward  
3 : Fulfill_Obligation  
4 : Initiate_Relationship  
5 : Provide_Guidance  
6 : Get_Date  
7 : Obtain_Permission  
8 : Obtain_Favor  
9 : Share_Time_Together  
10 : End_Relationship  
11 : Obtain_Information  
12 : Stop_Annoying_Habit  
13 : Change_Opinion  
Enter goal #s: 3
```

```
-----  
Do you know your min preferred levels of efficiency and politeness for each goal in this situation?  
[yn]: n
```

```
Defaults: min preferred efficiency = 0.5; min preferred politeness = 0.5
```

```
-----  
How goal-oriented are you in general? (not at all) [1...7] (very)  
Enter #: 4
```

```
Update: min preferred efficiency = 0.5; min preferred politeness = 0.5
```

```
-----  
How efficient is your audience? (not at all) [1...7] (very)  
Enter #: 7
```

```
Update: min preferred efficiency = 2; min preferred politeness = 0.5
```

How formal are the circumstances in which you are speaking? (not at all) [1...7] (very)
Enter #: **1**

Update: min preferred efficiency = 2; min preferred politeness = -1

How urgent is your goal? (not at all) [1...7] (very)
Enter #: **6**

Update: min preferred efficiency = 3; min preferred politeness = -1

How polite does your audience expect you to be in this conversation? (not at all) [1...7] (very)
Enter #: **3**

Update: min preferred efficiency = 3; min preferred politeness = **-1.5**

How polite is your audience? (not at all) [1...7] (very)
Enter #: **4**

Update: min preferred efficiency = 3; min preferred politeness = -1.5

How efficient does your audience expect you to be in this conversation? (not at all) [1...7] (very)
Enter #: **4**

Update: min preferred efficiency = 3; min preferred politeness = -1.5

How efficient do you think you should be in this conversation? (not at all) [1...7] (very)
Enter #: **5**

Update: min preferred efficiency = 3.5; min preferred politeness = -1.5

How attentive are you to social cues in general? (not at all) [1...7] (very)
Enter #: **6**

Update: min preferred efficiency = 3.5; min preferred politeness = -0.5

What is your social position relative to your audience? (much lower) [1...7] (much higher)
Enter #: **1**

Update: min preferred efficiency = 2; min preferred politeness = 1

How concerned are you for the status of your relational bond with your audience? (not at all) [1...7] (very)
Enter #: **5**

Update: min preferred efficiency = 2; min preferred politeness = 1.5

How private is the situation you are speaking about? (not at all) [1...7] (very)
Enter #: 5

Update: min preferred efficiency = 2; min preferred politeness = 2

How polite do you think you should be in this conversation? (not at all) [1...7] (very)
Enter #: 7

Update: min preferred efficiency = 2; min preferred politeness = 3.5

```
lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Fulfill_Obligation-2e-3.5p-v1.lp'...
lpx_read_cpxlp: 2 rows, 56 columns, 112 non-zeros
lpx_read_cpxlp: 56 integer columns, all of which are binary
lpx_read_cpxlp: 66 lines were read
lpx_simplex: original LP has 2 rows, 56 columns, 112 non-zeros
lpx_simplex: presolved LP has 2 rows, 56 columns, 112 non-zeros
lpx_adv_basis: size of triangular part = 2
0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)
3: objval = 2.602791878e+00 infeas = 0.000000000e+00 (0)
* 3: objval = 2.602791878e+00 infeas = 0.000000000e+00 (0)
* 7: objval = 1.360000000e+00 infeas = 0.000000000e+00 (0)
OPTIMAL SOLUTION FOUND
Integer optimization begins...
Objective function is integral
+ 7: mip = not found yet >= -inf
(1; 0)
+ 34: mip = 2.000000000e+00 >= 2.000000000e+00
0.0% (28; 0)
+ 34: mip = 2.000000000e+00 >= tree is empty
0.0% (0; 55)
INTEGER OPTIMAL SOLUTION FOUND
Time used: 0.0 secs
Memory used: 0.1M (96536 bytes)
lpx_print_mip: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...
```

There is an optimal solution containing these 2 act(s): thank question
For goal Fulfill_Obligation,
achieved efficiency: 3.25
achieved politeness: 3.6

But first we have to verify that the conditions of these acts are all satisfied...

