SANDIA REPORT

SAND2012-8113
Unlimited Release
Printed September, 2012

Diffusion Among Cognitively Complex Agents Final Report

Kiran Lakkaraju, Ann Speed, Craig M. Vineyard, Joseph Collard, and Stephen J. Verzi

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.

Sandia National Laboratories

Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: http://www.osti.gov/bridge

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online ordering: http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online

SAND2012-8113
Unlimited Release
Printed September, 2012

Diffusion Among Cognitively Complex Agents Final Report

Kiran S. Lakkaraju
Cognitive Systems
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-9999
klakkar@sandia.gov

Ann Speed
Cognitive Systems
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-9999
aespeed@sandia.gov

Craig M. Vineyard
Cognitive Systems
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-9999
cmviney@sandia.gov

Joseph Collard
Systems Research, Analysis & Applications
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185
jmcolla@sandia.gov

Stephen J. Verzi
Systems Research, Analysis, & Applications
Sandia National Labs
P.O. Box 5800
Albuquerque, NM 87185-9999
sjverzi@sandia.gov

Abstract

This is the final report for the Early Career LDRD "Diffusion Among Cognitively Complex Agents" (DACCA). This report consists of three main parts. In the first part we overview the motivation and summarize our accomplishments. The second part provides a detailed description of the model and simulation results. The third part describes an Amazon Mechanical Turk based experiment design to gather data for calibrating and validating this model.

Contents

1 Overview **9**

InformationDiffusion 9

Summary of Accomplishments 10

Anonymization and Data Sharing 11

Conclusions and Future Work 11

2 The DACCA Model and Simulations **13**

Introduction 13

Attitudes & Attitude Change 14

Attitude change and cognitive consistency 16

Summary of Factors that Influence Attitude Change 18

Consistency and Parallel Constraint Satisfaction 19

An Example of Constraint Satisfaction 21

A Socio-Cognitive Constraint Satisfaction Model of Attitude Change 22

Experiments 24

Cognitive network 24

Experimental Setup and Results 25

Discussion 29

Related Work 32

Conclusions & Future Work 33

3 Attitude resolve and population wide attitude change **35**

Introduction 35

Theoretical basis 36
Model Description 36
Experimental Results 38
Parameterizations: 38
Discussion 39
Experiment 1: Number of Advocates 39
Experiment 2: Position of Advocates 39
Experiment 1 39
Experiment 2 40
Conclusion & Future Work 41
Conclusion 41
Future Work 42

4 Online Data for Validation **43**
Amazon Mechanical Turk 43
Overview of Design 44
Experiment Design 45
Procedure 46

References **47**

List of Figures

2.1 Example of the updating procedure in a PCS network. 22

2.2 Illustration of the dynamics of one timestep in the simulation. 24

2.3 Cognitive network of the agents in the simulation. 25

2.4 Trajectory from the positive and negative states over 100 state updates. 26

2.5 Location in F1/X space after 99000 iterations on a complete graph. 27

2.6 Location in F1/F2 space after 99000 iterations on a small world graph. 27

2.7 Location in F1/F2 space after 99000 iterations on a regular graph. 28

2.8 Location in F1/F2 space after 99000 iterations on a lattice graph. 28

2.9 Agents in $F1/F2$ space for a run that did not result in agreement at iteration 99000. 29

2.10 Location in F1/F2 space after 99000 iterations on a complete graph. The trajectories from the positive and negative start points are labelled. 30

2.11 f_p vs. Number of iterations till agreement. 31

2.12 f_p vs. Number of runs that reached agreement. 31

3.1 Small ring and hierarchical network topologies 40

3.2 Drastically different results for different advocate positions within a network . 41

List of Tables

2.1 Attitude Change Phenomena 18

2.2 Important Factors 19

Chapter 1

Overview

InformationDiffusion

Information diffusion is the process by which information spreads throughout a population. Models of diffusion in societies have been developed for a variety of domains, such as the spread of innovations [46, 21], the diffusion of opinions [6], and the spread (and change) of language [64]. Models that exhibit qualitatively similar behavior have also been developed for the study of disease spreading [42] and even the diffusion of sensor readings in a sensor network [39]. The main thrust of these models has been on understanding how the social network impacts the diffusion process while leaving the internal cognitive models of agents extremely simple. These simple cognitive models do not take into account, for instance, the interdependence of beliefs and the tendency against changing ones beliefs. These factors can be important an individual who has a strong tendency against changing their beliefs might never do so, no matter how many friends have a different belief; in contrast, an individual with a strong tendency to change will do so with little prompting. Through understanding both social and cognitive factors, and how they interact, one can gain a better understanding of the diffusion process and consequently develop better techniques for enabling rapid diffusion of information.

Theories of cognition that center on the idea of the interdependence of beliefs are called Cognitive Consistency Theories (CCT) [51]. There are two hypotheses to CCTs: (1) individuals have a set of concepts that they accept (believe are true) or reject (believe are false); some of these concepts fit well together and others do not; and (2) individuals endeavor to accept a set of concepts that fit well together and reject the concepts that do not fit well together. Recent research [51] has indicated that CCTs can help shed light on human reasoning processes. For the current purposes, CCTs are particular interesting as they are easily placed into a computational framework - specifically using a constraint satisfaction approach.

The objective of this Early Career LDRD work is to study the diffusion of information in complex settings; i.e. ones that exhibit socially complex agents (differing numbers of neighbors, captured through a social network); cognitively complex agents (continuous, interdependent belief structures that change over time, rather than binary, independent belief structure, captured through a cognitive network); and limited resources (bounded rationality,

modeled by limiting the resources for communicating and integrating information).

Developing a model of diffusion can increase our ability to predict and affect the flow of information in society. This ability has great potential to help missions in many areas including the climate change mission of Sandia and the DOE, as well as missions of the DoD and DHS. Regarding climate change, understanding how individual beliefs about the validity of the science and how that interacts with messages received from their social network can provide an understanding of how readily anti-climate change messages will be adopted. From a DoD point of view, understanding the spread of tactics, techniques, and procedures (TTPs; e.g., IED tactics) can allow security forces to better estimate the capability evolution of the enemy especially between theaters of operation. This is important for DOE as well in its mission to provide physical security to DOE complexes. From a DHS perspective, in which community resilience (the ability of a community to recover from damage) is a critical area of work, understanding diffusion processes can help federal agencies develop communication strategies that more quickly and effectively spread critical safety information.

Summary of Accomplishments

This Early Career LDRD was funded for 2 years with a 6 month no-cost extension. The funding allowed the principal investigator, Kiran Lakkaraju, to allocate approximatly 60% of his time to the project.

The predominant work was in developing, implementing, and running simulations on a new diffusion model. The model and simulations are described in Chapter 2 and 3.

A second part of the project was in addressing issues related to data generation and gathering. This resulted in significant interaction with the DOE Human Subject Board. A description of this work is provided in Section 1.

The third part of the project was designing an experiment in Amazon Mechanical Turk (AMT) that can be used to calibrate/validate our model. The design of this experiment is described in Chapter 4.

The following is a summary of the accomplishments of this project:

Development and implementation of new diffusion model A new model of complex diffusion (the DACCA model) was developed and implemented.

Publications 6 publications; 4 papers at peer reviewed symposiums and conference, 1 extended abstract at a peer reviewed symposium, and 1 invited book chapter in a collection. See [27],[28], [60], [31], [30], [29].

Presentations The PI presented this work at several (> 6) conferences.

Anonymization and Data Sharing Provided expertise to DOE HQ Institutional Review Board (IRB) on de-identification for data sharing.

Transition The developed model was integrated into the Cultural Geography tool [1] as part of a project led by Zach Heath (8958) (this activity was separately funded). The model was also implemented in the CASoS platform.

Anonymization and Data Sharing

Calibration and validation of social system models requires data from human subjects with which to compare the model to. Using and getting this data is under the jurisdiction of an "Institutional Review Board" (IRB) which oversees human subject experimentation at a federal level. Organizations have their own local IRB boards who evaluate human subjects experiments for potential harm to the subjects. At Sandia, the IRB is called the "Human Subjects Board" (HSB).

On December 9th, 2009, a Secretarial Policy Memo was issued by Secretary Steven Chu on military or intelligence-related human subject research (referred to as human terrain mapping). This memo indicated that work by DOE program offices and field elements (including the National Nuclear Security Administration/NNSA) are to limit work in this area to only those projects "involving analysis and management of fully *de-identified data*" (emphasis added).

The memo directly impacts many projects at Sandia and potential collaborations. The DACCA project would also have been impacted by severely restricting data sources that could be used in the project.

Wendy Shaneyfelt and the PI interacted with the DOE HQ IRB to provide expertise on data de-identification/anonymization and validation. We lead and organized two workshops (funded by DOE HQ) to clarify de-identification issues. This interaction helped develop clarity on the difference between Human Terrain Mapping projects and non-Human Terrain Mapping projects; and the feasibility of de-identification.

Conclusions and Future Work

The sudden and surprising emergence of numerous protests in the Middle East underscores the need for understanding and anticipating behavior. Computational modeling of populations and influence operations has recently emerged as a means of gaining insight into these phenomena. Computational models allow us to rapidly test different hypotheses in order to understand how different elements can influence population behavior.

Developing appropriate models is often a "balancing act" between simplicity and incor-

porating all the critical aspects. On one hand, an "emulation" model [4] often contains many parameters and much detail. While these can be tailored to fit many situations, their sheer complexity can make them hard to understand and test. On the other hand, "intellective" models [4] are simplistic in nature and capture only the most necessary aspects of the situation. These models are relatively easy to test and draw conclusions from (for instance, see [6]).

The model developed in this project sits between these two extremes, but closer to the "intellective" side. The addition of an internal network of attitudes increases the complexity of our model, but it also allows us to capture the interplay between social influence and internal consistency, which is known to be an important aspect of attitude change. While we have increased the complexity, we believe that this model is amenable to analysis and testing. Our future work on this model will focus on gathering data that can be used to calibrate and validate the model. Chapter 4 describes one approach. Data from Massively Multiplayer Online Games may also be of use for calibration.

During the project execution we transitioned the model to the Cultural Geography toolkit [1]. In the future, we would like to identify other social simulation frameworks that could benefit from our model. The Behavioral Influence Assessment (BIA) model is one such internal framework. We are currently talking with the PI of this project to identify potential collaboration opportunities.

The rest of this report contains a thorough description of the model and simulation results. These were taken from two of the publications, [30] and [60]. We will end with a discussion of an experiment design for gathering data via Amazon Mechanical Turk.

Chapter 2

The DACCA Model and Simulations

This chapter contains the material from [30] which was authored jointly by Kiran Lakkaraju and Ann Speed. Here we describe the model and the underlying social psychology theories that informed it. We end with simulations of attitude diffusion among populations and importance of network topology.

Introduction

Attitudes are "general and relatively enduring evaluative responses to objects" where objects can be "a person, a group, an issue or a concept" [61, Page 1]. Attitudes are shown to have an impact on, and can sometimes predict, the behaviors of individuals (e.g., voting behavior [22], consumer purchases [20, 14]).

Understanding population wide attitude change is thus an important step to understanding the behavior of societies. For instance, consider the change in attitudes towards global warming and the environment that has resulted in a significant change in public policy and national priorities [7].

While there are a number of factors that influence attitude change [62], we focus on two. The first is social – individuals are exposed to various attitudes and information through interaction with others. Family, friends, acquaintances, and the media all influence the attitudes of individuals by providing new information/opinions. [63] demonstrated that the social context of an individual can impact a large number of aspects of that person.

The second factor is cognitive – individuals tend to hold a set of attitudes that are consistent with each other [47, 51, 54]. According to these *cognitive consistency theories*, an individual holding a strong positive attitude towards environmentalism should also hold a strong positive attitude towards recycling; if they do not, the attitudes are inconsistent with each other and could cause an uncomfortable feeling (i.e. *cognitive dissonance*) which tends to result in either attitude or behavior change [62].

