
DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU


HARVARD LIBRARY
Office for Scholarly Communication

Towards a General Purpose Political Communications Simulator

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

Citation	Glikstein, Ezequiel Dan. 2019. Towards a General Purpose Political Communications Simulator. Master's thesis, Harvard Extension School.
Citable link	https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364570
Terms of Use	This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Towards a General Purpose Political Communications Simulator: A Computational Model of Information and Attitudes

E. Dan Glikstein

A Thesis in the Field of International Relations
for the Degree of Master of Liberal Arts in Extension Studies

Harvard University

November, 2019

©2019 Ezequiel Dan Glikstein

Abstract

This thesis introduces a new way of modelling information for the purposes of agent-based political simulations. By representing information in the proposed fashion, it becomes possible for agents to compute the attitudinal effects of exposure to information. Moreover, the manner in which information is modelled allows for agents to hold attitudes towards an arbitrary number of objects – not only towards traditional entities of interest like the president of the United States, but also towards other agents and to individual pieces of information. By employing vector algebra to quantitatively represent attitudes, an arbitrary number of attitudinal dimensions such as sentiment or importance may also be represented in a simulation and influenced by exposure to information. As such, this method of representing information unlocks multi-dimensional multi-issue multi-agent simulations of attitude formation, reinforcement, and change. This approach to information enables future research into simulations where the critical elements of political communication theory are directly involved. A considerable amount of effort is devoted to recasting questions of priming, framing, agenda-setting, credibility, and salient characteristics of the American public into the computable language of this information model. I demonstrate that theoretical and empirically-validated results of the political communication literature with regards to attitude change and reinforcement – and formation, to a lesser extent – are guaranteed to be reproduced in simulation. The mathematical approach is complemented by a proof-of-concept which shows the expected differential attitudinal effects of equivalent information on a polarized artificial population is naturally emergent and due to the manner in which information is represented in the simulation. Classical emergent phenomena of opinion leadership and attitude convergence were observed and these were also caused by the use of the information model.

Acknowledgements

I am deeply indebted to Harvard University which has given me the freedom to pursue my intellectual curiosity and the friends who shared and encouraged my passions. I am grateful for the nurturing and I am proud to claim it as my *alma matter*. I extend my gratitude to Dr. Matthew Baum whose thought-provoking questions and enthusiasm sustained the motivation to research this topic when it dipped – thank you for believing in me. I am thankful for Dr. Wander Jager's encouragements and guidance, and for the inspiration to apply simulation studies to attitudes. I am thankful of Remo Aldiari and Mimi Goss' encouragement of free-expression and for refining the interest in rhetoric and communications which was sparked by John Aylen, so long ago, at the John Molson School of Business. I wish to thank Dr. Kryzanowski of Concordia and Dr. Neugeboren of Harvard for encouraging research into simulation models and their faith that it is a worthwhile endeavour that may bring insights into real world social phenomena. I am grateful Dr. Fabian Bastin at Université de Montréal taught me the foundations of simulations. I thank my uncle Diego Sztainberg for his insights into software engineering and his now-obvious remark to separate data from logic. I would also like to thank Dr. Robert Shapiro of Columbia for his numerous insights into public opinion and his encouragements. Byron Reynolds, Mieke Bon, Mike Honigstein, Joseph Scott McArdle, Anna-Rose Dijkstra, Nadya Nifantyeva, Charlotte Davidson, Matthias Knaelmann : I love you all! Thank you for the wonderful memories, warm friendships, and having taught me so much. I would never have had the opportunity to write this thesis without my parents and to them I express my deepest gratitude; I also wish to remark the unwavering support and patient listening of my grandfather. A.M.D., DMD: moi aussi.

Contents

Abstract iii

Acknowledgements iv

1 Introduction 1

I Information Models 6

2 Language Model 7

2.1 Introduction . . . 7

2.2 Subject-Verb-Object statements . . . 8

2.2.1 Notation . . . 8

2.2.2 Keywords . . . 8

2.2.3 Verbs to inform meaning . . . 9

2.3 Other types of sentences . . . 11

2.3.1 Conditional sentences . . . 11

2.3.2 Imperative sentences . . . 12

2.3.3 Interrogative sentences . . . 12

2.3.4 Exclamatory sentences . . . 12

2.4 Bundling Information for exchange . . . 13

2.4.1 Broadcasts . . . 13

2.4.2 Messages . . . 14

2.4.3 Queues . . . 15

2.5 Using real broadcasts . . . 16

2.5.1 Limitations of the model . . . 16

2.5.2 Vast data gathering operation . . . 17
2.5.3 Automated coding . . . 18

3 Agent Knowledge Model 20

3.1 Introduction . . . 20
3.2 Knowledge and memory . . . 23
3.2.1 Accumulation . . . 23
3.2.2 Forgetfulness and learning . . . 23
3.2.3 What an agent believes . . . 26
3.2.4 Spontaneous recall . . . 27
3.2.5 Continuous attitude changes . . . 27
3.3 Characteristics . . . 28
3.3.1 Physical characteristics . . . 28
3.3.2 Geographical location . . . 29
3.3.3 Religion . . . 30
3.3.4 Race . . . 32
3.3.5 Education . . . 34
3.3.6 Core values . . . 35
3.3.7 Partisanship . . . 35

4 Adding to Agent Knowledge 37

4.1 Introduction . . . 37
4.2 Credibility . . . 38
4.2.1 Credibility from homophily . . . 40
4.2.2 Credibility from belief of source knowledge . . . 41
4.2.3 Credibility from costly rhetoric . . . 43
4.2.4 Credibility from reputation . . . 44
4.2.5 Credibility from the medium . . . 45
4.2.6 Credibility from direct experience . . . 46
4.3 Framing effects . . . 46

5 A Mathematical Model of Attitudes **49**

5.1 Introduction . . . 49

5.2 Definitions . . . 50

5.2.1 The attitude vector and vector space . . . 50

5.2.2 Two verbs . . . 51

5.2.3 Inferring attitude and constructing attitude moments f^k . . . 53

5.2.4 Accepting/Ignoring/Rejecting information . . . 54

5.2.5 Attitude reinforcement and change . . . 55

5.2.6 Attitude formation . . . 55

5.3 Propositions . . . 56

5.3.1 In a general context . . . 56

5.3.2 In a world with a single attitude dimension . . . 57

5.3.3 In a world with two attitude dimensions . . . 59

II Simulation **63**

6 Creating a Synthetic Population **64**

6.1 Introduction . . . 64

6.2 Using a political typology . . . 65

6.3 Training synthetic population . . . 66

6.4 Obtaining more complete data . . . 68

6.4.1 Broadcasting . . . 69

6.4.2 Information seeking: deliberate exposure . . . 72

7 Proof-of-Concept Implementation **74**

7.1 Introduction . . . 74

7.2 Simulation description . . . 74

7.2.1 Mixing the collective . . . 75

7.2.2 The broadcast . . . 76

7.2.3 A second mixing results in . . . 78

7.3 Discrepancy between reality and simulation 78

8 Future Studies and Conclusion **84**

8.1 American Kulturkampf 84

8.2 Conclusion 87

III Appendices **88**

A Code snippets of implementation **89**

A.1 Knowledge retrieval query 89

A.2 What to broadcast 90

A.3 Data file for proof-of-concept simulation 91

A.4 Loader for simulation data 96

B Broadcasts **108**

B.1 The Washington Post, April 7, 2017. Page A1 108

B.2 Guidelines for Conservative Trolls 110

Glossary **112**

Bibliography **115**

1 Introduction

This thesis introduces a new way of modelling information for the purposes of agent-based political simulations. At the turn of the century, the academic enthusiasm for then-new computational technologies was boundless. The idea of simulating a society dates since at least Schelling (1971) who had shown that a society would become segregated on racial lines even if the constituents were not individually racist. When computers became affordable and more powerful, ambitious attempts were done to make more sophisticated studies, like Epstein and Axtell (1996) and Davidsson (2002). Later, more targeted studies like Jager and Amblard (2004) became concerned with the diffusion of information in simulated societies and the emergence of polarized opinions regarding a single object. The question of trust came up rapidly in these artificial societies (eg: Schillo et al. (2000) and Conte and Paolucci (2002)). As Squazzoni et al. (2014) reviews, the field grew, but more in the direction of business applications, particularly marketing. As the field grew, studies such as Van Eck et al. (2011) emerged where the primary motivation was the application to business cases and consumer adoption of beliefs. While this is indubitably important, the field seems to have mostly neglected the foundational motivation to simulate a society for social science purposes. My work in this thesis casts a foundation to reach this aim.