[thank] has the following preconditions:

- * There is some act that your audience has already done
 - * The event benefitted you
 - * You feel something good about your audience because of what they did
 - * Your audience will feel good knowing what you think about them
- Are all of these conditions satisfied in the current situation? [ynq]: n

[question] has the following preconditions:

- * There is something you suspect you don't know
- * It is something bad
- * Your audience might have info that would help you find out
- * Your audience could make it happen, but currently isn't intending to (and might be resistant)
- * You have a right to know
- * Your audience is obliged to comply

Are all of these conditions satisfied in the current situation? [ynq]: **n**

Searching for a new plan...

lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Fulfill_Obligation-2e-3.5p-v2.lp'...

lpx_read_cpxlp: 2 rows, 54 columns, 108 non-zeros

lpx_read_cpxlp: 54 integer columns, all of which are binary

lpx_read_cpxlp: 64 lines were read

lpx_simplex: original LP has 2 rows, 54 columns, 108 non-zeros

lpx_simplex: presolved LP has 2 rows, 54 columns, 108 non-zeros

lpx_adv_basis: size of triangular part = 2

0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)

3: objval = 1.939566705e+00 infeas = 0.000000000e+00 (0)

** 3: objval = 1.939566705e+00 infeas = 0.000000000e+00 (0)*

** 7: objval = 1.400000000e+00 infeas = 0.000000000e+00 (0)*

OPTIMAL SOLUTION FOUND

Integer optimization begins...

Objective function is integral

+ 7: mip = not found yet >= -inf

(1; 0)

+ 35: mip = 2.000000000e+00 >= 2.000000000e+00

0.0% (28; 0)

+ 35: mip = 2.000000000e+00 >= tree is empty

0.0% (0; 55)

INTEGER OPTIMAL SOLUTION FOUND

Time used: 0.0 secs

Memory used: 0.1M (96216 bytes)

lpx_print_mip: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...

There is an optimal solution containing these 2 act(s): disclose compliment

For goal Fulfill_Obligation,

achieved efficiency: 2.7

achieved politeness: 3.5

But first we have to verify that the conditions of these acts are all satisfied...

[disclose] has the following preconditions:

- * There is something your audience doesn't know
- * Your audience isn't likely to discover it on their own
- * Your audience doesn't know because it was hidden from them
- * Your audience half-believes it's not true
- * It is best for your audience if they know

Are all of these conditions satisfied in the current situation? [ynq]: **n**

[compliment] has the following preconditions:
* There is something good about your audience
* Your audience will feel good knowing what you think about them
Are all of these conditions satisfied in the current situation? [ynq]: y

Searching for a new plan...

lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Fulfill_Obligation-2e-3.5p-v3.lp'...

lpx_read_cpxlp: 2 rows, 53 columns, 106 non-zeros

lpx_read_cpxlp: 53 integer columns, all of which are binary

lpx_read_cpxlp: 63 lines were read

lpx_simplex: original LP has 2 rows, 53 columns, 106 non-zeros

lpx_simplex: presolved LP has 2 rows, 53 columns, 106 non-zeros

lpx_adv_basis: size of triangular part = 2

0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)

3: objval = 1.939566705e+00 infeas = 0.000000000e+00 (0)

** 3: objval = 1.939566705e+00 infeas = 0.000000000e+00 (0)*

** 7: objval = 1.400000000e+00 infeas = 0.000000000e+00 (0)*

OPTIMAL SOLUTION FOUND

Integer optimization begins...

Objective function is integral

+ 7: mip = not found yet >= -inf

(1; 0)

+ 35: mip = 2.000000000e+00 >= 2.000000000e+00

0.0% (28; 0)

+ 35: mip = 2.000000000e+00 >= tree is empty

0.0% (0; 55)

INTEGER OPTIMAL SOLUTION FOUND

Time used: 0.0 secs

Memory used: 0.1M (96056 bytes)

lpx_print_mip: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...

There is an optimal solution containing these 2 act(s): compliment tell

For goal Fulfill_Obligation,

achieved efficiency: 2.4

achieved politeness: 3.5

But first we have to verify that the conditions of these acts are all satisfied...

You have already told us that all of the conditions of [compliment] are satisfied.

[tell] has the following preconditions:

* There is something your audience doesn't know

* Your audience isn't likely to discover it on their own

* Your audience is likely to believe you

Are all of these conditions satisfied in the current situation? [ynq]: y

The final solution contains these 2 act(s): compliment tell

APPENDIX F

TRANSCRIPT FOR “THE EMBARRASSING HUSBAND”

SCENARIO: You and your husband are a middle-aged couple invited to a fancy dinner party. Your husband is slurping his soup and getting some in his lap. We assume stopping him from slurping his soup is an instance of `Stop_Annoying_Habit`, and that helping him avoid staining his trousers is an instance of persuading by `Providing_Guidance`.