The perspective of the authors is as follows, we argue that the phenomena of attitude change can be reasonably captured through a multi-scale *complex adaptive system* beginning with two levels: (1) a social level where autonomous agents representing individuals interact

and exchange information; and (2) a cognitive level, where individual concepts interact in order to preserve consistency.

These two factors are tightly linked. The social context of an individual can influence the information they are exposed to as well as how amenable to change the individuals are [63]; while the individual's existing attitudes may distort or diminish the impact of information [54, 16] and could impact the composition of an individuals social network.

Our research aim is to develop a model that can capture these two levels and the interaction between them. This paper presents preliminary results from the development of a new computational model of attitude change that incorporates both social and cognitive factors. The model represents individuals as networks of concepts (e.g., beliefs) that interact with each other. Social structure is implemented through a social network that limits the interaction between agents.

The dynamics of this model are explored through a simple attitude change experiment. The objective is to understand the *resiliency* of attitudes – how long does it take for attitudes to change and what do they change to? Two populations of agents with different attitude structures interact. We imagine this as a new population of individuals being introduced into an existing population. We show that the population does reach agreement (all agents have similar attitudes), and that the state that is agreed upon is a function of the initial distribution of attitudes in the population. In addition, we track time to agreement and show that it is a function of the social network and the initial distribution of attitudes in the population.

We first describe some of the empirical findings on attitude change and consistency and summarize some of the key phenomena and factors that can influence attitude change. We then describe the parallel constraint satisfaction model which we will use to implement cognitive networks. Finally, we describe our model and illustrate some properties through experiments on a simple attitude change scenario.

Attitudes & Attitude Change

Research on attitudes and attitude formation indicates that an attitude is generally considered an opinion or a feeling about something that has both a valence (i.e., positive or negative) and a strength [66]. These opinions influence behaviors and influence how new information is received and processed [13, 10, 48, 54, 66].

Even though research has traditionally focused on attitudes as singular feelings (e.g., one either has a positive or a negative attitude about American football), research in the last decade has highlighted the fact that an individual can hold two attitudes about the same object, and that some of these attitudes are explicit and some are implicit [48, 66]. Further, research has shown that such dual attitudes can even conflict with one another. For example, by explicit measures (e.g., questionnaires), a Spaniard might express his love of American

football, but when measured using implicit measures, he may reveal some disdain for the sport, and this may influence nonverbal behaviors or motivations [10]. These are in stark contrast to ambivalence (e.g., positive feelings for the flavor of a food but negative feelings about the calorie content of that food).

As an example, in five experiments, Rydell and McConnell 2006 [48] demonstrated the ability to create separate implicit and explicit attitudes about a fictional character, Bob, and that different manipulations impact the explicit and implicit attitudes differently. What is particularly interesting about this series of experiments is that the authors do not rely on previously held attitudes they create attitudes in the laboratory. However, studies of attitudes using pre-existing beliefs tend to yield similar results [13, 15, 16].

These findings are important because of the implications for behavior. Research has demonstrated that explicit, controllable behaviors change as explicit attitudes change. Research has also demonstrated that implicit attitudes change more slowly than do explicit attitudes, and that implicit attitudes predict a separate set of behaviors than do explicit attitudes [10, 48, 13, 14, 38]. Specifically, explicit attitudes predict behaviors over which the individual holding the attitude has control. Implicit attitudes tend to predict behaviors that are more subtle (e.g., nonverbal behaviors) and other behaviors over which the individual holding the attitude has little or no control.

If the findings in [48] about the factors that cause attitude change hold (e.g., the effect on attitude change of the relative proportion of negatively versus positively valenced evidence), that has implications for the kinds of things that need to be considered in efforts to change attitudes including specific actions to be taken as well as the amount of information needed to enact change in both explicit and implicit attitudes. Otherwise, the risk is that behaviors that are under conscious control may come to reflect the explicit attitude while behaviors that are more subtle, but possibly as powerful, will reflect the contrary implicit attitude.

Another variable that seems to impact explicit attitude change is the effect of cognitive load. Thus, current research suggests two routes to attitude change. In the systematic route individuals are assumed to expend deliberate effort to evaluate a message; including evaluation of the source of the message, the validity of the arguments, and relating the elements to their own attitudes. However, in situations of high cognitive load, that is periods where individuals are distracted and cannot expend cognitive effort, the peripheral route is taken. In these instances individuals do not evaluate the argument but rather use heuristics to decide whether to change their attitude. These heuristics can depend upon the message structure (if the message has many arguments individuals are more likely to be persuaded); or based on the context of the message (if many people already agree to the message the receiver is more likely to agree). Current work shows that this dual process model accounts for instances of attitude change that the systematic model does not [62].

Attitude change and cognitive consistency

One important question to ask is - when might an attitude change? The complimentary question is - when might the thinker distort new information in order to fit with an existing attitude or opinion? Several such factors have been mentioned above, but it is worthwhile to discuss those in greater detail. Generally, attitudes are less likely to change in an enduring fashion when (1) they are strongly held, (2) the thinker has less time or fewer cognitive resources to think through the logic of an argument that would result in attitude change, and (3) if they are implicitly, rather than explicitly held beliefs [16, 48, 66]. In terms of explicit attitudes, when the attitude isn't changed by incoming information, the thinker often distorts that incoming information to fit with his or her current attitude. Overall, this process of either changing ones attitude or biasing incoming information seems to be mediated by a drive for consistency between beliefs, attitudes and action [44, 47, 54, 32].

According to these cognitive consistency theories, individuals strive to hold a set of attitudes that are consistent with each other. For example, an individual holding a strong positive attitude towards environmentalism should also actively engage in recycling or composting household waste; if they do not, the attitudes are inconsistent with each other and could cause an uncomfortable feeling (i.e. cognitive dissonance) which tends to result in either attitude or behavior change [62]. The important idea is that attitudes interact with each other; a change in one attitude can have an effect on another.

Some of the classic studies on consistency demonstrate that people have a tendency to rate items differently when they occur in isolation than they do when they occur within an attitude-formation context.

In a subsequent series of studies, [54] demonstrated that attitudes about situations tend to form early in the evidence accrual process and that these forming attitudes tend to bias subsequent information, strengthening the original attitude. These authors also demonstrated that if the subject's attitudes later changed, they then biased information in the opposite direction. That is, if Information A is contradictory to a subject's initial attitude, he will discount it. However, if that same subject changes his attitude in the end, the discount on Information A will no longer be apparent. Rather, he will have discounted Information B, which was supportive of his initial attitude.

Specifically, the authors presented two sets of questions to subjects, one that involved a series of unrelated vignettes, and the other that consisted of questions all relating to a single, more extended story line. Importantly, the two sets of questions were otherwise identical. Thus, the authors measured the change in attitudes that was attributable to the subjects forming a coherent opinion in the extended story relative to the answers they gave in the unrelated vignettes. Over four studies, they found that people discount information that is contrary to their forming opinion, even before they have to officially declare that opinion. They also demonstrated that when people changed their attitude (i.e., they leaned one way initially and changed their attitude to the opposite opinion in the end), their evaluation of the evidence also changed to maintain coherence with their eventual decision. Thus, pieces

of information that were weighted highly initially were down-weighted in the end if they did not support the opposite view.

Additional evidence for the drive for consistency is demonstrated in [47]. These authors utilized a number of explicit and implicit measures to demonstrate that a tendency toward consistency is, indeed, behind information distortion, rather than other processing goals such as reduction in cognitive effort or as a method for increasing the distinctiveness of the alternatives under consideration.

Interestingly, one social factor that seems to influence the malleability of a given individual's attitudes is the heterogeneity of their social network with regards to that attitude. In a series of four experiments 2 in the laboratory and 2 using existing, real-world social networks [33] demonstrated that the more heterogeneous a given individual's network is with regard to an issue (e.g., attitudes about capital punishment), the more ambivalent the person is in their attitude, and the less certain they are of their attitude. Similarly, the more homogeneous an individual's social network is regarding a particular issue (i.e., the more attitudinally congruent an individual's network), the stronger that individual's attitudes are, and the less open that individual is to attitude change. Interestingly, these researchers also found that when subjects in the laboratory were exposed to attitudinally congruent versus heterogeneous sets of opinions about the same issues, only not in the context of a social network, subjects in the attitudinally congruent condition demonstrated increased attitude change relative to those in the attitudinally heterogeneous condition when exposed to a persuasive article about the topic of focus. Although this effect was not statistically significant, the reversal of trends is curious and warrants further investigation.

In a series of studies on information distortion, [47] manipulated both implicit goal priming and explicit reporting of goals used. They demonstrated that the only process goal that drove the evaluation of evidence in progressive choice tasks was consistency. The other goals tested in these experiments included the goal of reducing the effort necessary for evaluating incoming information and the goal of increasing the separation between alternatives.

In terms of attitude change, possibly the most well-known demonstration of attitude change in order to achieve consistency is the extensive literature on cognitive dissonance. In many experiments, Festinger and colleagues demonstrated the ability to get participants to express opinions that corresponded to their behavior, even when it was clear that their original attitude was quite the opposite [62]. One example of a typical dissonance study has participants engage in an extremely boring experimental task. Then subjects are asked to tell incoming participants about the task. When subjects are asked how they personally felt about the study, those who told future participants it was fun and engaging were more likely to rate it as more enjoyable than those who did not. Interestingly, if subjects can attribute their incongruent behavior to some external factor, such as receiving a large sum of money for their fib, or being instructed by the experimenter to fib, they are less likely to report that they personally felt the task was fun.

Another variant on the consistency theme was demonstrated by [48]. These researchers demonstrated that explicit attitudes can change fairly rapidly in the face of counterattitu-

Table 2.1. A summary of important attitude change phenomena

Phenomena	Description and its Effect
The drive for consistency	Maintaining a consistent set of attitudes can drive attitude change as well as information distortion.
Information distortion	One way in which to maintain consistency between attitudes - instead of changing an attitude one can decrease the importance or impact of incoming information.
Persuasion route	Peripheral and systematic routes determine which aspects play a role in attitude structure; systematic routes focus on logical arguments whereas peripheral routes utilize heuristics. The persuasion route also influences the durability of attitude change.
Implicit/explicit attitudes	These two types of attitudes influence different behavior. Implicit attitudes affect spontaneous behaviors while explicit attitudes affect deliberate behaviors. Their change profile is also different, with explicit attitudes tending to be more malleable than implicit attitudes.

dinal information. They also demonstrated that the amount of attitude change interacted with the homogeneity of the information that was considered in the development of the initial attitude. Specifically, in one of a series of 5 experiments concerning a fictional character, Bob, subjects learned about Bob's character through reading and responding to 100 facts. For some subjects, all 100 facts about Bob were consistent (he's a nice guy.) For the other subjects, 75% were consistent and 25% were counterattitudinal. When subjects then received additional information about Bob, those who were in the 100% condition, who received counterattitudinal information, displayed a more dramatic shift in explicit attitudes about Bob than did those in the 75% condition.

Summary of Factors that Influence Attitude Change

Table 2.1 summarizes four phenomena that we argue influence attitude change. Table 2.2 summarizes factors that parameterize the four phenomena. We believe that future models of population-wide attitude change should implement the four processes parameterized by the factors. In the next section we will develop one such conceptual model. Our research aim is to implement the model and manipulate the factors to understand attitude change.

While the preceding review begins to demonstrate a fairly complex picture of attitude

Table 2.2. Important factors that influence the phenomena in Table 2.1

Factor	Effect
Cognitive load	Impacts persuasion route.
Attitude relationships	Related attitudes influence each other through the drive for consistency.
Social network	The structure of relationships between individuals influences the rate of information diffusion; in addition the distribution of attitudes of an individuals acquaintances can impact the strength of an attitude.

change and information distortion, and while there are even more variables that seem to influence these effects [62, 66], for the current model, we chose to focus specifically on factors that might influence the change of explicit attitudes. As such, we have chosen some very particular variables to manipulate in the simulation, rather than trying to capture all of the nuances presented in the vast attitude change literature (see e.g. [62]). Two aspects are of concern to us. The first is the idea of *bidirectionality* in inference. As seen in [54] the notion that individuals gather information, then come to a conclusion is not always true. Rather, it seems that individuals are constantly integrating information into their growing conclusion, and that preliminary conclusions impact how one integrates and perceives information. Bidirectionality is a special case of the idea of *wholistic processing*, that all concepts influence and are influenced by each other [45].