The paradigm I introduce allows for sophisticated studies of how attitudes diffuse in a collective – a group of agents. As such, it is distinct and complementary from the efforts in simulated political psychology of Kim et al. (2010). It opens up new pathways for a "third way of doing science" (Axelrod, 1997) in political communications. The desire to estimate the effect of social media on presidential approval motivated this work. The initial experimental design was to compare the strike on Iraq ordered by President Clinton in June 1993 with the strike ordered by President Trump in April 2017. These two strikes are remarkably symmetrical, employing the same weapons systems, launched from similar

platforms, on nations in the same region of the world. Both strikes were ordered at the beginning of each presidency and constitute the President's first independently ordered military action. In both strikes, the objective was to send a message to a dictator, reminding him of American hard power and the limits of what it may tolerate. In the case of President Clinton, the strike was meant to destroy the intelligence headquarters in Baghdad where a plot to assassinate former president George Bush while on visit to Kuwait was hatched. The strike used twenty-three Tomahawk missiles launched from two ships striking an empty government building early Saturday morning. Nine civilians died. President Trump ordered his fifty-nine Tomahawk missile strike from two ships in response to a chemical weapons attack on Syrian civilians by the Syrian government. The aim was to destroy parts of the Shayrat airbase but not the suspected chemical weapons stock piles. Some materiel was damaged and nine civilians died along with seven to nine soldiers. These events were both single strikes without any follow-up. They were meant to redress an affront to American values and assert her power in a minor international crisis. Given such strong similarity between the two, why did President Clinton get a seven point increase of in approval ratings while President Trump got less than one? This short-term surge in approval ratings is termed a rally (Brody, 1991) and there are different hypotheses on why people rally, from patriotism (eg: Baker and Oneal, 2001) to the criticism, or absence thereof, by the political elite (eg: Zaller and Chiu, 1996), to predisposition to favourable views (eg: Edwards III and Swenson, 1997, who studied the Clinton strike), to, more generally, the interactions between the public, the media, and the political elite (eg: Groeling and Baum, 2008). It is to this latter view that I ascribe and for which I needed to develop a general purpose simulator of information diffusion.

In terms of information, a rally event may be seen as an impulse – a shock to the system. This is because rally events affect the entire nation by directly involving the president, are highly salient, and they gain widespread attention and interest. The missile strikes could be seen as producing shock and awe, both for the targeted foe and to members of the (worldwide) public. Rallies are interesting on their own, since they bring to the fore the question of whether a president could, at least for a short time, distract the domestic public's attention from domestic issues and focus them internationally. In the current

media landscape, would President Clinton have gotten the same surge, and from the same constituents (Edwards III and Swenson, 1997) as in 1993? Would President Trump have gotten a larger surge if thousands of ordinary people would not have been able to broadcast their opinions to millions in the minutes that followed the breaking news? These sorts of questions are not only interesting from a hypothetical and theoretical perspective: they allow one to gain insight on how these new communication mediums affect democratic power and public opinion. In order to explore these questions, one must study how information diffuses across a social network and how the public, the media, and the political elites interact on the basis of this information to form, change, or reinforce attitudes towards political objects. A number of researchers have studied how information spreads on real and experimental social networks according to their topologies. For example, Bakshy et al. (2012) research how information diffuses with millions of unwitting participants and static social networks to determine how users share information and the influence of so-called strong and weak ties but stopped short of manipulating the social network or considering the diffusion of *attitudes* towards the piece of information being tracked. On the other hand, Centola's (2010) experiment manipulates the network topology to study the so-called complex contagion of a single behaviour, and this comes closer to the questions that drove this thesis, but considers only diffusion among a single type of actor – there are no elites, nor media outlets with greater credibility and broadcasting reach. Centola (2018) would refer to information of missile strike spreading as a *simple* contagion; however, attitudes arguably diffuse as *complex* contagions. While controlled experiments like Centola (2010) and Centola (2011) may establish valuable results on diffusion dynamics of behaviours that may be translated to larger-scale social programmes, a simulation approach enables the researcher to study questions which are very difficult or impossible to reproduce in a laboratory setting. For example, one would find it difficult to study how attitudes of human subjects would evolve following a sharply dramatic event like a missile strike if they and their friends have access to Twitter or not. Nor can one practically study the effect of some variable in long-term evolution of attitudes – what if, all else being equal, the September 11 terrorists were not Muslim? Simulation, despite its stylized nature and modelling challenges, is the best tool to explore a range of "what if" questions. Since there did not seem to exist

a simulator that could model attitude and information diffusion, I set out to create one. This turned out to be more difficult and complex than expected. I never got to answer my questions on the effects of social network topology on presidential approval. Instead, I developed a model to represent information, the foundation of such a simulator, so that agents could compute attitudes.

The text that follows is both an outline and a guide to build a general-purpose political communications simulator. It is divided into two parts. The first deals with the information model. Chapter 2 describes a structured language that is used throughout to encapsulate information as well as the mechanisms for transmitting information. Chapter 3 takes information, assembles it into knowledge, and uses it to express the characteristics of agents. The first part of the chapter elaborates on how knowledge is created, how memory is simulated and its limitations, and how an agent forms beliefs. The second part of the chapter demonstrates the highly influential contribution certain characteristics have on the behaviour and cognition of agents. It argues for their inclusion in a sophisticated simulation while highlighting the expressive limits of the structured language from Chapter 2. Chapter 4 is concerned with how information is added to agent knowledge, particularly how an agent computes the credibility of the information to which it is exposed. Although the simulation of agent cognition I wrote for the proof-of-concept simulation of Chapter 7 does not take into account equivalency framing effects – in the sense of Tversky and Kahneman (1981) – I ensured that the language would allow for this in future developments. Chapter 5 establishes the mathematical foundations of the structured language. It describes how to compute attitudes from structured statements and how an agent decides whether to accept, reject, or ignore information to which its been exposed. It also proves a number of properties and most importantly, demonstrates that a number of empirical observations on political behaviour(Baum, 2002) emerge naturally from the model's construction, rather than from special-purpose instructions provided to agents. The second part of the text deals with the simulation aspect. Chapter 6, elaborates on the practicalities of generating a synthetic population and how to do it. Chapter 7 integrates a number of concepts from the text into a proof-of-concept simulation which demonstrates the operation of the paradigm as well as the inference of attitudes from structured statements shared among agents. Finally,

Chapter 8 provides motivation for a future study and concludes. A data file as well as a loader are provided in appendices, along with snippets of source code which are useful for to understand some precise mechanisms of agent memory and behaviour. The full source code – with its numerous inelegant programming choices – written in a combination of MySQL, Python, C++, and Matlab can be made available upon request.

I Information Models

2 Language Model

2.1 Introduction

In the study of language, a sentence encapsulates a complete unit of thought. Similarly, in this model, statements form complete declarative thoughts which state a fact or an opinion. Despite emulating the canonical Subject-Verb-Object word order (SVO) word order in main clauses of English for simplicity, I recognize that many other languages belong to other typologies, such as Verb-Subject-Object word order (VSO), eg: Arabic, and Subject-Object-Verb word order (SOV), eg: Japanese. As such, while the text is written employing the subject-verb-object structure, the paradigm is not limited to it and may made amenable to any basic word order with the caveat of *consistency* in the word order chosen. In other words, inverted word orders, such as using VSO in French, a SVO language, to denote a question is doubly forbidden as a *statement*.

This chapter will first go through the statement structure in Section 2.2. Other types of sentences and their uses is discussed in Section 2.3. Section 2.4 establishes how information is bundled and transmitted in the model while Section 2.5 looks into how real broadcasts may be integrated into it. Section 2.2.3 elaborates on how modifiers may be included, such as negation[1] prepositions of time and place, as well as quantifiers. Although these are important elements, it should be noted that they substantially complicate the interpretation and inference of meaning. This is made evident in Chapter 5.

[1] Conjunction, the logical AND as well as mutual exclusiveness, the logical XOR, present uniquely difficult problems in the simulation of reasoning and are omitted from discussion, notwithstanding their necessity to sustain formal logic, for simplicity.

2.2 Subject-Verb-Object statements

2.2.1 Notation

The components of a statement are delimited by XML tags, which is how any word order may easily be used. The subject of a statement is delimited by <s></s> tags, the object by <o></o> and the verb with <v></v>. so that a statement like "Trump is bad" is represented by

<s>Trump </s><v>is </v><o>bad</o>

Appendix A.4 provides a Python program to parse this XML data.