The planner suggests three acts, {give ask inform}, which are plausible when instantiated like this:

- Give:** Unfold your husband's napkin and hand it to him.
- Ask:** Ask your husband if he is aware that his trousers are being stained.
- Inform:** Tell your husband that he is slurping loudly and embarrassing you.

In the transcript below, the simulator's prompts are shown in plain text, our responses are shown in **bold**, and output from the GLPK solver is *italicized*.

distrib> perl cscPlanner.pl

These are the goal types for which I can generate speech act plans:

- 1 : Get_Advice
- 2 : Move_Relationship_Forward
- 3 : Fulfill_Obligation
- 4 : Initiate_Relationship
- 5 : Provide_Guidance
- 6 : Get_Date
- 7 : Obtain_Permission
- 8 : Obtain_Favor
- 9 : Share_Time_Together
- 10 : End_Relationship
- 11 : Obtain_Information
- 12 : Stop_Annoying_Habit
- 13 : Change_Opinion

Enter goal #s: **12 5**

Do you know your min preferred levels of efficiency and politeness for each goal in this situation?
[yn]: **n**

Defaults: min preferred efficiency = 0.5; min preferred politeness = 0.5

How goal-oriented are you in general? (not at all) [1...7] (very)
Enter #: **4**

Update: min preferred efficiency = 0.5; min preferred politeness = 0.5

How efficient is your audience? (not at all) [1...7] (very)
Enter #: **4**

Update: min preferred efficiency = 0.5; min preferred politeness = 0.5

How formal are the circumstances in which you are speaking? (not at all) [1...7] (very)
Enter #: **7**

Update: min preferred efficiency = 0.5; min preferred politeness = 2

How urgent is your goal? (not at all) [1...7] (very)
Enter #: **7**

Update: min preferred efficiency = 2; min preferred politeness = 2

How polite does your audience expect you to be in this conversation? (not at all) [1...7] (very)
Enter #: **5**

Update: min preferred efficiency = 2; min preferred politeness = 2.5

How polite is your audience? (not at all) [1...7] (very)
Enter #: 3

Update: min preferred efficiency = 2; min preferred politeness = 2

How efficient does your audience expect you to be in this conversation? (not at all) [1...7] (very)
Enter #: 6

Update: min preferred efficiency = 3; min preferred politeness = 2

How efficient do you think you should be in this conversation? (not at all) [1...7] (very)
Enter #: 7

Update: min preferred efficiency = 4.5; min preferred politeness = 2

How attentive are you to social cues in general? (not at all) [1...7] (very)
Enter #: 7

Update: min preferred efficiency = 4.5; min preferred politeness = 3.5

What is your social position relative to your audience? (much lower) [1...7] (much higher)
Enter #: 4

Update: min preferred efficiency = 4.5; min preferred politeness = 3.5

How concerned are you for the status of your relational bond with your audience? (not at all) [1...7] (very)
Enter #: 1

Update: min preferred efficiency = 4.5; min preferred politeness = 2

How private is the situation you are speaking about? (not at all) [1...7] (very)
Enter #: 2

Update: min preferred efficiency = 4.5; min preferred politeness = 1

How polite do you think you should be in this conversation? (not at all) [1...7] (very)
Enter #: 3

Update: min preferred efficiency = 4.5; min preferred politeness = 0.5

lpx_read_cpxlp: reading problem data from `Problems_cpxlp/Provide_Guidance-4.5e-0.5p-Stop_Annoying_Habit-4.5e-0.5p-v1.lp'...

lpx_read_cpxlp: 4 rows, 56 columns, 222 non-zeros

lpx_read_cpxlp: 56 integer columns, all of which are binary

lpx_read_cpxlp: 70 lines were read

lpx_simplex: original LP has 4 rows, 56 columns, 222 non-zeros

lpx_simplex: presolved LP has 4 rows, 56 columns, 222 non-zeros

lpx_adv_basis: size of triangular part = 4

0: objval = 0.000000000e+00 infeas = 1.000000000e+00 (0)

7: objval = 3.929487179e+00 infeas = 0.000000000e+00 (0)

* 7: objval = 3.929487179e+00 infeas = 0.000000000e+00 (0)

* 12: objval = 2.743589744e+00 infeas = 0.000000000e+00 (0)

OPTIMAL SOLUTION FOUND

Integer optimization begins...