The second important point is that there is a rich set of behavior based on the strength of an attitude. As [48] shows, it is important to capture how the strength and directionality of an attitude as it can impact how that attitude changes. Using a binary approximation (either one believes in X or does not believe in X) doesn't capture this rich set of behavior.

Consistency and Parallel Constraint Satisfaction

The drive for cognitive consistency has been discussed since the 1950's in a variety of contexts. Theories such as cognitive dissonance and balance theory utilize a notion of consistency. It is only recently through the use of *Parallel Constraint Satisfaction* (PCS) models, however, that computational models have been developed that allow a more complete exploration of these processes.

A PCS model is a type of connectionist, attractor neural network. In this model there are a set of variables that can take on a value from a specified range. The variables have weighted links to other variables, forming a *constraint graph*. Each variable is influenced by and influences its neighbors in the constraint graph.

Depending upon the domain that the PCS model is applied to the variables, values of the variables, and links represent different things. When studying impression formation, the variables represent stereotypes and traits, values represent belief, and links represent associations between stereotypes and traits [24]. Since our objective is the study of attitude change, we view variables as representing concepts, hypotheses or information, such as the concept of "environmentalism", the hypothesis of anthropogenic global warming, or information such as data on temperature fluctuations. Henceforth, we refer to variables as "concepts". Loosely following attitude change literature, we describe the value of a concept by its *valence* – indicating the sign and its *strength* indicating its magnitude. An attitude towards a concept is then represented by the valence and strength of the concept. Positive valences indicate support of the concept and negative valences indicate non-support of the concept. To simplify discussion, we refer to the *value* of a concept as the combination of valence and strength. In our model values are continuous and range from -1 to $+1$.

Weighted links between concepts indicate the strength of influence between the concepts. A positively weighted link indicates two concepts that positively influence each other – as one of the concepts increases in value, the other concept increases as well. For example, we would posit the concept of anthropogenic global warming to have a positive link to the concept of recycling. The more one believes humans caused global warming, the more one should believe in reducing human impact on the environment.

Negatively valued links indicate concepts that do not agree with each other; as one of the concepts increases in value, the other will decrease. For example, the more one believes humans have caused global warming the less one should believe that natural climate cycles are the cause of global warming.

Links between concepts can be viewed as constraints between the values of the different concepts. Two concepts with the same valence but a negative link between them would be considered inconsistent with each other. Two concepts with differing valences but a positive link between them would also be considered inconsistent. We can measure the consistency of a set of concepts by calculating what is called the *energy* of the system [19, 45]:

$$E = -\sum_i \sum_j w_{ij} a_i(t) a_j(t) \tag{2.1}$$

where w_{ij} is the weight of the link between concepts i and j; $a_x(t)$ is the valence of concept x at time t. The energy of the system decreases if the concepts satisfy the constraints between them and increases if the concepts do not satisfy the constraints. Because of this property we can view the energy of the system as a measure of the consistency of the set of concepts, the lower the energy the higher the consistency.

PCS models utilize a connectionist approach to find a consistent set of concepts: the value of each concept is updated according to a non-linear *activation function*. One activation function that is often used is [24, 56]:

$$a_j(t+1) = a_j(t)(1-d) + \begin{cases} net_j(\max - a_j(t)) & \text{if } net_j > 0 \\ net_j(a_j(t) - \min) & \text{if } net_j \leq 0 \end{cases} \tag{2.2}$$

where:

$$net_j = \sum_i w_{ij} a_i(t) \tag{2.3}$$

and d is a decay term that is set to 0.05.

The update rule modifies the valence of a concept based on the valences of other concepts. net_j captures the input to a concept j by summing the valences of all nodes against their weight to j. If a concept does not influence another concept the weight is set to 0. max and min are the maximum and minimum values that a concept can take.

Each agent synchronously updates all concepts in the network and repeats this process until no significant change[1] has occurred between iterations. We call the repeated application of the update rule until no significant change has occurred the *update process*.

Previous work on asynchronous update protocols have shown that an update rule like Equation 2.3 will result in a state that satisfies the constraints [19]. The result of the synchronous update rule has not been proven in the same manner, however empirical results indicate that this update rule does reach a state that satisfies the constraints.

Through the use of synchronous updating and establishing constraints between concepts, PCS models capture the important properties listed above. Value changes in concepts cascade to other concepts as a function of link strength which is one aspect of wholistic processing. The use of a continuous range for concept values allows the modeling of attitude strength.

An Example of Constraint Satisfaction

The left hand side of Figure 2.1 shows an inconsistent PCS network with 5 concepts, represented as circles. The value is indicated within the node and can range from -1 to $+1$. Link weights are shown adjacent to the link and also range between -1 and $+1$.

Two pairs of concepts are inconsistent with each other, Concepts 3 and 4 have a positive link but have opposite valences; and Concepts 3 and 5 have a negative link but have the same valence. The energy of the system, calculated according to Equation 2.1 is 0.005. After executing the update process we have the network on the right, which has an energy

[1] Defined as a maximum of .01 difference between concept valences between iterations

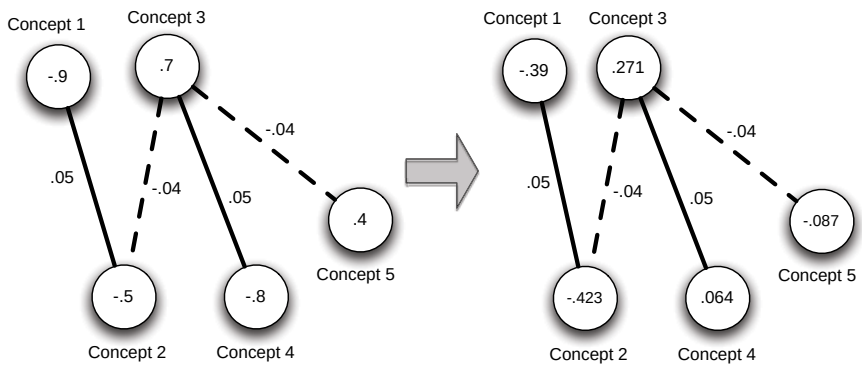

Figure 2.1. Example of the updating procedure in a PCS network.

of 0.029. In the updated network, the inconsistencies have been resolved, concept 4 now has a positive valence and concept 5 now has a negative valence.

Note that the decay parameter can have an effect on the final valences of the concepts – a decay value closer to 0.0 will result in less change of the concepts.

A Socio-Cognitive Constraint Satisfaction Model of Attitude Change

We propose a model of attitude change that brings together both social and cognitive factors that influence such change. Our model is composed of three elements:

$$< A, G, C > . \tag{2.4}$$

A represents a set of n agents that represent individuals. C is a *cognitive network* that represents the internal cognition that drives change in an attitude. We use the PCS model defined above for the cognitive networks of the agents. Each agent has the same number of concepts (m), and the same weights between links (W), however each agent can differ in the strengths and valences of those concepts.

The social network, G, is a graph of size n that defines the social structure of the population. Agents are restricted to interacting only with their neighbors.

We assume turn based dynamics – at each timestep the following actions are taken:

1. A single agent is uniformly randomly chosen from the population. We call this agent the *speaker*. One randomly chosen neighbor of the speaker is designated the *hearer*.

2. A single concept is chosen as the *topic* of communication.

3. The speaker communicates the valence and strength of the topic in its cognitive network to the hearer. This is the *communication* step.

4. The hearer changes the value of the topic (capturing the idea of integrating a speakers beliefs). This is the *information integration* step.

5. The hearer modifies the strengths and valences in its cognitive network in order to find a more consistent set of beliefs. This is the *hearer update* step.

Figure 2.2 illustrates one turn, progressing from left to right. First a speaker and a hearer are chosen from the social network. Then one of the concepts is chosen as a topic (denoted by the label "topic" in the example). All agents have the same set of concepts so it is guaranteed that the hearer and speaker will have the topic. However, they might differ in value; we can see that the speaker strongly disbelieves the concept (value is -1) while the hearer moderately supports the concept (value is .5).

The communication step represents interaction and information exchange between agents. This is, of course, a highly simplified version of what occurs in real human dynamics. For this model, we assume that agents honestly communicate their values with each other.

Information integration is the process by which an agent changes its beliefs in response to the received message. Motivated by the work of [48] our information integration process depends upon the strengths and valences of the speaker and hearer. The larger the difference between the valences, the larger the change in the hearer's valence. Specifically, we take the average of the two valences in the instantiation of the model. After the information integration step, the hearer's valence has changed to $-.25$ as it is the average of -1 and $+0.5$.

Finally, in the last step the hearer updates the valences of the concepts by running the update process. In the example, we can see that the information integration step has resulted in an inconsistency between the topic and another concept. The update process resolves this consistency, however the values of the other concepts have been altered as well. Even though consistency has been achieved, and the hearer has, essentially, the same attitude as before, there has been some change in response to the communication. It is this ability to capture the strength of an attitude, not just whether an agent has an attitude, that sets this model apart.

Note that there are two timescales here. The first is an inter-agent timescale where the protocol defined above is followed. Within each "tick" of this timescale is an updating of the cognitive network of the hearer according to Equation (2.3) which is the second timescale. The timescale will be apparent from the context.

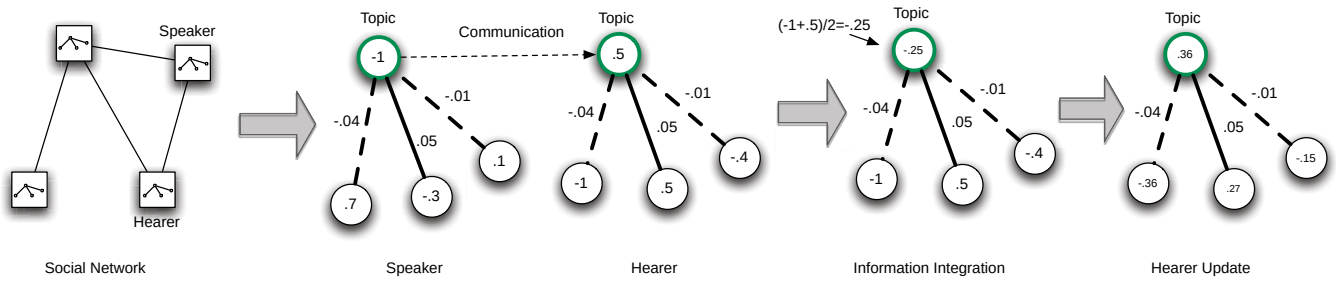

Figure 2.2. Illustration of the dynamics of one timestep in the simulation.

Experiments

To explore attitude resiliency and agreement we have implemented the model discussed above using a simple, manually designed cognitive network, and have executed numerical simulations while varying the distribution of attitudes among the agents. We would like to see whether the population can agree on a particular attitude and how long it takes to do so.

Cognitive network

Since our objective is to understand the interaction of populations with different attitudes we designed a simple cognitive network that can represent two conflicting attitudes, shown in Figure 2.3.

The concept of interest is the central node labeled X. The two nodes $F1$ and $F2$ represent information or factors that can influence the value of X. $F1$ has a positive connection to X so an increase in $F1$ increases X; $F2$ negatively influences X and thus an increase in $F2$ decreases X . This cognitive network represents a situation where individuals are receiving conflicting information; for instance evidence supporting and rejecting the theory of Anthropogenic Global Warming[2]. The weights are 0.05 and 0.05 for the positive and negative links respectively.