2.2.2 Keywords

All meaning in the context of this model is rooted in a set of key words which are akin to elemental particles providing fundamental quantities and acting as the building blocks of a physical model. Each keyword has an associated, opposite keyword for convenience – "good" is opposed by "bad" and both are used as objects in a statement to provide a positive or negative emotional charge to the subject [2]. By using other keywords denoting orthogonal areas of interest, such as "urgent" and "important", an attitude (vector space) is constructed[3]. (See Section 5.2.1) It should be noted that keywords are entities, if a special kind. Keywords are used in the agent logic implementation as a stopping criterion. More specifically, in order to infer the attitude towards the subject of a statement, a recursive procedure will accumulate attitudes towards the object by evaluating statements where it is the subject. The recursive procedure stops when it reaches a keyword. A second particularity of keywords which distinguish them from other entities is that they define the dimensions of the attitude space. Agent knowledge, which is where all knowledge is stored, including self-knowledge on attitudes towards entities, is constructed to reflect

[2]Because of the intensity parameter of Section 2.2.3 and the usual construct of these sorts of statements, <statement intensity="-1"><s>x </s><v>is </v><o>bad</o></statement> is attitudinally equivalent to <statement intensity="1"><s>x </s><v>is </v><o>good</o></statement>. However, this is contrary to the findings of Tversky and Kahneman (1981) and Section 4.3 discusses the question of framing in more detail.

[3]The opposite keywords, eg: "unimportant", were omitted to lighten the text. Strictly speaking, only one of each keyword pair is in the attitude space basis.

the dimensions of the attitude vector space and, concretely, the database table of agent knowledge will have a column for each keyword pair.

2.2.3 Verbs to inform meaning

Verbs are a unique component of the model since unlike nearly all other components, they are not entities. Verbs are what link objects to subjects in statements. Unlike entities which may be associated to an attitude vector, verbs are *operators*, mapping the attitude vector of an object to the subject of a statement.[4] The attitude map of each verb must be individually defined (see Section 5.2.2) and programmed and, as a result, the vocabulary (ie: set) of verbs is quite limited. Because of the rigid SVO statement structure, there is no conjugation which may complicate the mathematical model and the programming implementation. Instead, when written, verbs are often used in a third-person singular form. This is a personal choice since verbs are represented by an integer in a specific computer memory location and the representative integer is defined in the simulation data file. You could easily define the human-readable representation for your own verbs. (See Appendix A.3.) One may also define verbs without defining an associated attitude map for special purposes, such as indexing knowledge. For example, the verb "know" may be used to link two entities, representing the knowledge of the subject for a particular object, eg: `<s>McCain </s><v>know </v><o>Military operations</o>`. Statements with the verb "know" are used prominently when deriving credibility from expertise (See Section 4.2.2) but also when computing trust from homophily (Section 4.2.1). Similarly, statements with the verb "agree" and "like" are used to used for credibility computations based on homophily and on costly rhetoric (Section 4.2.3.)

Intensity to nuance meaning. While verbs are not conjugated, there is still the need to inflect verbs to nuance meaning. The most important inflection in the context of this model is to denote the "strength" or intensity of the verb. This is encapsulated in a Subject-Verb-Object with an intensity parameter (SVOi) statement. For example, a person (subject) may *agree* (verb) with a statement (object) enthusiastically or tepidly (intensity). The intensity

[4] It is because of this design than conjunction (logical AND) and exclusive conjunction (logical XOR) were particularly difficult to implement.

of a statement is represented by a real number, a scalar. In the XML form, this distinction is shown in Statement Set 1.

```
<statement intensity="0.01">
<s>P </s><v>agree </v><o>stmt</o>
</statement>

<statement intensity="0.99">
<s>P </s><v>agree </v><o>stmt</o>
</statement>
```

Statement Set 1: An example of tepid agreement and strong agreement, respectively

This sort of parametrization could also be used to provide chronological inflections. However, a better sense could be obtained using modifiers, as described in the following subsection, but at the cost of complexity as illustrated in Section 5.2.2. This is arguably an incomplete model of language.

Modifiers: Adding a sense of time and scale. With the use of XML tags, modifiers to statements may increase the expressive range in the simulation. The difficulty does not arise from the notation but rather from the computation of meaning in a statement which includes modifiers. Negation is not only important from a logical standpoint but also from a political one. Negation allows agents to compute attitudinal changes based on pre-existing policy preferences (their agenda). Statements like "Trump not bomb Syria" have political implications which may not be captured with simple parametrization of the statement like in the above subsection. In order to express preferences and expand the model to analyze the effects of exposure to agenda-setting communications, modal verb constructions that express obligation must be added, eg: "Trump should bomb Syria". Modals for possibilities (eg: "Trump could bomb Syria") are also important for agenda-setting communications. Prepositions of time (eg: "Trump yesterday bomb Syria") add substantial richness, acting on memory recall and being useful to for communications that

prime their targets. The inclusion of superlatives (eg: "Trump is best president") may be easier than comparatives (eg: "Trump is better president than Obama"). Both seem very important for a richer modelling of agent interactions since even judiciously parametrized SVOi statements would not link or rank entities. Although these are important elements, it should be noted that they greatly complicate the interpretation and inference of meaning and, for now, implementing this, even from a theoretical perspective, exceeds my abilities.

2.3 Other types of sentences

This model allows only for declarative sentences with a fixed word order. Nevertheless, due to the design and implementation of the model, language features, including conditional, imperative, interrogative, and exclamatory sentences could be added. With careful parametrization, not unlike the intensity parameter of Section 2.2.3, the sentence type could be denoted. Whether these sentences are accepted as true by an agent requires the same credibility evaluations as for declarative sentences. Pointing to the modularity of the design, the procedures of Chapter 4 would only need to be adapted to operate on these additional sentence types.

2.3.1 Conditional sentences

Conditional sentences are particularly important for a more sophisticated study of media effects for two reasons. Firstly, conditional sentences are a foundational element to endow agents with propositional logic. While the present model allows for some inferred reasoning, it cannot not allow for deduction because the *if-then* structure of conditional sentences is missing. Secondly, conditional sentences may also be employed to transmit behaviour rules such that agents respond to stimuli. This is very important to conditioning stimulus-response behaviour in agents which may then drive group identification behaviours. As an extreme example, consider the infamous Nazi salute. Fundamental to this is the transmission of new behaviour rules. The infamous salute and exclamation may be seen as the stimulus-response result of conditioning which occurred in person, through mass gatherings

and in normal interactions, as well as through media, for example in *Triumph of the Will*[5]. The salute is an emission of information which identifies members of a group and provides a strong belief of homophily or heterophily, which is an important component of credibility evaluations. (See Section 4.2.1)

2.3.2 Imperative sentences

Imperative sentences are useful in the broader context of studying media effects since this sort of sentence is often used to direct the attention of the recipient of a broadcast (eg: "stay tuned!") and influence the information diet. Moreover, this sort of sentence could be used to transmit behaviour cues (eg: "Drink Coca-Cola").

2.3.3 Interrogative sentences

Interrogative sentences may be employed to prompt memory recall of a subject through mass-media in order to reinforce existing positions. Both the headline of WorldNetDaily (2012) and the images of billboards it contains ask "Where's the real birth certificate?" in reference to former President Obama's documents. This is an example of the importance of implementing interrogative sentences in future development to better capture the nuanced effects of the media on attitudes. Interrogative sentences used in this way may also prompt information-seeking behaviour, including choosing to read an article headlined by a question. The second, more frequent use is in agent interactions as a mechanism to request information. (Consider the questions asked when meeting someone new: "what's your name?", "what do you do?", "where are you from?", etc.) This is important for a richer simulation of dynamic social networks where the bonds between people evolve over time.

2.3.4 Exclamatory sentences

Exclamatory sentences may be communicated with appropriate declarative sentences and keywords, namely "is important", "is urgent", and "is surprising". Addition of this class

[5] *Triumph of the Will*. Directed by Leni Riefenstahl. Performed by Adolf Hitler et al . Germany: Universum Film (UFA), 1935. Film.

of sentences seems to be of dubious utility since judicious use of keywords and statements renders an equivalent result.