Objective function is integral
+ 12: mip = not found yet >= -inf (1; 0)
+ 47: mip = 4.000000000e+00 >= 3.000000000e+00 25.0% (35; 0)
+ 49: mip = 3.000000000e+00 >= 3.000000000e+00 0.0% (13; 45)
+ 49: mip = 3.000000000e+00 >= tree is empty 0.0% (0; 71)
INTEGER OPTIMAL SOLUTION FOUND
Time used: 0.0 secs
Memory used: 0.1M (99776 bytes)
lp_solve: writing MIP problem solution to `Problems_cpxlp/gnuGLPKSolver.out'...

There is an optimal solution containing these 3 act(s): give ask inform
For goal Stop_Annoying_Habit,
achieved efficiency: 4.51
achieved politeness: 4.87
For goal Provide_Guidance,
achieved efficiency: 5.58
achieved politeness: 7.11

But first we have to verify that the conditions of these acts are all satisfied...

[give] has the following preconditions:

- * There is a certain event that would be good for your audience
- * You can cause the event to happen
- * You DO have to follow what your audience tells you
- * Your audience won't tell you not to

Are all of these conditions satisfied in the current situation? [ynq]: y

[ask] has the following preconditions:

- * There is an event you want to take place
- * The event involves your audience sharing info with you
- * Your audience could make it happen, but currently isn't intending to (and might be resistant)
- * You expect your audience to act as you would prefer
- * You don't know what the info will be

Are all of these conditions satisfied in the current situation? [ynq]: y

[inform] has the following preconditions:

- * There is something your audience doesn't know
- * Your audience isn't likely to discover it on their own
- * You feel obliged to mention it to your audience
- * Your audience will understand that it must be true if you say so

Are all of these conditions satisfied in the current situation? [ynq]: y

The final solution contains these 3 act(s): give ask inform

APPENDIX G

PILOT QUESTIONNAIRE

Pursuing Multiple Social Goals Simultaneously

[Exit this survey >>](#)

1. Overview

Estimated time to complete questionnaire: 15-30 mins

This questionnaire begins by setting the stage for an imaginary conversation, providing you with a pair of goals and some facts about the situation and the person you'll be speaking with (i.e., "the audience"). Then, eight sets of sentence types are presented, representing some of the ways one might pursue your side of the goal-driven conversation. We ask you to answer two questions for each set: "If I said all of the things in this set, would it be an acceptable way of achieving the **first** goal?" and "If I said all of the things in this set, would it be an acceptable way of achieving the **second** goal?"

Some sets will be acceptable for both goals, some sets might work only for one goal, and some might not work for either – that's why we ask about the two goals separately for each set. And we are asking you to judge each set as a whole -- don't ignore a sentence type if it seems not to fit with the others; just make that lack of fit part of your acceptability judgment.

The goals you should imagine pursuing are:

- **Obtain a favor** – You are going on a trip and need to find someone to feed your pet.
- **Get advice** – Your flight is during rush hour, and you'd like to know a quick way to the airport.

And here are some more facts about the imaginary you:

- You are moderately goal-oriented (5 on a 1-7 scale).
- You are moderately attentive to social cues in general (4 on a 1-7 scale).

You've picked out a person to approach for help, "the audience". Actually, your parents suggested this person, a family friend, and since your flight is just a few days away and everyone else is busy, you don't have much choice. This person is your parents' age, a bit old-fashioned, formal, and opinionated. You're a little concerned about what they might think of your place, and what they might tell your parents, but there's no one else you can ask. Besides, they'll probably be happy to tell you their secret fast route to the airport.

Imagine the following things are true about the audience and the situation in which you'll be speaking:

- The audience is somewhat polite (4 on a 1-7 scale).
- The audience expects you to be fairly polite (6 on a 1-7 scale).
- The audience is fairly efficient (6 on a 1-7 scale).
- The audience expects you to be somewhat efficient (4 on a 1-7 scale).
- Your social position is much lower than that of the audience (2 on a 1-7 scale).
- The situation is fairly formal (6 on a 1-7 scale).
- The situation you are talking about is fairly private (6 on a 1-7 scale).
- Your goals are fairly urgent (6 on a 1-7 scale).
- You think you should be moderately efficient (5 on a 1-7 scale).
- You think you should be very polite (7 on a 1-7 scale).
- You are fairly concerned about the status of your relationship with the audience (6 on a 1-7 scale).