$F1$ and $F2$ do not represent specific pieces of information but rather, in general, all positive and negative information. So a particular piece of information (for instance, "the poloar ice caps are melting") would increase the value of $F1$ or $F2$. This abstraction allows us to focus our attention on the impact of information on X rather than on how to interpret information with regards to a concept.

[2]See: http://climatedebatedaily.com/ for a large set of articles both for and against anthropogenic global warming

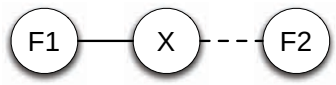

Figure 2.3. Cognitive network of the agents in the simulation.

Although simple, this network does exhibit several important characteristics that we want to capture. First, all concepts have continuous valences thus the ability to represent the strength of belief. Second, the network exhibits bidirectionality – the valence of concept X is influenced by the evidence from $F1$ and $F2$; but also belief in the evidence is influenced by belief in X.

There are two organizations of the nodes that satisfy all the constraints. In the first, which we call the *positive state*, $F1$ and concept X have positive valences while $F2$ has a negative valence. The second is called the *negative state* where $F1$ and attitude X are both negative, and $F2$ is positive. The two states reflect polar opposites – in the positive state individuals believe the positive evidence, reject the negative evidence and believe in concept X. The negative state is the opposite situation where individuals accept the negative evidence and thus reject the positive evidence and reject the concept X. For the positive state we used the values $F1 = 1, X = 1, F2 = \quad 1$ and for the negative state we used the values $F1 = \quad 1, X = \quad 1, F2 = 1$.

Experimental Setup and Results

The experimental setup was as follows:

- A population of n agents was initialized with f_p fraction of the population assigned the positive state and $f_n = 1 \quad f_p$ assigned the negative states.
- The protocol described above was run for $t = 100000$ number of steps.

Note that due to the decay parameter, the positive and negative states are not attractors under repeated application of the update process. Figure 2.4 shows the trajectory from the positive/negative states $F1/F2$. Each point represents the valences of $F1/F2$ after an application of the update process. Initially, there is a large jump from the positive/negative state. However, we see a gradual reduction in the amount of change. The two endpoints of the trajectory are designated the positive and negative endpoints and are indicated in the figures below.

To simulate the exchange of information that can impact an attitude we restricted the topics to be only $F1$ and $F2$. X was only modified through the hearer update process. The simulations here are with $n = 484$ agents.

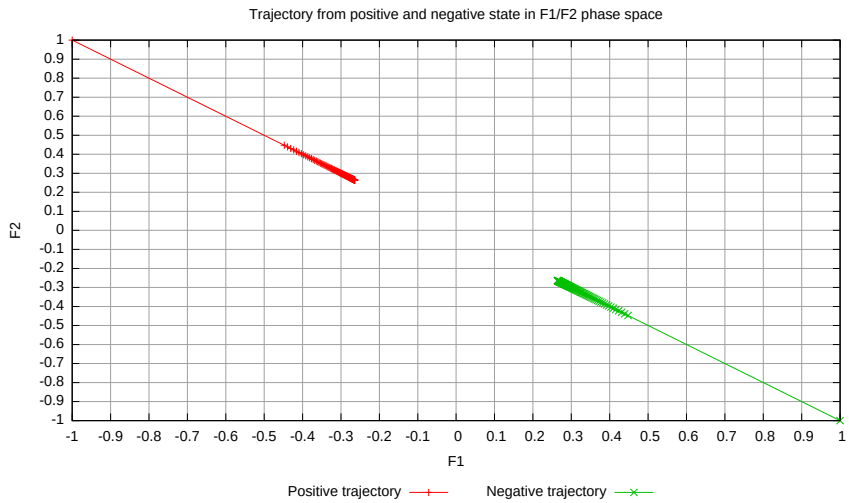

Figure 2.4. Trajectory from the positive and negative states over 100 state updates.

We experimented with four social network topologies: a complete graph, a 2-D lattice, a regular graph and a small world network. The regular graph was of degree 10. Small world (SW) networks have a high clustering coefficient (a higher likelihood that a two neighbors of an agent are also neighbors), while also having a low characteristic path length (mean distance between nodes). These types of networks are known to describe a variety of real world networks, such as the power grid of the U.S. [65]. We generated SW networks using the Watts-Strogatz algorithm with $k = 10$ and $p = .01$ [65].

Unless otherwise stated, results are averaged over 10 runs.

Figure 2.11 shows the time to reach agreement for the three network topologies topologies as f_p ranges from 0.1 to 0.9. We define agreement as when the mean euclidean distance between the valences of all pairs of agents is less than 2% of the maximum distance. The error bars indicate one standard deviation. Means and standard deviations were calculated based only on the runs that reached agreement. Figure 2.12 shows the number of runs out of 10 that reached agreement within $100,000$ iterations. Notice that for $f_p = 0.5$ the lattice, SW and regular networks did *not* converge within $100,000$ iterations and thus no points are provided.

Figure 2.10 shows the valences of all agents for a range of f_p values in $F1/F2$ space on a complete graph for a single run. The positive and negative end points are indicated as well. Each smudge of color is 484 agents, the color indicates the value of f_p. Figure 2.5 is the same as Figure 2.10 except it shows the valences in $F1/X$ space. Both of the figures combined indicate the valences of the entire population.

Figures 2.6,2.7,2.8 show the end states of all agents for a range of f_p values. End states for $f_p = 0.4, 0.5$ were not shown as system rarely converged in those cases. The presented figures are indicative of the ending state of all the runs.

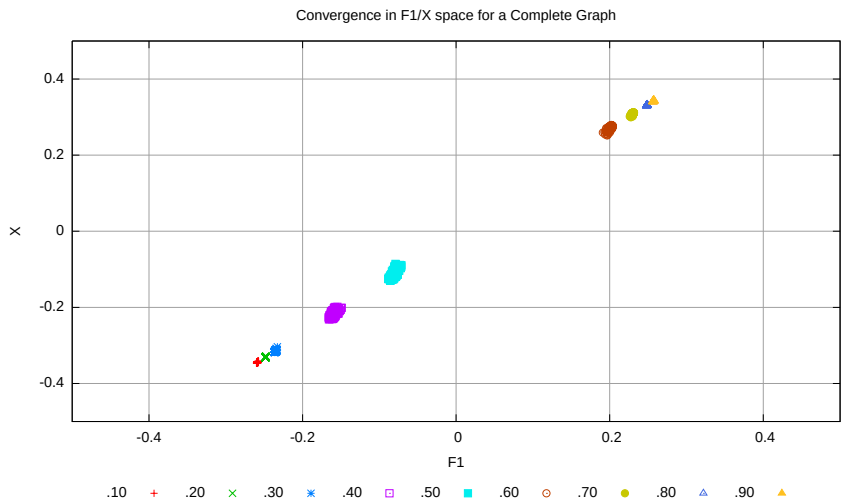

Figure 2.5. Location in F1/X space after 99000 iterations on a complete graph.

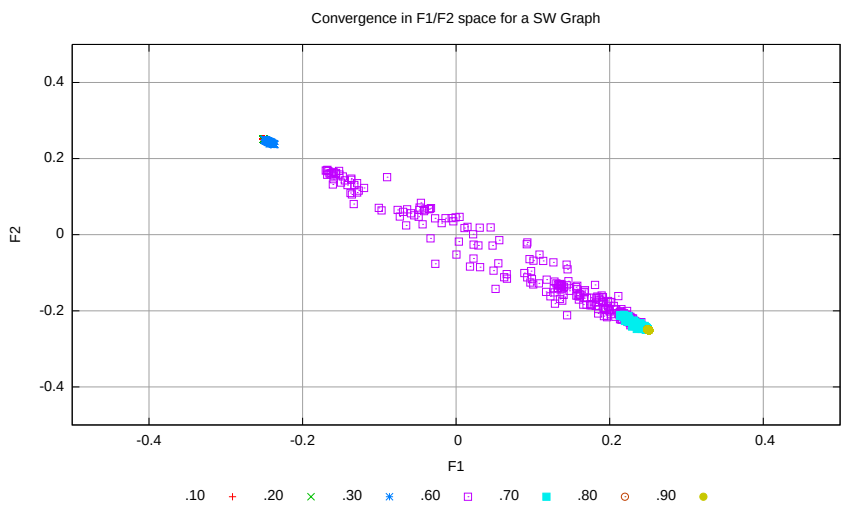

Figure 2.6. Location in F1/F2 space after 99000 iterations on a small world graph.

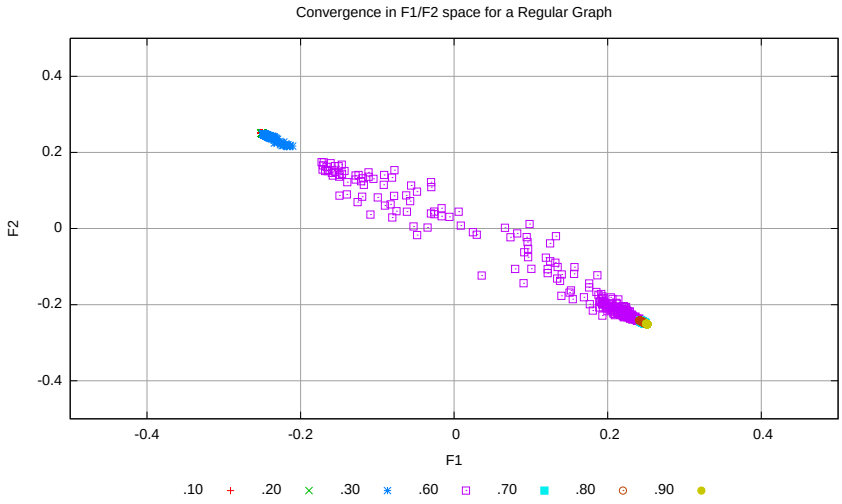

Figure 2.7. Location in F1/F2 space after 99000 iterations on a regular graph.

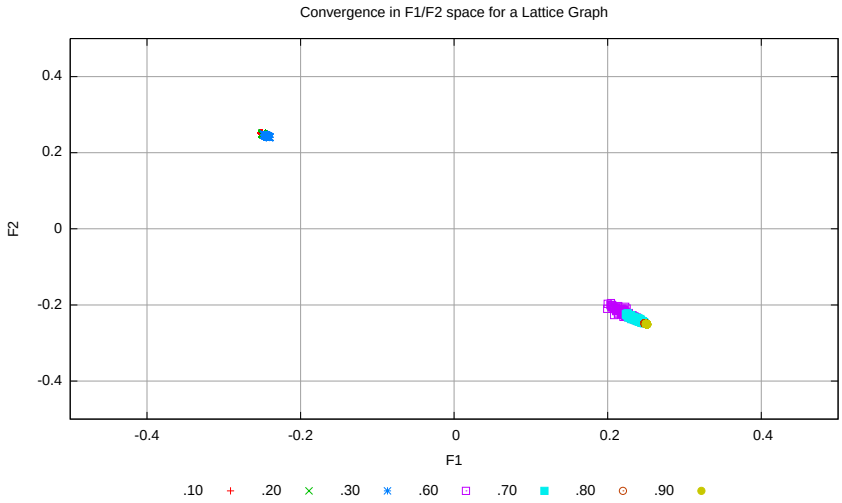

Figure 2.8. Location in F1/F2 space after 99000 iterations on a lattice graph.

Discussion

First, does the system reach agreement, and if so, how long does it take to reach agreement? We can see that in most cases the majority of runs do reach agreement. However, for all non-complete graphs very few runs reached agreement when f_p was close to 0.5. In these cases it seems that the population has split into two groups. Figure 2.9 shows the valences of agents in a run that did not result in agreement for the small world network with $f_p = 0.4$. We can see that two groups have formed that do not seem to coming closer.