2.4 Bundling Information for exchange

2.4.1 Broadcasts

A broadcast is the most encompassing bundle of information. It joins messages and/or other broadcasts. For example, a physical copy of The Washington Post may be seen as one broadcast. The newspaper contains many pages, each of them a separate broadcast contained in the main broadcast – the physical newspaper. Appendix B.1 reproduces page A1 of the April 7, 2017 newspaper. As it may be seen, there are many articles on this page, each of them a broadcast. Each article contains both messages and other broadcasts. The headline of an article is a (influential) message. Headlines are important for journalistic reasons because it influences whether readers choose to expose themselves to the broadcast it headlines. This is retaken in Section 2.4.3. Inside each article, such as the one in Figure 2.1, there are messages (eg:"The U.S. military launched 59 cruise missiles at a Syrian military airfield.") and there are still more broadcasts (eg: "President Trump said the strike was in the 'vital national security interest.'") A broadcast is an entity. A broadcast is conveyed through a channel and is intended to reach at least one target. It has a time of creation and, optionally, a time of destruction or expiration. The latter is useful for simulating broadcasts that do not linger perpetually, like online newspaper articles, but rather have a limited window of time during which agents may be exposed to it, like television broadcasts. A broadcast may be shared and referenced by agents – that is, agents can formulate statements that have a broadcast as a subject or object. An example of such a statement is <s>B </s><v>is </v><o>lie</o> where "lie" is a keyword in the dimension of truth and B refers to a broadcast which the agent believes is not true. Additionally, messages may be appended to an existing broadcast. This is meant to simulate both 'developing' broadcasts, like television or radio broadcasts, as well as online commenting: it helps model Twitter threads and comments on blogs. When a user x "likes" a tweet B, it is appending to the



Firing of cruise missiles comes days after chemical attack against civilians

BY DAN LAMOTHE, MISSY RYAN AND THOMAS GIBBONS-NEFF

The U.S. military launched 59 cruise missiles at a Syrian military airfield late Thursday, in the first direct American assault on the government of President Bashar al-Assad since that country's civil war began six years ago.

The operation, which the Trump administration authorized in retaliation for a chemical attack killing scores of civilians this week, dramatically expands U.S. military involvement in Syria and exposes the United States to heightened risk of direct confrontation with Russia and Iran, both backing Assad in his attempt to crush his opposition.

President Trump said the strike was in the "vital national security interest" of the United States and called on "all civilized nations to join us in seeking to end the slaughter and bloodshed in Syria. And also to end terrorism of all kinds and all types."

"We ask for God's wisdom as we

SYRIA CONTINUED ON A13

Figure 2.1: Article at the top of The Washington Post's front page, April 7, 2017

broadcast B the broadcast consisting of a single message with the statement <s>user x </s><v>like </v><o>B</o>.[6]

2.4.2 Messages

A message is composed of a statement – representing the content – along with information on the source, medium, time of creation, and optionally, a time of expiration. The medium of a message is its format, such as sound, image, writing, video, etc. It is distinct from the channel that carries the broadcast which contains the message and a broadcast may contain messages in different mediums. The front page of a newspaper, for example the one in Appendix B.1, is mostly composed of messages using the written medium but there are a noticeable few that are in the image medium. The medium of a statement affects the propensity a target will expose itself to it. While further research needs to be conducted to

[6]The message source would be user x and the channel the same as the channel of the broadcast.

precisely determine the different propensities of voluntary exposure by message medium and population type, a casual observation is sufficient to suspect the existence of a differential propensity of voluntary exposure between mediums for all population types.

2.4.3 Queues

Because agents have an overabundance of information on all sorts of subjects, they cannot process it all. Broadcasts that target an agent are placed in that agent's queue, to be scanned at its convenience and sorted by its beliefs, such as the degree of homophily and reputation of the channel or source, by its preferences on message medium, by information demands, and by "momentum" – the last channel or source, so that an agent is more likely to stay "tuned-in" after the first voluntary exposure. Medium preferences and information demands are dynamic sorting criteria. When an agent wants to know more on a particular subject, the sort criteria changes to reflect this demand for information.

When an agent scans its sorted queue, it automatically exposes itself to the headlines and processes them as any other statement, using the *channel* as the *source*. This stipulation ensures that the headline of Figure 2.1 is attributed to The Washington Post rather than the article authors. Broadcasts that have no titles or headlines, such as the one in Figure 2.2, use the statement of the first message it contains as the headline. The implication of these two stipulations is that when an agent scans through a large list of broadcasts received from social media – analogous to a person "scrolling" through posts – each influential headline scanned is attributed to the *channel* and *not* the actual source. Not only does this mimic anecdotal evidence of how people interact with headlines and social media feeds, it also mimics the increased credibility of headlines because they are attributed to a reputable entity. This is key to modelling the persuasiveness and astounding credibility of messages from unknown, possibly malignant, sources. However, once an agent "tunes-in" to a broadcast by choosing to expose itself to it, then the source is correctly attributed to the message. As such, this model is capable of simulating persuasive anonymous Russian election trolls on Facebook as capably as Walter Cronkite's untarnished credibility.

2.5 Using real broadcasts

In normal, natural language there are sometimes many declarative sentences bundled as a single message. A broadcast from President Trump in Figure 2.2 which appeared on Twitter on February 24, 2017 provides a useful referent for the following exposition.



Figure 2.2: A broadcast by US President Donald Trump on Twitter implying that The New York Times and CNN are not credible news sources, February 24, 2017.

2.5.1 Limitations of the model

The broadcast in Figure 2.2 can be converted into a series of statements with some human interpretation. Crucially, the presidential tweet does not explicitly call The New York Times or CNN liars. Nevertheless, in an attitude vector space with the dimension of truth (defined by the keyword pair "truth" and "lie") and the verb vocabulary $\{is, threaten\}$,[7] the tweet could be translated to the Statement Set 2.

The relative difficulty in creating Statement Set 2 from Figure 2.2 is due to the necessary inference that must be performed given the president's persistent difficulty in clearly

[7]The verb "threaten" would be defined similarly to the verb "bomb" of Section 5.2.2. The verb "is" is defined in that same section.

```
<s>New York Times </s><v>is </v><o>fake news</o>
<s>CNN </s><v>is </v><o>fake news</o>
<s>fake news </s><v>is </v><o>lie</o>
<s>fake news </s><v>threaten </v><o>USA</o>
```

Statement Set 2: Statements corresponding the the presidential tweet in Figure 2.2

articulating his message. While it is (usually) possible to parse his words into a coherent, intelligible message, Figure 2.2 nevertheless highlights a major limitation of the model: When considering these sorts of influential messages, the subtleties of the uncomfortable or rallying echoes of a head of state labelling critical domestic media outlets as *Lügenpresse*[8] cannot be fully captured. Furthermore, the declaration that The New York Times and CNN should not be taken seriously is imperfectly captured by the statements of Statement Set 2. Similarly, the claim of *deliberate* lies, and it's implied malignancy, is imperfectly captured and hinges on a positive opinion of "USA". The sense of a bygone era, when The New York Times and CNN *could* be trusted, as expressed through the use of the present perfect "has become [a joke]" along with the exclamation "Sad!", fails to be captured; yet the implicit claim that what *can* be trusted is @realDonaldTrump is part of the model when agents are programmed to be more credulous of sources with whose messages they agree.

2.5.2 Vast data gathering operation

In order to recreate the media environment of a particular period, actual broadcasts must be gathered for three reasons. The first is that the simulation is a self-contained informational environment and while *commentary* can be generated with appropriate programming and calibration to real broadcasts, *news*, especially news of foreign military action, cannot be endogenously generated. The second reason is for calibration purposes. The methodologies of Groeling and Baum (2008) and Baum and Groeling (2008) provide a guiding example of how to calibrate media coverage for major networks and partisan media, respectively. Studies like that of Vosoughi et al. (2018) provide for calibration of agent sharing. Similarly useful are the results from de Benedictis-Kessner et al. (ND), as well as the methodology it employs to reveal media preferences, the influence of the media on attitudes, and the

[8] Nesbit (2016)

propensity to share information. Programming and training agents to reflect these findings requires real broadcasts. Lastly, some agents, like the one representing the president, may be considered so important and influential, that an experimenter may wish to use actual broadcasts, appropriately translated into the simulation's language, in order to drive it.