With these conditions and goals in mind, please indicate how acceptable each of the following sets would be to you, as the speaker. Note that the sentence examples have been kept very general in order to keep the focus on the sentence type. The lack of specifics may make rating these sets feel awkward, but we ask that you try to imagine a way that each type might apply to the situation; for example, the Permit type might seem a strange way to Obtain a Favor, but permitting the audience to use your cable TV is one way to imagine how it could be relevant and acceptable. But some types still might not seem acceptable, so you should rate them accordingly. Also note that sentence types are in no particular order within the set. Feel free to reorganize them in whatever order would make the most sense. And if you aren't sure how to rate something, just check the 'Not sure' option.

Imagine that you've asked the audience person to meet with you at a cafe, and you're going to pursue your goals over conversation there.

Next >>

Pursuing Multiple Social Goals Simultaneously

[Exit this survey >>](#)

2. Sentence Set #1

Thank - I'm so grateful.

Warn - I would avoid that if I were you.

Apologize - I'm sorry to impose on you.

Forbid - Don't mess with that.

* 1. Is this set an acceptable way to Get Advice?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

* 2. Is this set an acceptable way to Obtain a Favor?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

Pursuing Multiple Social Goals Simultaneously

[Exit this survey >>](#)

3. Sentence Set #2

Assure - There's not much you'd have to do.

Explain - ...and that's why I need your help.

Confess - If I had my act together, it wouldn't be this way.

Question - Are you sure that's right?

* 3. Is this set an acceptable way to Get Advice?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

* 4. Is this set an acceptable way to Obtain a Favor?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

Pursuing Multiple Social Goals Simultaneously

[Exit this survey >>](#)

4. Sentence Set #3

Offer - Let me know if I can return the favor.

Request - Would you do this for me?

Assure - There's not much you'd have to do.

Advise - This is the best way to do it.

Apologize - I'm sorry to impose on you.

Question - Are you sure that's right?

* 5. Is this set an acceptable way to Get Advice?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

* 6. Is this set an acceptable way to Obtain a Favor?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

Pursuing Multiple Social Goals Simultaneously

[Edit this survey >>](#)

5. Sentence Set #4

Request - Would you do **this** for me?

Compliment - It's so nice of you to help me.

Thank - I'm so grateful.

Explain - ...and that's why I need your help.

* 7. Is this set an acceptable way to Get Advice?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

* 8. Is this set an acceptable way to Obtain a Favor?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

6. Sentence Set #5

Challenge - I bet you don't know a better way.

Compliment - It's so nice of you to help me.

Confess - If I had my act together, it wouldn't be this way.

Advise - This is the best way to do it.

Question - Are you sure that's right?

*** 9. Is this set an acceptable way to Get Advice?**

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

*** 10. Is this set an acceptable way to Obtain a Favor?**

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

7. Sentence Set #6

Assure - There's not much you'd have to do.

Thank - I'm so grateful.

Promise - I'll definitely return the favor.

Explain - ...and that's why I need your help.

Forbid - Don't mess with that.

Apologize - I'm sorry to impose on you.

*** 11. Is this set an acceptable way to Get Advice?**

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

*** 12. Is this set an acceptable way to Obtain a Favor?**

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

Pursuing Multiple Social Goals Simultaneously

[Exit this survey >>](#)

8. Sentence Set #7

Request - Would you do this for me?

Compliment - It's so nice of you to help me.

Advise - This is the best way to do it.

Criticize - There's a better way to do it than that.

Explain - ...and that's why I need your help.

Question - Are you sure that's right?

* 13. Is this set an acceptable way to Get Advice?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

* 14. Is this set an acceptable way to Obtain a Favor?

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

9. Sentence Set #8

Offer - Let me know if I can return the favor.

Assure - There's not much you'd have to do.

Thank - I'm so grateful.

Apologize - I'm sorry to impose on you.

* **15. Is this set an acceptable way to Get Advice?**

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

* **16. Is this set an acceptable way to Obtain a Favor?**

- ☐ 1 (not at all)
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ 6
- ☐ 7 (absolutely)
- ☐ Not sure

[<< Prev](#)

[Next >>](#)

APPENDIX H

SURVEY QUESTIONNAIRE

SURVEY OF ACCEPTABLE SOCIAL COMMUNICATION

This questionnaire begins by setting the stage for an imaginary conversation, providing you with a pair of goals and some facts about the situation and the person you'll be speaking with (i.e., "the audience"). Then, eight sets of sentence types are presented, representing some of the ways one might pursue your side of the goal-driven conversation.