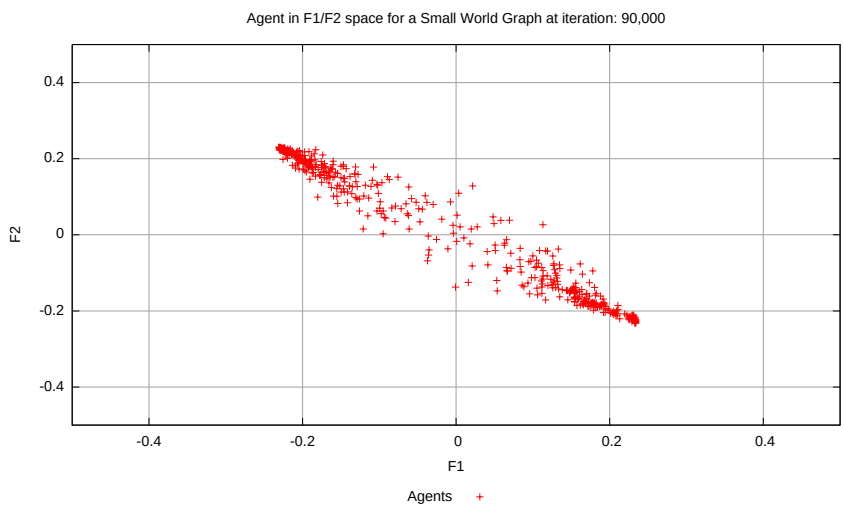

Figure 2.9. Agents in $F1/F2$ space for a run that did not result in agreement at iteration 99000.

The lattice, SW and regular graphs take much longer to reach agreement than the complete graph. This has been noted in a variety of models, including the voter [55], and the majority-state update [9] models. However, in our case the small-world networks takes longer than even the lattice and regular networks to reach agreement. This is counter-intuitive, since SW networks have a lower characteristic path length one would expect a decrease in the time to agreement. This is a matter for further study. Results for the voter model on SW networks can provide some insight into this behavior as in the thermodynamic limit (where $n \to \infty$) agreement does not occur [5].

Note the symmetry around $f_p = 0.5$ in terms of time to agreement. This symmetry is also found in voter models for complete graphs [55].

Secondly, what state does the system agree upon? Figures 2.10,2.6,2.7,2.8 indicate that the agreed upon state is a function of the topology of the social network and f_p.

In the complete graph case we see a linear dependence, as f_p increases the agreed upon state comes closer to the the positive state. However, for non-complete graphs the agreed upon states cluster around the positive and negative endpoints. Below $f_p = 0.4$, the systems

reach agreement, but very close to the negative state, above $f_p = 0.6$, the system reaches agreement as well, but on the positive state.

These results are especially interesting for the SW networks, as these are found in many real-world networks. For the simple cognitive network described here, these results indicate that a population will either split into two groups (for $f_p = 0.4, 0.5, 0.6$) or one attitude will overtake the other attitude. Unlike the complete graph case where attitudes change in strength, we see no such variance here.

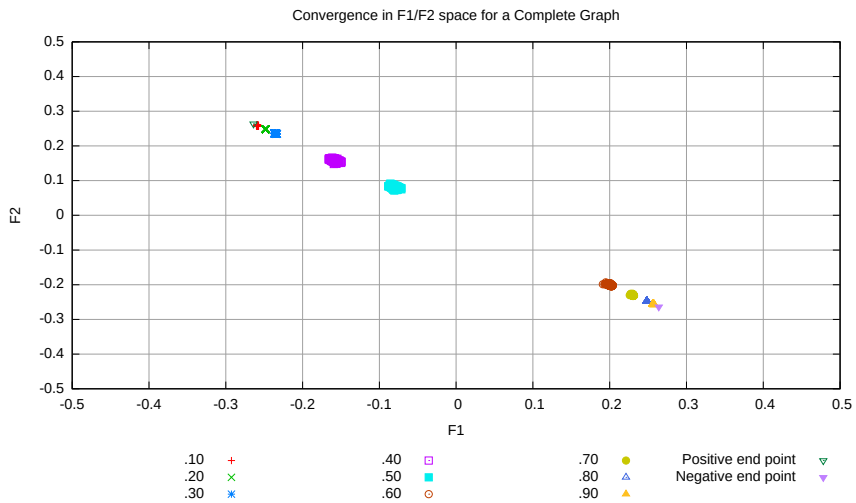

Figure 2.10. Location in F1/F2 space after 99000 iterations on a complete graph. The trajectories from the positive and negative start points are labelled.

The experiments described here served as a preliminary illustration of the model developed in the prequel. We can see, however, from these simulations several important things:

1. We have shown, through numerical experiments, that our model can capture the process of reaching agreement on a concept, even when there are agents with contradictory attitudes

2. We have shown that the distribution of attitudes in a population has a large effect on the time to agreement as well as the strength of the attitude. This is interesting as it allows us to study not only whether populations agree, but on what they agree upon.

3. The social network can have a great affect on the final agreement state – for lattices and small-world and regular graphs we see a polarization effect – agents agree on either the positive or negative endstate.

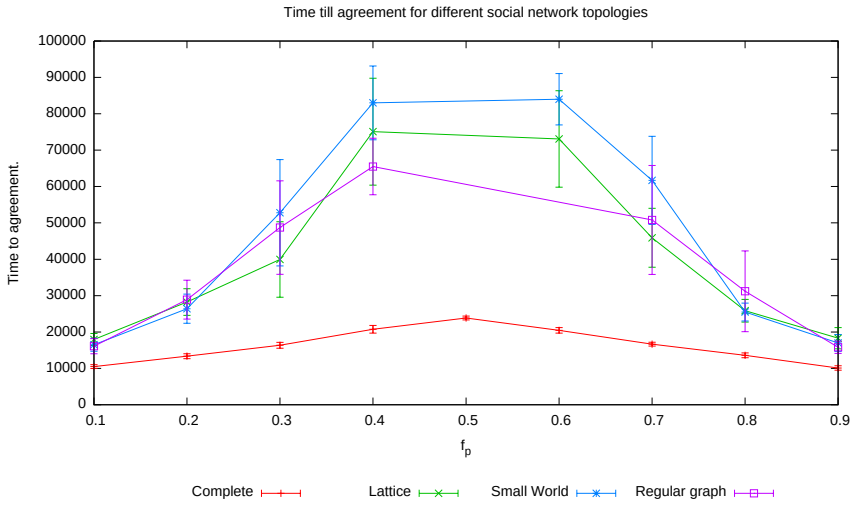

Figure 2.11. f_p vs. Number of iterations till agreement.

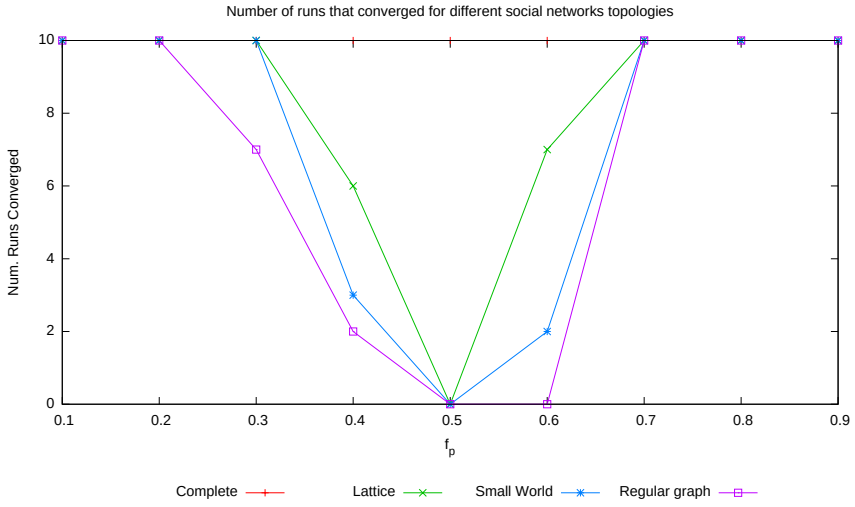

Figure 2.12. f_p vs. Number of runs that reached agreement.

Related Work

Many models have been proposed that focus on the social aspect or the cognitive aspect of attitude change; however not many have put together both of these aspects. While a full review of these are out of scope for this work, we refer the interested reader to [26, 36, 6, 18]. Importantly, most models do not account for a complex cognitive aspect for the agent – individuals are often represented as having binary attitudes.

The communication and interaction dynamics of our model are closely related to the *voter model* a well explored model from the physics domain [55]. In this model there are a set of sites each of which is endowed with a variable that can take either state 0 or state 1. At each timestep a random site changes its state to a randomly chosen neighbors. The voter model has been studied on lattices as well as small-world and heterogenous degree networks [58, 55]. Multiple extensions have been developed that incorporate a majority rule (sites change to the state that is in the majority of their neighborhood) [34, 43, 8]. The main issues with voter models are that they are limited to one concept with only two states; this simplicity is, however, a boon for analysis.

Similar to the voter model, at each timestep we choose one agent which communicates with one of its neighbors. However, our agents are more complicated as they utilize a PCS network.

The conformity-consistency model (CCM) is a model that represents both social and cognitive factors [3, 40]. In the CCM, there are N agents, each endowed with a binary vector of size M that represents their cognitive state. At each time step an agent is chosen and it will execute the standard voter model process on of its variables with probability p; with probability $q = 1 - p$ the agent engages in a voter model *with itself*, between the elements of its cognitive state. The voter process within itself can be viewed as reducing dissonance between the elements of its cognitive state – with the extremes of all variables at 0 or all variables at 1 to be the no dissonance situations.

The CCM is very similar to our model, however the representation of the the cognitive state of an agent is significantly limited in only representing binary values. In addition, the model only captures positive relations between concepts; it does not capture negative relations nor does it specify a strength for the relation.

Chapter 10 of [59] describes the "consensus = coherence + communication" (CCC) model. In this model each agent has a PCS network where concepts represent hypotheses and evidence, and links represent explanatory relations. The purpose of the model was to explain the diffusion of scientific theories in a population, so pairs of agents could interact (simulating a complete network) and agents could execute "lectures" for many other agents.

The main difference between our model and the CCC model is in the interaction. In the CCC model agents exchange concepts with others, thus changing the structure of the network. We have focused on a different perspective, given that individuals already know the links between concepts, how do persuasive messages between agents lead to attitude change.

The parallel constraint satisfaction model has been explored in a wide variety of contexts, such as impression formation [25], legal inference [52, 50, 53], and as a model of change in attitude to the persian gulf war [56].

Conclusions & Future Work

Attitudes play a significant role in determining how individuals process information and behave. In this paper we have developed a new computational model of population wide attitude change that captures the *social* level: how individuals interact and communicate information, and the *cognitive* level: how attitudes and concept interact with each other. The model captures the cognitive aspect by representing each individual as a parallel constraint satisfaction network.

As a way of understanding the dynamics of this model we have explored a simple attitude change experiment. A population of agents was divided into two groups with opposing attitudes. We wanted to know if the population can reach agreement, how long that would take, and what attitudes would the population reach. We found that: (1) the time to agreement varies with the social network and the initial distribution of attitudes in the population, with the complete graph taking the least amount of time to reach agreement; (2) in complete graphs the final state of the population has a linear dependence with the f_p (the proportion of agents initially with a positive attitude); however for non-complete graphs there was a polarization effect where the population converged on the positive or negative endpoint.

We plan on greatly extending and exploring this model in the future. First, we will try to utilize a more realistic social network. Scale-free networks are a widely known topology that also appear in many physical domains [57].

Developing a more realistic cognitive network is an important part of our future work. We intend to draw from existing literature, for instance from [49] which created a network on breast cancer testing decisions that was developed from survey data.

Chapter 3

Attitude resolve and population wide attitude change

This chapter looks at modifying the DACCA model to account for individual differences in susceptibility to influence. This was joint work with Craig M. Vineyard, Joseph Collard, and Stephen J. Verzi of Sandia National Labs. This work was presented as a poster at the 2012 Social-Computing, Behavioral-Cultural Modeling, and Prediction Conference [60].