2.5.3 Automated coding

Coding an article, or translating it to the language of the model, is the exercise that was performed to generate Statement Set 2 from Figure 2.2. Given a fragmented media landscape with numerous news outlets publishing many articles per day and with prolific influential Tweeters, like Mr. Trump, the amount of work of assembling data and translating it to the simulation language would be overwhelming even for a costly army of research assistants (cf. Groeling and Baum (2008)). In order to resolve this impediment, future research would likely require the use of natural language processing, a subset of artificial intelligence, to automate the task. However, given that the data collected on broadcasts by individuals and amateur organizations is not subject to same copy-editing standards of professional organizations, there are likely to be substantial challenges in interpreting the content due to unusual spelling, grammar, or epithets (eg: "towel heads" for Muslims in Figure 3.2a). This difficulty is compounded by the central importance the US President has in any political simulation and the difficulties of the 45^{th} holder of the office to deliver grammatically correct, coherent messages. Making sense of the Presidential tweet in Figure 2.3 requires several cognitive leaps which may prove too difficult for natural language processing, yet it is an important and influential message which had an effect on the American public's perception of immigrants. Likewise, the Presidential tweet of Figure 2.2 is a message which influenced the perception of two media channels and reinforced attitudes on "Fake News", but the heterodox grammar as well as the use of the idiom "has become a joke" would make it very difficult to automate translation this tweet into a structured statement.

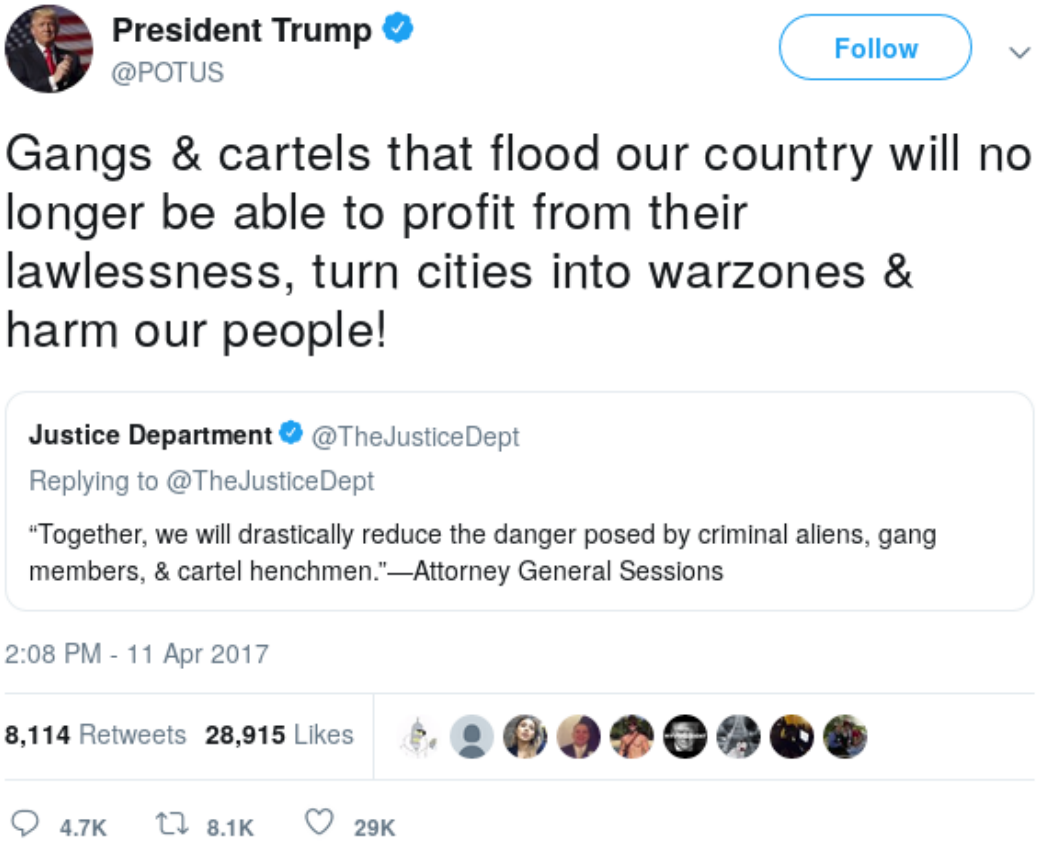


Figure 2.3: An example of a Presidential tweet which poses substantial interpretation challenges, even for humans.

3 Agent Knowledge Model

3.1 Introduction

This chapter builds on the representation of information from Chapter 2 and assembles it into a foundation through which agents can make sense of the world. An agent thus has knowledge of the world and itself. The self-knowledge takes the form of characteristics and it conditions the agent's cognitive processes, its behaviour, and its exposure to information. The major characteristics considered are its physical attributes, geographical location, religion, race, education, and core values. While some of these are mutable – for example an agent can move, age, or gain education – these are distinct from attitudes which are much more responsive to information exposure. Bafumi and Shapiro (2009) shows that core values change over time and most likely due to exposure to information. However, the time scale at which core values (eg: views on abortion) change, barring an extreme adoption of new beliefs akin to a conversion, is measured in years – or perhaps, for a particular person, never. I contend that attitudes, on the other hand, consist of those opinions (of what is good, important, etc) which change in much shorter time scales. They are relatively volatile and are based on the exposure to information which is interpreted using individual characteristics and previous information. The degree with which an agent will process information in a systematic manner rather than relying on heuristics depends on its capacity and willingness to process a received message (Perloff, 2013, p. 265). When processing information with heuristics, an agent *still* relies on its knowledge – particularly what it knows about the source and channel. It uses this knowledge to determine the credibility (Section 4.2) of the message and determine whether to accept it, reject it, or ignore the message (Section 5.2.4).

Despite my early efforts being centred on modelling belief structures with spreading

activation, I was unable to do so while retaining the design for arbitrary statements and pre-defined attitude dimensions. The impressive efforts of Kim et al. (2010) have created a sophisticated model of memory and belief structures. My model shares several similarities to Kim et al.'s (2010) but they differ in their focus and in technical details. Kim et al.'s (2010) *John Q. Public* (JQP) model is concerned with the individual and is at its core a simulation of political psychology; my model being chiefly concerned with how information diffuses in a collective of agents is at its core a simulation of political communications. Thiriot et al. (2008) present a model which also relies on associative networks to represent knowledge and they employ this to study the diffusion of innovation in a collective. Their model is much lighter computationally than Kim et al.'s (2010) and they were able to simulate a collective with thousands of agents. However, unlike both Kim et al. (2010) and my models, Thiriot et al.'s (2008) relies on a *finite* set of concepts – which I've termed entities – and this proves to be one of its limitations for a political simulation. In a general purpose political communications simulation, agents must be able to form opinions on the information to which they've been exposed and, in so doing, generate new information: their opinion on a concept is, in itself, information and consequently a concept. Therefore, the number of entities must be boundless.[1] A second limitation of associative networks which is apparent in *both* Kim et al. (2010) and Thiriot et al. (2008), but principally in the latter's, is the difficulty in computing attitudes towards entities. The motivated reasoning underlying Kim et al.'s (2010) JQP model rescues it from being reduced to an emotionless automaton.[2] Thiriot et al.'s (2008) model, on the other hand, only *seems* amenable to emotional associations but they do not elaborate on the "functions to evaluate attractivity" (p. 267). Although my model struggles with cascading memory activation of associative models, it is capable of representing both the conceptual associations *and* computing on them through verbs to form attitudes.[3] The knowledge in the associative models may be seen as a collection of "is"

[1] Consider a statement s with which an agent x disagrees. In disagreeing, agent x generates a *new* statement <s>s </s><v>is </v><o>wrong</o> which it then may broadcast to other agents. A second agent, agent y may agree with x's disagreement and in doing so, generate a third statement <s><s>s </s><v>is </v><o>wrong</o> </s><v>is </v><o>right</o>, which it then may broadcast, and so on...

[2] In fact, Kim et al. (2010) place a substantial emphasis on augmenting the ACT-R model, on which their JQP model is based, precisely to incorporate "affective mechanisms." (p. 2)

[3] See Section 5.2.2 for a mathematical description of verbs as operators on attitudes attached to simulation entities.

statements, eg: <s>Bush </s><v>is </v><o>honest</o>(Kim et al., 2010, p. 12) or <s>iPod design </s><v>is </v><o>beauty</o> (Thiriot et al., 2008, p. 263). Although both these models are excellent examples of knowledge and memory representation, they fall short in being useful in employing knowledge to condition agent responsiveness to subsequent messages and to form opinions on other agents. That is, although JQP may hold a belief that <s>Bush </s><v>is </v><o>honest</o>, it would not be more receptive, more prone to accepting, the statements of messages with "Bush" as a source.