We ask you to answer two questions for each set:

"If I said all of the things in this set, would it be an acceptable way of achieving the first goal?" and

"If I said all of the things in this set, would it be an acceptable way of achieving the second goal?"

Some sets will be acceptable for both goals, some sets might work only for one goal, and some might not work for either – that's why we ask about the two goals separately for each set. And we are asking you to judge each set as a whole – don't ignore a sentence type if it seems not to fit with the others; just make that lack of fit part of your acceptability judgment.

Scenario Background:

You are planning a trip and need to find someone to take care of your pet. You also need to get advice on the quickest way to the airport during rush hour in order to catch your flight.

So, the two goals you are pursuing are:

- **Goal: Obtain a favor** – You need to find someone to feed your pet.
- **Goal: Get advice** – You'd like to learn the quickest way to the airport during rush hour.

And imagine that these facts are true about you in this scenario:

- You are moderately goal-oriented (5 on a scale of 1 - 7)
- You are moderately attentive to social cues in general (4 on a scale of 1 - 7)

You've found only one person who can help you with both of these goals. Actually, your parents suggested this person, a family friend, and since your flight is just a few days away and everyone else is busy, you don't have much choice.

This person is your parents' age, a bit old-fashioned, formal, and opinionated. You're a little concerned about what they might think of your place, and what they might tell your parents, but there's no one else you can ask. In this survey, we will refer to this person as "the audience". Imagine that you've asked the audience person to meet with you at a cafe, and you're going to pursue your goals through conversation there.

Now imagine that the following things are true about “the audience” and the situation:

- The audience is somewhat polite (4 on a scale of 1 - 7)
- The audience expects you to be fairly polite (6 on a scale of 1 - 7)
- The audience is fairly efficient (6 on a scale of 1 - 7)
- The audience expects you to be somewhat efficient (4 on a scale of 1 - 7)
- Your social position is much lower than that of the audience (2 on a 1 - 7 scale)
- The situation is fairly formal (6 on a scale of 1 - 7)
- The situation you are talking about is fairly private (6 on a scale of 1 - 7)
- Your goals are fairly urgent (6 on a scale of 1 - 7)
- You think you should be moderately efficient (5 on a scale of 1 - 7)
- You think you should be very polite (7 on a scale of 1 - 7)
- You are fairly concerned about the status of your relationship with the audience (6 on a scale of 1 - 7)

With these conditions and goals in mind, please indicate whether each of the following sets of sentence types would be acceptable to you, as the speaker.

Note that the sentence examples have been kept very general in order to keep the focus on the sentence type. The lack of specifics may make rating these sets feel awkward, but we ask that you try to imagine a way that each type might apply to the situation; for example, the ‘Permit’ type might seem a strange way to ‘Obtain a Favor’, but permitting the audience to use your cable TV is one way to imagine how it could be relevant and acceptable. But some types still might not seem acceptable, so you should rate them accordingly.

Also note that the sentence types are in no particular order within the set. Feel free to reorganize them in whatever order would make the most sense. And if you aren’t sure how to rate something, just check the ‘Not sure’ option.

Estimated time to complete questionnaire: 15-30 minutes.

SETS of SENTENCE TYPES

Advise - This is the best way to do it.

Explain - ...and that's why I need your help.

Question - Are you sure that's right?

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Offer - Let me know if I can return the favor.

Compliment - It's so nice of you to help me.

Question - Are you sure that's right?

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Offer - Let me know if I can return the favor.

Assure - There's not much you'd have to do.

Apologize - I'm sorry to impose on you.

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Forbid - Don't mess with that.

Complain - It's a shame, don't you think?

Question - Are you sure that's right?

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

continued...

Assure - There's not much you'd have to do.
Confess - If I had my act together, it wouldn't be this way.
Promise - I'll definitely return the favor.
Apologize - I'm sorry to impose on you.

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Warn - I would avoid that if I were you.
Criticize - There's a better way to do it than that.
Request - Would you do this for me?

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Assure - There's not much you'd have to do.
Thank - I'm so grateful.
Explain - ...and that's why I need your help.

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Compliment - It's so nice of you to help me.
Assure - There's not much you'd have to do.
Advise - This is the best way to do it.
Apologize - I'm sorry to impose on you.