Introduction

Attitudes are "general and relatively enduring evaluative responses to objects" where objects can be "a person, a group, an issue or a concept" [61, Page 1]. Attitudes are shown to have an impact on, and can sometimes predict, the behaviors of individuals (e.g., voting behavior [22], consumer purchases [20, 14]).

Understanding population wide attitude change is thus an important step to understanding the behavior of societies. For instance, consider the change in attitudes towards global warming and the environment that has resulted in a significant change in public policy and national priorities [7].

While there are a number of factors that influence attitude change [62], we will focus on three in this paper. The first is social – individuals are exposed to various attitudes and information through interaction with others. Family, friends, acquaintances, and the media all influence the attitudes of individuals by providing new information/opinions.

The second factor is cognitive – individuals tend to hold a set of attitudes that are consistent with each other [47, 51, 54]. According to *cognitive consistency theories*, an individual holding a strong positive attitude towards environmentalism should also hold a strong positive attitude towards recycling; if they do not, the attitudes are inconsistent with each other and could cause an uncomfortable feeling (i.e. *cognitive dissonance*) which tends to result in either attitude or behavior change [62].

The third factor is individual differences in influence. Intuitively, some individuals seem more likely to change than others. This intuition has been supported by research from the

marketing and social psychology fields (see Section 3 for more details).

Previous work [28, 27] has described a socio-cognitive model that captures the social and cognitive aspects of attitude diffusion in individual interaction in a social setting. This work extends previous work by incorporating individual differences in influence. In our model we have two types of individuals, "susceptibles" and "advocates". "Susceptibles" represent individuals that are strongly influenced by others. Thus in any interaction they change significantly. "Advocates" represent individuals that do not change as much.

In this paper we present preliminary work on a new model that captures these three elements together to help understand how information, in the form of attitude, is spread in a social setting. In our model, individual differences will be modeled using *game theory*. We describe the basis of this model, then show some simple simulations that explore the dynamics of this model.

Theoretical basis

Intuition and common experiences seems to indicate that people vary in how they are influenced by others. Some people stand firm and rarely change their attitude opinion, while others often vacillate. Research in social psychology and marketing has provided some evidence to this folk psychological idea.

Decades of research has occurred in the marketing domain to understand how consumer purchasing decision are influenced by others. Two types of influence are usually described, (1) informational – where individuals are influenced by obtaining new information from peers; and (2) normative – where individuals are influenced to conform to others decision in order to be liked [2]. The unit of analysis here is on individuals, and the demographic characteristics that can make them change, such as age or gender [17].

Another branch of research has focused more on attitudes themselves rather than individuals. *Strong* attitudes are: "resistant to change, stable over time, and have powerful impact on information processing and behavior" [23, page 279]. Several characteristics of an attitude, such as its importance to an individual or its accessibility, can influence it's strength.

Model Description

An attitude dissemination model consists of various levels of interactions. At the social level, some mechanism selects which agents are to interact with one another. Numerous possibilities exist such as teacher-student interactions, co-worker interactions, or a discussion amongst friends/peers to name a few. Each of these interaction types may have subtle effects with regards to how the selected interacting agents affect one another. Upon selecting the

interacting agents, their affect upon one another must be assessed in some manner. And finally, at the individual level, any internal changes brought about by the social interaction may lead to a cascade of changing beliefs or a resilience of the agent's original attitudes.

Game theory is a branch of mathematics pertaining to strategic interactions. As such, it is applicable to the development of a framework which can address each level of interaction in an attitude dissemination model. In this work, we have investigated applying game theory to the agent interaction level. Before describing our game theoretic approach to agent interaction, we will first give an overview of the cognitive and social aspects of our model (for a more detailed description see [27]).

Within each agent in the model is a Parallel Constraint Satisfaction (PCS) model to represent their individual cognitive network. A PCS model is a type of connectionist, attractor neural network. Nodes within an agent's cognitive network represent concepts, hypotheses, or information. These values are continuous and range from -1 (min) to 1 (max). Weighted links between concepts (w_{ij}) indicate the strength of influence between the concepts. PCS models utilize a connectionist approach to find a consistent set of concepts: the value of each concept is updated according to a non-linear *activation function*. One activation function that is often used is [24],[56]:

$$a_j(t+1) \;=\; a_j(t)(1-d) + \begin{cases} net_j(\max - a_j(t)) & \text{if } net_j > 0 \\ net_j(a_j(t) - \min) & \text{if } net_j \leq 0 \end{cases} \tag{3.1}$$

where:

$$net_j = \sum_i w_{ij} a_i(t) \tag{3.2}$$

and d is a decay term that is set to 0.05. This update rule modifies the value of a concept based on the values of other concepts. An agent synchronously updates all concepts in their network and repeats this process a finite number of times defined as the cognitive effort.

Given this basic inter-agent cognitive framework, our overall model is then composed of three elements:

$$< A, G, C > . \tag{3.3}$$

A represents a set of n agents that represent individuals, C is a *cognitive consistency network* that represents the internal cognition that drives change in an attitude, and G is the social network graph that defines the social structure of the population. Each agent has the same number of concepts (m), and the same weights between concepts (W), however each agent can differ in the value of those concepts.

We assume turn based dynamics for simulation of individual interaction – where at each timestep the following actions are taken:

1. A single agent is randomly chosen from the population, uniformly across all agents.

We call this agent the *speaker*. One random neighbor of the speaker is designated the *hearer*, chosen uniformly from all neighbors.

2. A single concept is chosen as the *topic* of communication.

3. The speaker and hearer communicate with one another regarding the topic of communication. This is the *communication* step.

4. The interacting agents engage in an *information integration* step.

5. The interacting agents modify the values of their individual PCS model cognitive networks in order to find a more consistent set of beliefs. This is the *agent update* step.

In the communication step, our game theoretic approach to agent interactions incorporates agent personality types such that each agent is either an *advocate* or *susceptible*. Much like the canonical Hawk-Dove game, this approach yields differing payoffs depending upon what types of players interact. If an advocate speaks to a susceptible agent, then the advocate will have great success in influencing the suspectible agent towards their attitude with little concession on their own part. However, if two advocates interact with one another, neither is able to persuade the other to budge in their attitude. When two susceptibles interact they are both moderately successful in influencing the other. And so, while being an advocate and aggressively trying to influence the agent you are interacting with is quite successful when dealing with susceptible agents it is a poor strategy in dealing with another advocate. The level of influence one agent has on another is the information integration step. Consequently, new strategic dynamics emerge regarding societal interactions.

Formally, the following update equations govern this agent interaction. Let $x_i(t+1)$ be the activation of node x for agent i at time $t+1$. Let $u_a \in [0,1]$ be the strength of influence of an advocate. Let $u_s \in [0,1]$ be the strength of influence of a susceptible agent. Let $u_a + u_s = 1$, where the update equations are:

$$x_i(t+1) = \begin{cases} u_a x_i(t) + u_s x_j(t) & \text{for Advocate } i \text{ and Susceptible } j \\ u_a x_j(t) + u_s x_i(t) & \text{for Susceptible } i \text{ and Advocate } j \\ 0.5x_i(t) + 0.5x_j(t) & \text{for Susceptibles } i \text{ and } j \\ x_i(t) & \text{for Advocates } i \text{ and } j \end{cases} \tag{3.4}$$

Experimental Results

Parameterizations:

In order to select functionally realistic values for the influence strengths of both the advocate and susceptible agent types we performed an empirical value sweep and found intuitively plausible behaviour to occur with an advocate agent influence strength (u_a) of 0.7 and susceptible agent influence strength (u_s) of 0.3. Weights between concepts in each

agent's PCS network were set at 0.05, and we selected a small percentage (10% to 30%) of the population of agents to be of the advocate type while the rest became susceptible agents.

Discussion

Experiment 1: Number of Advocates

To visualize and more easily understand the influential power of advocates who may significantly change susceptible agents and are not easily influenced themselves, we experimented with small, randomly connected networks and varied the percentage of advocates in the overall population (see results below). We used the same three node cognitive consistency network in each agent and initialized all of their networks to be cognitively consistent with an attitude value ranging from zero to the minimum defined by the negative attractor value (determined by the weight values in the PCS model). Then we selected a minority percentage of agents to adopt the opposing positive attitude value (ranging between 0 and the maximum postive attractor value, again determined by the weights in the PCS model). The negative attitude value agents are susceptible agents and the positive attitude value agents are advocates.

Iterating over this initialization, we investigated whether 10, 20, or 30 percent positive attitude advocate agents would be sufficient to persuade the opposing majority of susceptible agents to change their attitude belief or conversely if the majority of susceptible agents would be able to overwhelm the minority advocates.

Experiment 2: Position of Advocates

Beyond investigating the affect of the number of influential advocates in a network, we also examined the significance of the particular position of an advocate within the social network topology. To place advocates in a meaningful location we used small hierarchical and ring network topologies in which we could manually determine specific agents as advocates due to their connectivity or position within the topology.

Experiment 1

We have used a small ten node, randomly connected network so that it is tractable to visualize and analyze the interactions and attitude values over each time step of our simulation. A single advocate (ten percent of the population) was never able to convert a single susceptible agent and eventually is overwhelmed to adopt the negative attitude itself. Using two advocates (twenty percent of the population) allows for possible advocate-advocate interactions as well as doubles the possible conversion influence. We began by randomly

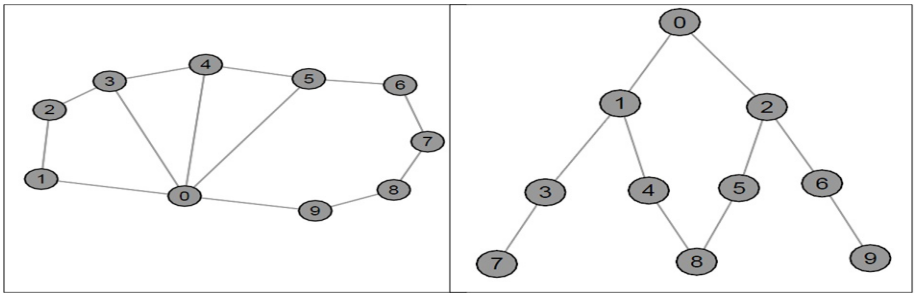

Figure 3.1. Small ring and hierarchical network topologies

selecting two random agents as advocates and as before, despite their increased influential capabilities, the two minority advocates were not sufficient to persuade the majority of susceptible agents. Selecting two connected agents, each with low social connectivity, as the advocates was sometimes able to convert the opposing eight susceptible agents. Increasing the advocate population to three (thirty percent) typically was sufficient to convert the seven opposing attitude susceptible agents simply by using randomly selected advocates.

Experiment 2

To study the significance of the position of advocates within a network, we used the small ring and hierarchical network topologies shown on the left and right side of Figure 3.1 respectively.

For the ring network we experimented with setting agent 0 as the sole advocate since it is the most connected agent. However, despite being the most connected agent it was unable to overcome the nine opposing attitude susceptibles. As one might expect, it was able to raise the attitude values of the agents it was connected to, but not sufficiently enough to convert any of them to the postive attitude. Additionally, we also experimented with setting agent 2 as the sole advocate. As a member of the perimeter of the ring which only interacts with its two adjacent neighbors agent 2 was also unable to convert the opposing 9 susceptibles. Finally, setting both agents 0 and 5 as advocates was typically enough to convert the other 8 agents despite the fact that neither is directly connected to many of the other agents in the ring. In this pairing, agents 0 and 5 were able to reinforce one another and eventually persuade the other agents despite being outnumbered.