JQP operates on simplified natural language or "simple campaign statements attributable to some known actor," (p. 15) and is superior to mine in the sophistication of simulated cognition and memory. The similarities extend from the simplification of language to the integration of knowledge – compare Equations 3a and 3b in Kim et al. (2010) to Equation 5.6 in my model. However, the simulations of Kim et al. (2010) do not involve a collective sharing information: instead, JQP is simply exposed to (natural language) information and its attitudinal changes are observed. In other words, the JQP simulations are analogous to forced exposure experiments where the experimental subject is an isolated individual who answers survey questions.

My model also codes natural language news into statements but these are *structured* statements, thereby further simplifying the language. The simpler language allows for greater computation: while JQP can answer how it feels about an issue or entity, it cannot infer how it *should* feel about "Trump" when exposed to "Trump bomb Syria" – yet it is this inference which is crucial to simulate attitude dynamics. Instead of a single agent forcibly exposed to information, my model is for a self-contained collective where agents both emit and receive information. Forming opinions based on information and sharing them is the core consideration. In the process of evaluating whether information should be believed, a measure of credibility independent of the message content is constructed and is subsequently used along with message content to determine whether an agent accepts or not the information – JQP only evaluates credibility based on message content.

With JQP, Kim et al. (2010) were able to track cognitive processes to minute details; my model does not allow for that. Instead, it enables to track how a specific piece of information circulates and evolves within a group of agents. JQP does not need to have behaviour, it

only needs to *feel*. On the other hand, the agents in this model need to both feel and act. The rest of this section elaborates what an agent knows about itself and others, the importance and effects of certain characteristics, and the limits of what may be expressed in the language of the model.

3.2 Knowledge and memory

3.2.1 Accumulation

Agents who accept or reject a statement (see Section 5.2.4) will add it to their knowledge, along with a measure of belief in its veracity. All the statements to which an agent has been exposed are kept in this knowledge. Without the addition of forgetfulness, agents have infinite memory and all exposures from the very first to the very last are equally weighed. Of course, this is not representative of human memory. It has long been recognized that a important factor in (political) evaluations is whatever information comes to mind, rather than the sum of knowledge. To model the priming effects (Iyengar, 1987), there *needs* to be forgetfulness so that the concepts to which an agent has been more recently exposed are more salient and thus more heavily weighed in the evaluation of a political entity. Kim et al. (2010) offer a sophisticated model that uses both long-term memory and working memory. My memory model is much simpler, consisting of an exponentially decaying amount of memory for a each statement in an agent's knowledge. However, because of the separation of data and logic inherent in my model, the manner in which memory is modelled may be improved in the programming of agents without necessitating significant changes in how information is modelled or accumulated as knowledge.

3.2.2 Forgetfulness and learning

My memory model is relatively simple and does not distinguish between different memory types. It is a high-level view of memory, in contradistinction to low-level views like those of Kim et al. (2010)[4] There has been substantial research in modelling memory more generally.

[4]It is conceivable that memory could be simulated at the molecular level (Goelet et al., 1986) but this is far more than is necessary or useful for political simulation.

(eg: Hochreiter and Schmidhuber, 1997). In my model of memory, statements that are in long-term memory are simply more memorable because the rate at which memory decays is slower. I model the transition from short-term to long-term memory in the same manner that I model learning: by decreasing the rate at which a statement is forgotten. This is an oversimplification.

Forgetfulness is important to model because it allows for topics to sink into the deeper recesses of an agent's "mind" while allowing others to float up to the forefront over longer periods of time. Today's preoccupations are not yesterday's and the importance of a topic's salience in evaluating the president has been well established in the literature. (eg: Iyengar, 1987). Not only is endowing agents with forgetfulness a method of varying the salience of concepts over time, it also allows agents to "learn" by repeated exposure. How quickly an agent forgets a statement is determined by the memory decay rate λ. To model something that cannot be forgotten, like the characteristics of Section 3.3, a rate of $\lambda = 0$ is used. By progressively decreasing the rate of decay as an agent is repeatedly exposed to a statement, learning can be simulated.

While a crude approximation of reality[5] the learning model is as follows. Suppose an agent has a statement s in its knowledge through exposure at time t_0, that the statement has a decay rate of λ_0, and that this is the agent's most recent (or only) exposure to s. When re-exposed to s at time t_1, the remaining memory of the *a priori* most recent exposure to s is

$$r(t_1) = e^{-(t_1 - t_0)\lambda_0} \tag{3.1}$$

The decay rate for the newest statement is chosen to be

$$\lambda_1 = \frac{1}{1 + r(t_1)}\lambda_0 \tag{3.2}$$

Figure 3.1 illustrates how memory of a statement improves as an agent is re-exposed to

[5]This model allows for an agent to always learn. An agent can be easily "brainwashed" ensuring concepts are essentially never forgotten by simply exposing an agent to a sufficiently large number of uninterrupted repetitions of the same message. As any student who has crammed for an exam and forgotten the material soon after can attest, large, concentrated doses of information are not retained very long in reality. Moreover, this model of memory and learning does not account for information saturation – where people avoid or ignore topics following too much exposure. This should be taken into account for accurate simulations.

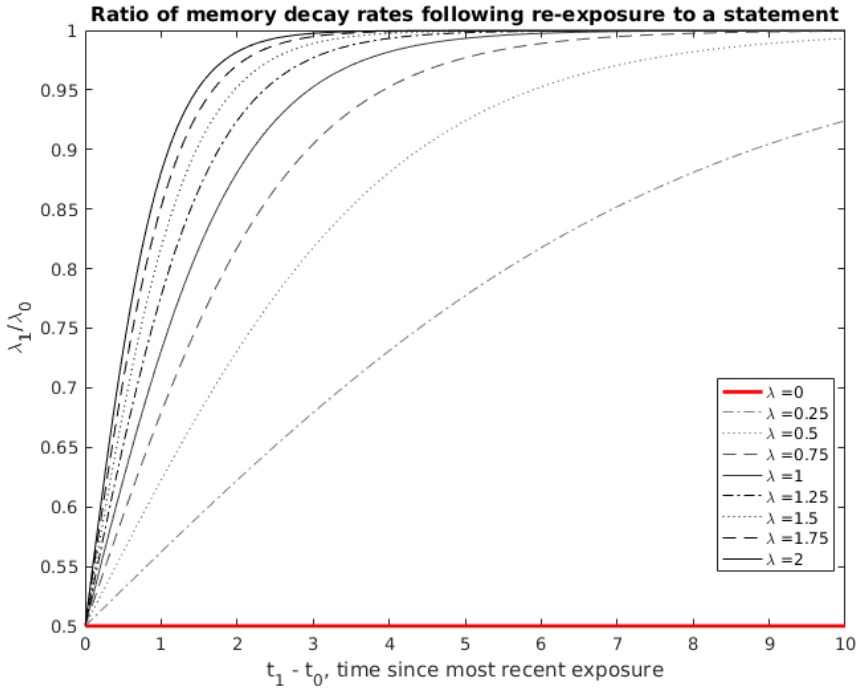


Figure 3.1: Illustration of how forgetfulness decreases after re-exposure to a statement depending on time elapsed since most recent exposure. The time scale is adjusted to fit the the simulation objectives and empirical findings

it. If the time between exposures is too long, then there is essentially no difference in the memory's half-life. On the other hand, if the time between exposures is short, the half-life of a statement's memory increases. The choice of Equation 3.2 is arbitrary and while it exhibits the desirable effect of decreasing forgetfulness following re-exposure, it is not rooted in the literature on memory and a priority of future work should be to solidify this weakness in the conceptualization of learning based on information. Moreover, while agents were programmed to select their most salient issues through a memory sort in the simulation of Chapter 7, there were no restrictions on what was memorable, possibly violating the findings of Miller (1956). It should be noted that this is a weakness of the logic applied to the data – a weakness of the processing of information rather than of the information model *per se*: the data necessary for a more sophisticated model of memory is available and I stress that it is a shortcoming of the logic I programmed to demonstrate the use of the information model.