Is this set an acceptable way to Get Advice? Yes__ No__ Not sure__

Is this set an acceptable way to Obtain a Favor? Yes__ No__ Not sure__

Gender: ___ Female ___ Male Age: ___ years

Please make sure you've answered all of the questions. Thank you!

- END OF SURVEY -

REFERENCES

- Abdi, H. 2007. Binomial Distribution: Binomial and Sign Tests. In N.J. Salkind (ed.): *Encyclopedia of Measurement and Statistics*. Thousand Oaks (CA): Sage.
- Allen, J. F., L. K. Schubert, G. Ferguson, P. Heeman, C. H. Hwang, T. Kato, M. Light, N. G. Martin, B. W. Miller, M. Poesio, and D. R. Traum. 1994. *The TRAINS project: A case study in building a conversational planning agent*. Technical Report 532, Computer_ Science Department, University of Rochester, Rochester, NY.
- Austin, J. L. 1975. *How to do things with words*. Cambridge, MA: Harvard University Press.
- Bateson, G. 1958. *Naven*. Stanford, CA: Stanford University Press.
- Belz, A. 2003. *And now with feeling: Developments in emotional language generation*. Technical Report No. ITRI-03-21. Information Technology Research Institute, University of Brighton. 17 pages.
- Berger, C. R. 1997. *Planning strategic interaction: Attaining goals through communicative action*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Bickmore, T. 2003. *Relational agents: Effecting change through human-computer relationships*. PhD dissertation, MIT.
- Brown, R. and Gilman, A., 1972. The pronouns of power and solidarity, in *Language and Social Context*, (ed.) P. Giglioli. Harmondsworth: Penguin, pp. 252-282.

- Brown, P. and S.C. Levinson. 1987. *Politeness: Some Universals in Language Usage* (Studies in Interactional Sociolinguistics). Cambridge University Press.
- Bruce, B. 1975a. *Belief systems and language understanding*. Technical Report 2973. Cambridge, MA: Bolt, Beranek, and Newman.
- Bruce, B. 1975b. Generation as a social action. In *Theoretical issues in natural language processing*, (eds.) R. Schank and B. Nash-Webber. Cambridge, MA: Association for Computational Linguistics.
- Castelfranchi, C. 1992. No more cooperation, please! In search of the social structure of verbal interaction. In *Communication from an artificial intelligence perspective*, (eds.) A. Ortony, J. Slack, and O. Stock. New York: Springer Verlag.
- Cody, M.J., J.O. Greene, P.J. Marston, H.D. O'Hair, K.T. Baaske, M.J. Schneider. 1986. Situation perception and message strategy selection. *Communication Yearbook* 8. 390-420.
- Cohen, P. R. and C. R. Perrault. 1979. Elements of a plan-based theory of speech acts. *Cognitive Science* 3. 177-212.
- Daly, J. A. and J. M. Wiemann (eds.) 1994. *Strategic interpersonal communication*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- D'Andrade, R. G. and M. Wish. 1985. Speech act theory in quantitative research on interpersonal behavior. *Discourse Processes* 8. 229-259.

- Elliott, C. 1992. *The Affective Reasoner: A process model of emotions in a multi-agent system*. Ph.D. dissertation, Computer Science Department, Northwestern University. Technical Report 32, Institute for the Learning Sciences, Evanston, IL.
- Forbus, K. 2002. *Simulation and modeling: Under the hood of The SIMs™*. Lecture notes http://www.cs.northwestern.edu/~forbus/c95-gd/lectures/The_Sims_Under_the_Hood_files/v3_document.htm
- Free Software Foundation. 2006. GNU Linear Programming Kit. <http://www.gnu.org/software/glpk/glpk.html>
- Goffman, E. 1967. *Interaction ritual*. Garden City, NY: Doubleday Anchor.
- Green, N. 1994. *A computational model for generating and interpreting indirect answers*. Ph.D. dissertation, Department of Computer and Information Sciences, University of Delaware. Technical Report 95-05.
- Grice, P., 1989, *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Grosz, B. and Sidner, C. 1986. Attentions, Intentions, and the Structure of Discourse. *Computational Linguistics*, 12:175–204.
- Gullahorn, J. T. and J. E. Gullahorn. 1963. A computer model of elementary social behavior. In *Computers and thought*, (eds.) E. Feigenbaum and J. Feldman. New York: McGraw-Hill.

- Haslam, N. 1994. Mental Representation of Social Relationships: Dimensions, Laws, or Categories? *Journal of Personality and Social Psychology*, Vol. 67, No. 4, pp. 575-584.