Figure 3.2 depicts the affect of different advocate positions within a network. This image plots each agent's attitude value across time. On the left half of Figure 3.2, two arbitrary agents were selected as advocates, and were unable to persuade the other agents to adopt the positive attitude. In fact these advocates were themselves converted. This is illustrated by by the full convergence of all agent attitude values to the negative attractor at -0.33. Alternatively, the right half of Figure 3.2 plots the simulation results where agents 0 and

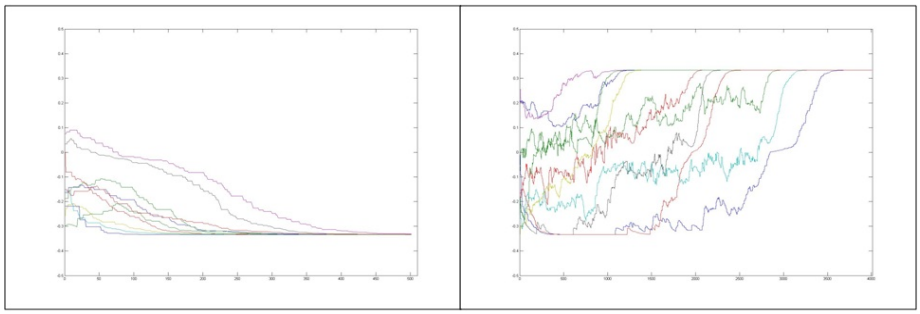

Figure 3.2. Drastically different results for different advocate positions within a network

5 were selected as advocates. Although they were the only two agents initially in favor of the positive attitude value, over time they are successfully able to pull all the other agents attitude values up to the positive attractor at 0.33.

For the hierarchical network (seen in the right half of Figure 3.1), we likewise began by experimenting with single agents as the lone advocate and investigated whether the position at the topmost node in the network would be sufficient to propagate down and convert the rest of the nodes, or whether being a bottom node would be able to spread upwards and convert the others. In both cases, neither sole advocate was sufficient to overcome the other nine susceptibles. Alternatively, utilizing both a top down and bottom up influence by setting agents 0, 7, and 9 as advocates was largely sufficient to spread their attitude and convert all of the remaining agents despite the fact that their influence is constrained such that they can only interact with one or two other agents whom they must convert and in affect those most continue the spread with their further neighbors.

Conclusion & Future Work

Conclusion

In this paper we have presented a new computation model of attitude resolve which incorporates the affects of player interaction dynamics using game theory in an integrated model of socio-cognitive strategy-based individual interaction. Running a variety of experiments on this model with different network topologies and population mixtures we have investigated attitude adoption and resolve within small networks. These experiments have indicated that selecting specific advocate locations within a population tends to result in a stronger spread of influence, even with only a few advocates in a small network, as opposed to random placement. And the enhanced spread of influence can be observed by the ability to convert agents with opposing attitude values that greatly outnumber the advocates as well

as through lower diffusion time (iterations in the PCS network - also referred to as cognitive effort) for the adoption of the opposing attitude.

Future Work

Having investigated interaction dynamics on computationally tractable small networks, we would like to extend this work to larger real world networks to assess whether the same behavior occurs. Additionally, we can also extend the model to incorporate factors such as more sophisticated cognitive networks.

In addition to incorporating game theory at the agent level we are also interested in looking at the affects of applying game theory at both the social and individual, cognitive level. At the social level, agent interactions are typically dictated by given static social network topology. An alternative approach would be to investigate coalition formation games from coalitional game theory. This particular branch of game theory seeks to build network structure in a strategic manner. Furthermore, if a particular agent, selected as a speaker, is seeking to maximize the spread of their influence they might want to select the hearer they share their attitude view with strategically rather than randomly from everyone they know. Various aspects of game theory such as coalitional graph games or network routing could provide an alternative framework to model this interaction.

As an alternative mechanism to implement parallel constraint satisfaction (PCS) attitude updates within a single agent, a game may be played amongst two attitudes competing to influence one another. Rather than requiring multiple iterative updates the effect is achieved by a single instance of the game. For attitude nodes connected to (influenced by) several other nodes, the game is played simultaneously with all neighboring nodes and the net attitude change is a synchronous aggregated update of all games played in a pairwise interaction with neighbor nodes.

Chapter 4

Online Data for Validation

Validation of agent based models is an important issue, especially with models of social systems that could influence decision making at higher levels. As part of this project, we explored using online data sources to generate data for validation. In the rest of this chapter we will outline an new, innovative experiment we designed to gather data using Amazon Mechanical Turk (AMT).

Amazon Mechanical Turk

Amazon Mechanical Turk (AMT) is an online marketplace of tasks. Workers can login to the site and complete Human Intelligence Tasks (HITs) which are posted by requesters. Anyone can be a worker and/or requester. Workers receive small payments for completing tasks, on the order of $.10 per task. After workers complete a task, the requester can review the work and accept or reject the task. Requesters can also provide bonuses for good work. HITs vary significantly in terms of content. Some HITs are simple image classification tasks while other HITs require evaluating/summarizing text content, or completing surveys. Once workers complete an HIT, the requester can decide whether to accept (and pay the described amount) or reject the HIT. The number of rejected HITs is tracked for individual workers. Several surveys of workers on AMT have been conducted[35] ; the answers are not validated but can provide insight into the worker populations. These surveys indicate that (1) workers come from many countries (¿100, with the majority of workers from India and the U.S.); (2) workers have a mean age of 32, median of 30; (3) the majority of workers earn around $30K, some ¿$100K; (4) of workers who chose to give their gender, 55% female, 45% male. These results indicate that the worker population is highly diverse, especially as compared to standard population used in university based laboratory experiments. Requesters can require HITs to be completed by specific types of workers. This is implemented through a qualification system. Requesters set certain qualifications workers must meet. Amazon provides system wide qualifications such as location that one is logging in from, whether the user can view adult content, and a minimum percentage of accepted HITs. Requesters can also create their own worker qualifications.

AMT is rapidly becoming a popular method for conducting experiments due to its large and diverse subject pool, inexpensiveness, and the rapidity with which to get results. For

instance, Paolacci, et. al [41] replicated the results of several classic psychology studies on framing effects, the conjunction fallacy, and outcome bias. Experiments can be rapidly completed; Mason and Suri [35] show that for a simple survey, at a cost of $.05 per survey response, 500 people responded within 10 days. A laboratory experiment of the same size would take months to complete and be extremely expensive. The location qualification of AMT allows us to conduct experiment specifically designed to understand the impact of culture on behavior. For instance, Ericksson and Simpson [12] studied risk preferences as a function of gender and culture (US or India).

Overview of Design

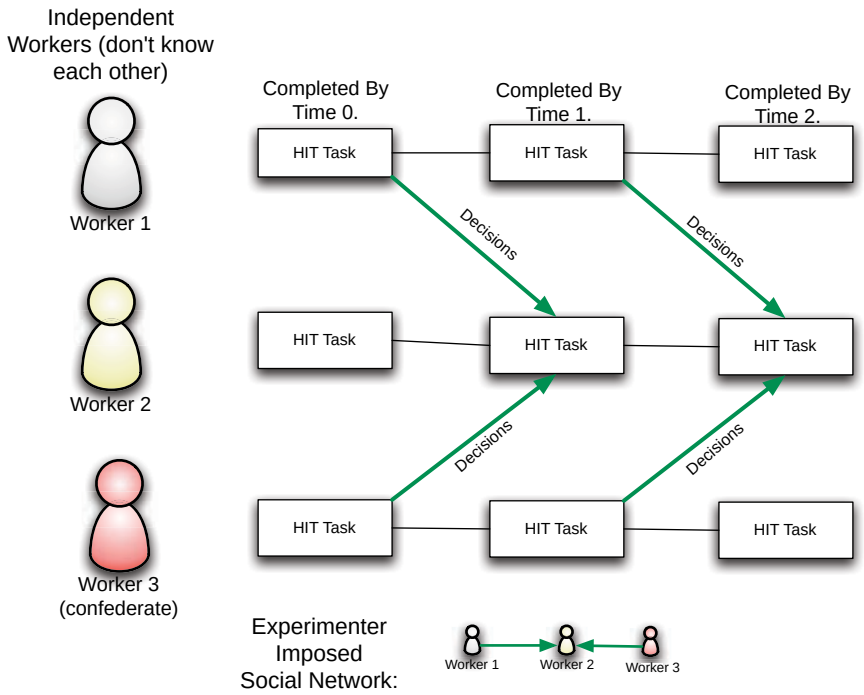

Figure 4 illustrates the general design of experiments we have in mind. We can view the experiment as a discrete time system. At time point 0, worker 0 logs into AMT and chooses to execute one of our tasks, for instance a task in which workers have to pick a product to purchase from a set of products. When first executing the task, the worker will not get any information about the product choices of others. At time point 1, worker 0 will again choose to execute our task. At this point, worker 1 gets information about the choices that other workers made on the task from the previous time point (green arrows in Figure 1) so the product decisions they made. By manipulating who the worker gets information from, we can impose different social networks. In Figure 1, Worker 2 gets information about the results of both Worker 1 and 3, but Workers 1 and 3 do not get any information from anyone else. The first key aspect of this design is the imposition of a social network rather than elicitation of a social network. By manipulating the information workers see we can impose and quickly change the social network of individuals allowing us to systematically identify

critical topological factors. The second key aspect is the use of confederates experimenter controlled entities to manipulate and systematically perturb the system. Most studies of behavioral influence are post-hoc data is collected after an event, and thus one has to wait for an interesting event to occur. This might occur rarely or not at all. In our design, we can manually insert events via confederates. By manipulating the location of the confederates within the social network we can systematically test different hypotheses about diffusion and attitude change over populations. The third key aspect is the asynchronous interaction workers are executing tasks at different times, yet sharing information between tasks. In related experiments workers were asked to join a waiting room, when enough workers had joined all workers completed tasks together [37]. This allowed rapid completion of tasks and ease in sharing information. However, this limited the number of participants because of the difficulty of getting workers together at the same time. In our design, individuals can finish a task within some short period of time, say 24 hours, and then can receive a new task. This means we can scale to larger numbers of workers, and can experiment with cross-cultural populations of workers (since members of different countries can complete tasks during their comfortable times).

The main drawback to this experimental design is that the constructed social network is based on imposed ties; it is not clear yet whether individuals will be influenced similarly if they were interacting with their actual friends or family . Laboratory studies on social influence suggest that workers will be influenced by random other individuals in the population. In addition, we also know that many people interact with semi-random individuals via Online media. For instance, the number of Facebook friends is often much higher than actual friends. This means individuals are routinely exposed to messages by semi-random people online, as well as messages from their friends.

Experiment Design

In collaboration with Ann Speed, we designed an experiment to test the diffusion of attitudes related to environmentalism. The procedure is outlined below. The goal is to measure the change in attitude strength (measured through a Likert scale) as a function of the social network topology and number of exposures to contrasting beliefs.

To assess attitudes towards environmentalism we will use the revised NEP scale [11]. This scale has items designed to evaluate 5 dimensions of ecological worldview: the reality of limits to growth, antianthropocentrism, the fragiliy of nature's balance, rejection of exemptionalism and the possibility of an ecocrisis. Each dimension as approximately 2-3 questions assigned to it.

The design of this particular experiment relies on providing subjects contrasting information about specific dimensions of the ecological worldview and then testing their attitudes. As described above, this contrasting information could come from experimenter created confederates or other subjects.

Procedure

- Informed consent
- Subject takes an attitude test.
- Experimenter assigns subject to a place in the network structure (or excludes subject)
  - Subjects that are in the middle on the attitude and on the average of dimensions for each attitude will be excluded from further participation.
  - Subjects who are at the extremes will be assigned to the key parts of the network, surrounded by people who have the opposite views.
  - If there are not enough extreme people, we will use confederates.
- Subjects are introduced to experiment. Subjects told that they will interact with other subjects at the end of the experiment for a bonus. Subjects will be told they will be providing their opinions about certain beliefs and behaviors and will be reading opinions about similar beliefs and behaviors of their team members. They will be told they will answer questions about their team members beliefs.
- Subjects are presented with material on 2 dimensions of the attitude, which they summarize.
- Subjects are presented with the opinions of other subjects on 2 dimensions of the attitude.
- Once subjects read their other subjects answers and asked questions about the opinions of the others.
- Process repeated.
- At the end, there will be a final attitude test.