3.2.3 What an agent believes

The beliefs of an agent are determined by the belief an agent has in the truthfulness of a statement and how memorable it is. For simplicity, the memorable belief in a statement, π in Equation 5.4, was implemented as the product of remaining memory ($r(t)$ from Equation 3.1) which is strictly positive, and the belief associated to a statement, which ranges from 1 when certain of its veracity to -1 when certain of its falsity. Equation 5.4 accumulates the attitudes towards a subject using the statements in an agent's memory, weighing them by their belief. This represents the model of Brousmiche et al. (2017), the observation of (Iyengar, 1987, p. 64), and Zaller's (1992) Axiom 4, "individuals answer survey questions by averaging across considerations that are immediately salient or accessible to them." (p. 49) However, Zaller (1992) would likely dispute how I model immediacy and he convincingly argues that it is unrealistic to "use each piece of incoming information to update all of the 'attitudes' to which it might be relevant" (p. 50). I attempt to preemptively mitigate this criticism by setting a minimal threshold of remaining memory for statements that are retrieved for attitude computations – that is, only statements that are "sufficiently" memorable are used to construct the set $\mathfrak{I}$ of Equation 5.4. These are the `agent_knowledge.mem_remaining` lines in the query listed in Appendix A.1. Increasing this threshold reduces the set of accessible statements to those that are most memorable; this could be augmented by adding criteria using `agent_knowledge.timestamp` but it is still not exactly the same as the most immediate and accessible. By choosing the threshold in accordance to the expected memory decay rate, the choice of this threshold could be improved. This is admittedly not nearly as satisfactory as Kim et al.'s (2010) memory model. The reason is structural. Rather than isolated concepts like "Republican" or "Small-Government" in Kim et al. (2010), in my model agent knowledge is composed of statements – complete units of thought eg: `<s>Small-Government </s><v>is </v><o>good</o>`, are the smallest, atomic, units of knowledge.

3.2.4 Spontaneous recall

Spontaneous recall of information is possible to implement using the data in the information model and requires very simple programming to simulate it. Everyone will have experienced the sudden, (apparently) unprompted remembrance of a piece of information. In the psychology literature this is often referred to as involuntary memory (Mace, 2007). Although this subset of memory functions surely plays a role in political psychology, it was not implemented due to its relative unimportance in the proof-of-concept simulation performed in Chapter 7. By adding a process that randomly selects a "forgotten" statement – that is, one with low levels of remaining memory – and restores its initial memory quantity, spontaneous, involuntary recall is possible.

3.2.5 Continuous attitude changes

Fink and Cai (2013) reports several studies which examine the dynamic aspect of belief change. When an agent accepts (or rejects) a statement according to Equation 5.5, the incorporation into its knowledge is immediate. As such attitude changes in my model are *discontinuous* and beliefs over time, when memory decay is nonexistent or negligible, are a step-function. Continuous attitude changes may influence the information dynamics in a collective – that is, the diffusion and generation of information within it – because the findings reported in Fink and Cai (2013) point towards a belief change taking several seconds to occur. In simulations on the very-short term dynamics of attitude changes in a collective, like the one in Chapter 7, the resolution of time is very fine – but not extremely fine – grained: the smallest unit of simulated time is a second and the duration between agents actions is in the range of several simulated seconds. As such, the timescale of the simulation in Chapter 7 places an ambiguous need for dynamic attitude changes following an exposure. At smaller timescales like those required in Kim et al. (2010) to analyze the cascade of memory activations, dynamic belief changes become much more important. At larger timescales, where the smallest unit of time is in excess of five simulated minutes, the findings of Fink and Cai (2013) would lead one to assume that beliefs *ex post* exposure reach their steady-state within a single time unit and thus that static models of belief change

are sufficient. Due to time constraints, I programmed agents to incorporate information immediately. Nevertheless, the logic operating on the information model may be without great difficulty adapted to incorporate information progressively.

3.3 Characteristics

In contradistinction to knowledge, characteristics are not fluid for the purposes of this simulation. Nevertheless, they are still described through statements using the "is" verb. The characteristics of an agent refers to immutable descriptors, like age and sex, but also less-mutable characteristics, like geographic location or even car ownership. A female agent x would therefore be described through <s>x </s><v>is </v><o>female</o>; a black agent x by <s>x </s><v>is </v><o>black</o>; a college-educated agent by <s>x </s><v>is </v><o>education.college</o>. In more sophisticated agent programming than in my proof-of-concept, some of these characteristics are automatically broadcast in certain mediums and fuzzy beliefs on the broadcasting agent are formed or reinforced. These are crucial for both the determination of homophilic bonds (Section 4.2.1) and for an agent to form expectations of another agent's attitudes (Section 4.2.3). Characteristics may also condition an agent's exposure to information: for example, an agent who commutes by car is much more likely to be exposed to radio news rather than print news compared to a public transit commuter who can read without risking life and limb. The following characteristics were deemed to be of particular importance as they seemingly strongly condition agent behaviour. A number of limits to my model are highlighted and may provide inspiration for future research.

3.3.1 Physical characteristics

Race is a physical characteristic like sex and age, but it seems to carry a great deal of influence on communications and is expounded upon in its own section. Age, despite its importance in framing some political discourse, is relatively undeveloped in my model because I did not see a particular distinction in the literature on how this influences on the capacity to rally. Nevertheless, this characteristic is present in the model and may be used

in homophily calculations. I use sex, rather than gender, given the population data of Pew Research Center (2017) – I assume a binary gender construct and the existence of a bijection between sex and gender.[6] Gender matters for the capacity to rally: in early stages of the conflict, "women are less supportive than men of military actions" (Huddy and Cassese, 2011, p. 473). While corroborating the general aversion to the use of force, Eichenberg (2003) also finds that women respond more positively than men when the use of force is motivated by humanitarian reasons. Defining a humanitarian reason in the language of my model proves to be a challenge. While a statement like <s>Assad </s><v>gas </v><o>children</o> may be used to communicate and prompt negative sentiment and importance towards the "Assad" entity[7], it is *suffering* – which is necessary to justify a humanitarian intervention – that my model struggles to capture.

3.3.2 Geographical location

Geography is a crucial factor in conditioning what sort of information an agent will be exposed to. Channels are often geographically-bound (more so in past media eras,) have an emphasis on local news, and often carry local sources. It is often the case that a source is linked to a geographic area. American federal politicians are generally presented with both their party and state, eg: Sen. McCain (R-AZ); at the state-level, politicians are named with their district, eg: Rep. Marjorie C. Decker (D-Cambridge). When introducing an entity this way, a statement of the sort <s>McCain </s><v>is </v><o>Arizona</o> is implied and this conditions credibility calculations. Geography also plays into derogatory statements that mean to imply heterophilic bonds, eg: "East Coast Liberal", "hillbilly"[8]. When these are applied to a source, they compromise its credibility both because of the heterophily and because of the attitudes derived from known statements which describe the geography-laden term, eg: <s>East Coast Liberal </s><v>is </v><o>bad</o>. Furthermore, geography creates a set of agents with whom an agent may interact on a "day-to-day" basis, and this is further conditioned by the employment status of the agent.

[6] Of course, these assumptions may be loosened as necessary by including appropriate entities in addition to male and female.

[7] Using a verb definition similar to the one in Section 5.2.2.

[8] A poor, white, rural resident of the southern Appalachians

That is, an agent with full-time work is likely to interact with other agents in full-time work in the same geographic area in so-called "water-cooler talk" or through professional relationships; an agent in full-time studies is likely to interact with other students; a retired agent is likely to share opinions with other retirees.

3.3.3 Religion

The religious denomination to which an agent ascribes is used to compute a measure of homophily and, compounded with attendance of religious services, also used to determine the exposure to information. Religious leaders sometimes venture into political discourse[9] and often political leaders will attend religious services. When they do, they may have strong political effects on the congregants (Cole, 2019). In a religious, predominantly Christian, nation like the United States (Pew Research Center, 2017), religious affiliations form part of the electoral scrutiny for the presidency – and a source of disinformation for presidential evaluations (Jaffe, 2015). Religion is often intermingled with politics – despite the First Amendment to the US Constitution – and justifications for controversial bills and lawsuits are often rooted in biblical interpretations. Religious organizations are also "quite active" in foreign policy and different denominations advocate for different positions (Barbeau et al., 2011, p. 543). Combining both the domestic politics and foreign policy influences of religion, I hypothesize that it is a major driving force in the current and ongoing American *Kulturkampf* which I would suggest testing in a future study. (See Section 8.1)

Religion also plays a substantial role in the perception of a media channel's credibility, the term "liberal news media" intermixing political labels with implied anti-semitism (Gillis, 2017). In addition to influencing the credibility of a channel, the religion (or assumed religion) of a subject in a statement influences the credibility of the statement and may also, with more refined future modelling of behaviours, produce a conditioned responses as those in Figure 3.2. Furthermore, some (questionable) media outlets play on religion to re-orient the political discourse, as evidenced by the tweet in Figure 3.3. It is interesting to note that my model is capable of capturing the information contained in the World Net

[9] See `https://www.christiancentury.org/article/opinion/do-politics-belong-church` Accessed 2019-06-06

Daily News tweet but struggles to capture the sarcasm of Figure 3.2b ("Let's not forget the real victims of this strike against Syria: Liberals") and the hatred in the exclamation of Figure 3.2a ("f--- those towel heads"). This latter difficulty is compounded by the use of an epithet that would likely trip an automated coding system. (See Section 2.5.3.)