- Hovy, E. H. 1988. *Generating natural language under pragmatic constraints*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Huang, J. 2006. *Combining knowledge compilation and search for conformant probabilistic planning*. Proceedings of International Conference on Automated Planning and Scheduling (ICAPS).
- Hughes, L. 1992. *Societal simulation: An artificial intelligence approach*. PhD dissertation, Northwestern University.
- Jonsson, I-M., B. Reaves, C. Nass, & H. Harris. 2005. *Thank you, I did not see that: In-car speech based information systems for older adults*. Proceedings of the 2005 CHI Conference, Portland, OR.
- Jonsson, I-M., C. Nass, J. Endo, B. Reaves, H. Harris, J. Le Ta, N. Chan, and S. Knapp. 2004. *Don't blame me I am only the driver: Impact of blame attribution on attitudes and attention to driving task*. Proceedings of the 2004 CHI Conference, Vienna, Austria. 1219-1222.
- Kellermann, K. 2006. *Theory of Selection* (chapter). Unpublished manuscript.
- Kellermann, K. 2004. A goal-directed approach to gaining compliance: Relating differences among goals to differences in behavior. *Communication Research* (31)4. 397-445.

- Kellermann, K. and H.S. Park. 2001. Situational urgency and conversational retreat: When politeness and efficiency matter. *Communication Research* (28) 1. 3-47.
- Kellermann, K. and C. Lee. 2001a. *Seeking Disaffinity: Making Others Dislike and Feel Negatively About You*. National Communication Association Conference.
- Kellermann, K. and C. Lee. 2001b. Coauthor's preparation notes for: *Seeking Disaffinity: Making Others Dislike and Feel Negatively About You*. National Communication Association Conference.
- Kellermann, K. and T. Cole. 1994. Classifying compliance gaining messages: Taxonomic disorder and strategic confusion. *Communication Theory*, 4, 3-60.
- Kiesler, D. 1983. The 1982 interpersonal circle: A taxonomy for complementarity in human transactions. *Psychological Review* (90) 3. 185-214.
- Lasseter, J. 1987. Principles of traditional animation applied to 3D computer animation. *Computer Graphics*, Vol. 21, Number 4.
- Lee, M.G. 1998. *Belief, rationality, and inference: A general theory of computational pragmatics*. PhD dissertation, Department of Computer Science, University of Sheffield.
- Maes, P. 1989. How to do the right thing. *Connection Science Journal*, 1(3).

- Mao, W., and J. Gratch. 2003. The social credit assignment problem. *The Fourth International Working Conference on Intelligent Virtual Agents (LNCS/LNAI 2792)* (Kloster Irsee, Germany, September, 2003).
- Meehan, J. 1976. *The Metanovel: Writing stories by computer*. PhD Dissertation, Yale University.
- Oakman J., Gifford S., Chlebowsky, N. 2003. A Multilevel Analysis of the Interpersonal Behavior of Socially Anxious People. *Journal of Personality*, Volume 71, Number 3, June 2003, pp. 397-434(38). Blackwell Publishing.
- Ortony, A., G. Clore, and A. Collins. 1988. *The cognitive structure of emotions*. New York: Cambridge University Press.
- Pautler, D. & A. Quilici. 1998. *A computational model of social perlocutions*. Presented at COLING-ACL, Montreal, Quebec.
- Pautler, D. 1997. *Strategic social communication and its basis in natural reasoning*. Unpublished manuscript.
- Pentland, A. 2005. Socially-aware computation and communication. *Computer*, (38) 3. 33-40.
- Pérez y Pérez, R. 1999. *MEXICA: A computer model of creativity in writing*. PhD Dissertation, University of Sussex.
- Picard, R.W. 2000. *Affective computing*. MIT Press.

- Reilly, W. and J. Bates. 1992. *Building emotional agents*. Technical Report 143, Computer Science Department, Carnegie Mellon University, Pittsburgh, PA.
- Rosenschein, J. and G. Zlotkin. 1994. Designing conventions for automated negotiation. *AI Magazine* (15) 3. 29-46.
- Searle, J. 1969. *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- Searle, J. and D. Vanderveken. 1985. *Foundations of illocutionary logic*. Cambridge University Press.
- Shoham, Y. and M. Tennenholtz. 1992. On-the synthesis of useful social laws for artificial agent societies. *Proceedings of the Tenth National Conference on Artificial Intelligence*. Menlo Park, CA: American Association for Artificial Intelligence. 276-281.
- Svennevig, J. 2000. *Getting acquainted in conversation: A study of initial interactions*. John Benjamins Publishing Co.
- Wierzbicka, A. 1987. *English speech act verbs*. Academic Press.