References

[1] Jonathan K. Alt, Leroy A. 'Jack' Jackson, David Hudak, and Stephen Lieberman. The cultural geography model: Evaluating the impact of tactical operational outcomes on a civilian population in an irregular warfare environment. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 6:185, 2009.

[2] William O. Bearden, Richard G. Netemeyer, and Jesse E. Teel. Measurement of consumer susceptibility to interpersonal influence. *Journal of Consumer Research*, 15, 1989.

[3] Jenna Bednar, Aaron Bramson, Andrea Jones-Rooy, and Scott Page. Conformity, consistency, and cultural heterogeneity. Mimeo, University of Michigan, 2006.

[4] Kathleen M. Carley. Validating computational models. Technical report, Carnegie Mellon University, 1996.

[5] C. Castellano, D. Vilone, and A. Vespignani. Incomplete ordering of the voter model on small-world networks. *Europhysics Letters*, 63(1):153–158, 2003.

[6] Claudio Castellano, Santo Fortunato, and Vittorio Loreto. Statistical physics of social dynamics. *Reviews of Modern Physics*, 81, April-June 2009 2009.

[7] Yale Project On Climate Change. Climate change in the american mind: Americans' climate change beliefs, attitudes, policy preferences, and actions. Technical report, George Mason University, 2009.

[8] P. Chen and S. Redner. Majority rule dynamics in finite dimensions. *Physical Review E*, 71, 2005.

[9] Jordi Delgado. Emergence of social conventions in complex networks. *Artificial Intelligence*, 141(1-2):171–185, October 2002.

[10] John F. Dovidio, Kerry Kawakami, and Samuel L. Gaertner. Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, 82(1):62–68, 2002.

[11] Riley E. Dunlap, Kent D. Van Liere, Angela G. Mertig, and Robert Emmet Jones. Measuring endorsement of the new ecological paradigm: A revised nep scale. *Journal of Social Issues*, 56(3):425–442, 2000.

[12] Kimmo Eriksson and Brent Simpson. Emotional reactions to losing explain gender differences in entering a risky lottery. *Judgement and Decision Making*, 5(3), June 2010.

[13] Russell H. Fazio, J. R. Jackson, B. C. Dunton, and C.J. Williams. Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology*, 69:1013–1027, 1995.

[14] Russell H. Fazio and Michael A. Olson. Attitudes: Foundations, functions, and consequences. In M. A. Hogg and J. Cooper, editors, *The Sage Handbook of Social Psychology*. Sage, 2003.

[15] B. Gawronski and G.V. Bodenhausen. Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin*, 132:692–731, 2006.

[16] B. Gawronski and F. Strack. On the propositional nature of cognitive consistency: Dissonance changes explicit but not implicit attitudes. *Journal of Experimental Social Psychology*, 40:535–542, 2004.

[17] Tulay Girard. The role of demographics on the susceptibility to social influence: A pretest study. *Journal of Marketing Development and Competitiveness*, 5(1), 2010.

[18] Robert L. Goldstone and Marco Janssen. Computational models of collective behavior. *Trends in Cognitive Science*, 9(9), 2005.

[19] J. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences of the United States of America*, 79:2554–2558, 1982.

[20] Christopher R. M. Jones and Russell H. Fazio. Associative strength and consumer choice behavior. In C. P. Haugtvedt, P. M. Herr, and F. R. Kardes, editors, *Handbook of Consumer Psychology*, pages 437–459. Psychology Press, 2008.

[21] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of 2003 SIGKDD*, 2003.

[22] Jon A. Krosnick. The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluation, and voting behavior. *Journal of Personality and Social Psychology*, 55(2):196–210, 1988.

[23] Jon A. Krosnick and Wendy R. Smith. Attitude strength. In V.S. Ramachandran, editor, *Encyclopedia of Human Behavior, Volume 1*, volume 1. Academic Press, 1994.

[24] Ziva Kunda and Paul Thagard. Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103(2):284–308, 1996.

[25] Ziva Kunda and Paul Thagard. Forming impressions from stereotypes, traits, and behaviors: A parallel-constraint-satisfaction theory. *Psychological Review*, 103(2):284–308, 1996.

[26] Kiran Lakkaraju. *Agreement, Information and Time in Multiagent Systems*. PhD thesis, University of Illinois at Urbana-Champaign, 2009.

[27] Kiran Lakkaraju and Ann Speed. A cognitive-consistency based model of population wide attitude change. In *Proceedings of the 2010 AAAI Fall Symposium on Complex Adaptive Systems*, 2010.

[28] Kiran Lakkaraju and Ann Speed. Key parameters for modeling information diffusion in populations. In *Proceedings of the 2010 IEEE Homeland Security Technologies Conference*. IEEE, 2010.

[29] Kiran Lakkaraju and Ann Speed. Population wide attitude diffusion in community structured graphs. In *Proceedings of the 2011 AAAI Fall Symposium on Complex Adaptive Systems*, 2011.

[30] Kiran Lakkaraju and Ann Speed. A cognitive-consistency based model of population wide attitude change. In *Complex Adaptive Systems: Views from the Physical, Natural, and Social Sciences*. Springer-Verlag, 2012.

[31] Kiran Lakkaraju, Samarth Swarup, and Les Gasser. Consensus under constraints: Modeling the great english vowel shift. In *Proceedings of the 2012 Social-Computing, Behavioral-Cultural Modeling, and Prediction Conference*, 2012.

[32] Mark R. Lepper and Thomas R. Shultz. Dissonance. In Robert A. Wilson and Frank C. Keil, editors, *The MIT Encyclopedia of Cognitive Science*. MIT Press, 1998.

[33] Lindsey Clark Levitan and Penny S. Visser. The impact of the social context on resistance to persuasion: Effortful versus effortless responsees to counter-attitudinal information. *Journal of Experimental Social Psychology*, 44:640–649, 2008.

[34] F.W.S. Lima, A.O. Sousa, and M.A. Sumour. Majority-vote on directed erdos-renyi random graphs. *Physica A*, 387:3503–3510, 2008.

[35] Winter Mason and Siddharth Suri. Conducting behavioral research on amazon's mechanical turk. SSRN eLibrary.

[36] Winter A. Mason, Frederica R. Conrey, and Eliot R. Smith. Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and Social Psychology Review*, 11, 2007.

[37] Winter A. Mason, Andy Jones, and Robert L. Goldstone. Propagation of innovations in networked groups. *Journal of Experimental Psychology: General*, 137(3):422–433, 2008.

[38] A. R. McConnell and J. M. Leibold. Relations among the implicit association test, discriminatory behavior, and explicit measures of racial attitudes. *Journal of Experimental Social Psychology*, 37:435–442, 2001.

[39] R. Olfati-Saber and R.M. Murray. Consensus problems in networks of agents with switching topology and time-delays. *Automatic Control, IEEE Transactions on*, 49(9):1520–1533, 2004.

[40] Scott E. Page, Leonard M. Sander, and Casey M. Schneider-Mizell. Conformity and dissonance in generalized voter model. *Journal of Statistical Physics*, 128:1279–1287, 2007.

[41] Gabriele Paolacci, Jesse Chandler, and Panagiotis G. Ipeirotis. Running experiments on amazon mechanical turk. *Judgement and Decision Making*, 5(5), August 2010.

[42] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical Review E*, 63(066117), 2001.

[43] Luiz F. C. Pereira and F. G. Brady Moreira. Majority-vote model on random graphs. *Physical Review E*, 71(016123), 2005.

[44] Stephen J. Read, Chadwick J. Snow, and Dan Simon. Constraint satisfaction processes in social reasoning. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society*, 2003.

[45] Stephen J. Read, Eric J. Vanman, and Lynn C. Miller. Connectionism, parallel constraint satisfaction processes, and gestalt principles: (re)introducing cognitive dynamics to social psychology. *Personality and Social Psychology Review*, 1(1):26–53, 1997.

[46] Everett M. Rogers. *Diffusion of Innovations*. Free Press, fifth edition, 2003.

[47] J. Edward Russo, Kurt A. Carlson, Margaret G. Meloy, and Kevyn Yong. The goal of consistency as a cause of information distortion. *Journal of Experimental Psychology: General*, 2008.

[48] Robert J. Rydell and Allen R. McConnell. Understanding implicit and explicit attitude change: A systems of reasoning analysis. *Journal of Personality and Social Psychology*, 91(6):995–1008, 2006.

[49] Yuichi Shoda and Walter Mischel. Personality as a stable cognitive-affective activation network: characteristic patterns of behavior variation emerge from a stable personality structure. In Stephen J. Read and Lynn C. Miller, editors, *Connectionist Models of Social Reasoning and Social Behavior*. Lawrence Erlbaum, 1998.

[50] Dan Simon. A third view of the black box: Cognitive coherence in legal decision making. *University of Chicago Law Review*, 71(511), 2004.

[51] Dan Simon and Keith J. Holyoak. Structural dynamics of cognition: From consistency theories to constraint satisfaction. *Personality and Social Psychology Review*, 6(6):283–294, 2002.

[52] Dan Simon, Daniel C. Krawczyk, Airom Bleicher, and Keith J. Holyoak. The transience of constructed preferences. *Journal of Behavioral Decision Making*, 21(1):1–14, 2008.

[53] Dan Simon, Daniel C. Krawczyk, and Keith J. Holyoak. Construction of preferences by constraint satisfaction. *Psychological Science*, 15(5):331, 2004.

[54] Dan Simon, Chadwick J. Snow, and Stephen J. Read. The redux of cognitive consistency theories: Evidence judgments by constraint satisfaction. *Journal of Personality and Social Psychology*, 86(6):814–837, 2004.

[55] V. Sood, Tibor Antal, and S. Redner. Voter models on heterogenous networks. *Physical Review E*, 77, 2008.

[56] Barbara A. Spellman, Jodie B. Ullman, and Keith J. Holyoak. A coherence model of cognitive consistency: Dynamics of attitude change during the persian gulf war. *Journal of Social Issues*, 49(4):147–165, 1993.

[57] Steven H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, March 2001.

[58] Krzysztof Suchecki, Victor M. Eguiluz, and Maxi San Miguel. Voter model dynamics in complex networks: Role of dimensionality, disorder and degree distribution. *ArXiv*, 2008. cond-mat/0504482v1.

[59] Paul Thagard. *Coherence in Thought and Action*. Life and Mind: Philosophical Issues in Biology and Psychology. MIT Press, 2000.

[60] Craig M. Vineyard, Kiran Lakkaraju, Joseph Collard, and Steve J. Verzi. The impact of attitude resolve on population wide attitude change. In *Proceedings of the 2012 Social-Computing, Behavioral-Cultural Modeling, and Prediction Conference*, 2012.

[61] Penny S. Visser and Lindsey M. Clark. Attitudes. In A. Kuper and J. Kuper, editors, *Social Science Encyclopedia*. Routledge, 2003.

[62] Penny S. Visser and Joel Cooper. Attitude change. In M. Hogg and J. Cooper, editors, *Sage Handbook of Social Psychology*. Sage Publications, 2003.

[63] Penny S. Visser and Robert R. Mirabile. Attitudes in the social contex: The impact of social network composition on individual-level attitude strength. *Journal of Personality and Social Psychology*, 87(6):779–795, 2004.

[64] Kyle Wagner, J.A. Reggia, J. Uriagereka, and G.S. Wilkinson. Progress in the simulation of emergent communication and language. *Adaptive Behavior*, 11:37–69, 2003.

[65] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442, June 1998.

[66] Timothy D. Wilson, Samuel Lindsey, and Tonya Y. Schooler. A model of dual attitudes. *Psychological Review*, 107(1):101–126, 2000.

DISTRIBUTION:

	MS	,
		,
1	MS 0899	Technical Library, 9536 (electronic copy)
1	MS 0359	D. Chavez, LDRD Office, 1911

Sandia National Laboratories