(a) (b)

Figure 3.2: Examples of reactions on Twitter on the basis of religion following the April 7, 2017 strike in Syria



Figure 3.3: World Net Daily News attempting to reorient the discussion following the April 7, 2017 strike in Syria by leveraging religion

3.3.4 Race

Race relations play a substantial role in a nation that considered "other Persons" – overwhelmingly slaves of African descent – three fifths of a person for the purposes of apportioning taxes and seats in the lower house of government in its constitution (U.S. Constitution, Art. 1, Sec. 2); contended that citizenship as defined in the constitution did not extend to blacks, regardless of whether they were free or enslaved (*Dred Scott v. Sanford*, 1857); and in which politicians systematically manoeuvre to disenfranchise black voters (eg: Lousiana Constitution 1898, Art. 197 Sec.5; *Cooper v. Harris*, 2017). Although "the [young] United States government's support of slavery was based on an overpowering [economic] practicality,"(Zinn, 1999, p. 171), it could be argued that American slaveholders fostered racial hatred among lower classes of whites to preserve the social order which benefited them: "The need for slave control led to an ingenious device, paying poor whites... to be overseers of black labor and therefore buffers for black hatred" (Zinn, 1999, p. 177). After the Civil War and the ratification of the 15^{th} amendment (1870) in which the right to vote was extended to all males, regardless of race or previous conditions of servitude, "ex-slaveholders... turned to terror and intimidation...[and] found collaborators among poor and middling southern whites whose social identities rested on their sense of racial superiority over blacks" (Kornblith, 2011, p. 98) to restore the social order's *status quo ante bellum.*

Racist views in America continue to exist and have even entered into presidential evaluations (Samuel, 2016). Pew Research Center (2017) was unable to ask survey respondents the blunt question "Are you a racist?" but approximates it through a number of indirect questions (Q25f, Q50hh, CB82, CB71). Particularly interesting are the wide differences uncovered by Pew Research Center (2017) in views on racial equality. In some clusters of public opinion, up to 81% believe "the US has made changes to give blacks equal rights with whites" while in others, up to 98% believe the US needs "to continue making changes to give blacks equal rights with whites." As such, the immutable characteristic of race is important for sophisticated political simulations because they are likely to strongly condition computations homophilic (dis)trust.

However, a lot of racism is communicated through memetic images and this poses a

substantial challenge in the translation to the structured statements required by this information model. Although the meme in Figure 3.4a is a derivative from the "Sopa de Macaco, Uma Delicia" meme circulating on 4Chan's /int/ board in 2017, it may be also be interpreted as a white supremacist warning against miscegenation yet carrying no text, this dog-whistle political message is truly silent. Other memetic images have text and express obvious racist undertones but are similarly difficult to translate to structured statements, such as Figure 3.4b. The challenge of incorporating the full nuance contained in these sorts of political messages into a simulation using my information model may be insurmountable. Programming agents to behave with varying degrees of racism, from conscious and overt hatred to unconscious and covert race-based distinctions of homophilic trust, should, on the other hand, be relatively simple when agents are further endowed with racial attitudes. Though due to time constraints this was not attempted, it is important for future studies.



(a) A white supremacist meme warning against miscegenation.



(b) A shirt from a Romney 2012 supporter with explicit racist connotations.

Figure 3.4: Examples of racist communications which are difficult to convert to structured statements. Communications of this sort are likely beyond the expressive limits of the model.

3.3.5 Education

Education is important to being able to comprehend (political) messages. *In extremis*, exposure of an illiterate person to newspapers will result in a complete incapacity to receive any information. Baum (2002) used education as a proxy for political awareness, for which Zaller (1992) established the importance.[10] Chaiken and Eagly (1976), who find that difficult-to-understand issues were more persuasive in the written form and easier-to-understand issues were more persuasive in the video form, didn't separate participants by education levels who were all students at the University of Massachusetts. Nevertheless, they find that easier-to-understand messages were – as obvious as it may seem – more often understood and recalled. Notwithstanding the racial assumptions on education in the findings of Dupree and Fiske (2019), the white, self-reported liberals who dumb down their communications with minorities may be attempting to " 'get on their level'" (Dupree and Fiske, 2019, p. 2) and thereby be perceived as more credible on the basis of homophily. Chaiken (1980) finds that a message is evaluated systematically (rationally) in a high-involvement situation and through heuristic cues in a low-involvement situation. Although the forced exposure methodology leaves some doubts on the generality of the results, one can assume that education levels serve as a useful proxy for the cognitive capacity of a person[11] (Van Hooren et al. (2007), Matarazzo and Herman (1984)) to understand an argument that is systematically evaluated. Baum (2003) finds that education is strongly and positively correlated to voluntary exposure during campaigns (p. 188) and that while there is *some* attentiveness gap during foreign crises with more educated individuals being more politically aware, it has been narrowing in the decades leading up to the study (p. 167). Furthermore, (Baum, 2003, p. 218) finds correlations between party identification and education level – with more-educated individuals leaning more towards Republicans – which he uses to predict the magnitude of a rally given the affiliation of the sitting president. While I agree with the greater propensity (and availability) for opponents to rally and I demonstrate that the

[10] It should be noted that Iyengar (1987) didn't find conclusive evidence for education in mitigating the effects of priming or agenda-setting communications, but this is likely due to his forced exposure experimental design.

[11] And a signal of cognitive capacity in Spence (1973)

model naturally exhibits larger approval rating increases in groups that disapprove *ex ante* of the president in Corollary 1, I would point out that the implicit assumption of stationarity in the correlation between party identification and educational attainment is possibly misplaced, given the findings of Bafumi and Shapiro (2009). Corroborating this impression is Pew Research Center's (2017) data where the group with the highest educational attainment is also the one which is most solidly Democratic and Democratic-leaning. The two least-educated groups, "Devout and Diverse" (57% high-school or less) and "Core conservatives" (50% high-school or less) are respectively mostly Democratic or Democratic-leaning (59%) and overwhelmingly Republican or Republican-leaning (97%).

3.3.6 Core values

Agents may differ in their deeply held values and opinions. This influences how they perceive the world and how they make sense of the information to which they are exposed. For example, in Pew Research Center (2017), a question divides the American public into so-called Hawks, who favour peace through military might, and Doves, who favour peace through diplomacy. Similarly, there are core values of internationalism and isolationism. These sorts of core values colour the perception of news. An isolationist Dove is likely to disapprove of news that the president bombed a far-away land, despite harbouring a negative sentiment towards the nation on the receiving end of the strike. Core values are reflected in my model by pre-existing knowledge (eg: believing <s>Foreign intervention </s><v>is </v><o>bad</o> and <s>Foreign intervention </s><v>is </v><o>wrong</o>) as well as by modifying parameters in the functional definitions of verbs, in the manner of Proposition 6.

3.3.7 Partisanship

Partisanship shapes the way the political behaviour of Americans in uniquely powerful ways. As Jacoby (2011) reports, parity affiliation not only has a "strong impact on foreign policy attitudes" (p. 443), it seemingly has a causal influence on all attitudes. Moreover, party affiliation structures attitudes and conditions the conceptual hierarchy in the mind of an

individual. Initial knowledge structures are fundamental in any political simulation because they are the basis on which an agent will interpret the information to which it is exposed. Kim et al. (2010) uses NAES survey data to create five agent archetypes. Although my model is amenable to being used with a collective initially clustered solely on partisan lines, I suggest a more sophisticated clustering and initialization of agent knowledge structures. Although partisanship is weighted heavily in Pew Research Center's (2017) clustering, I would argue along the same lines that it is a necessary but not a *sufficient* factor and that other core values, such as views on immigration (Q25g), foreign policy (Q50cc, Q50ee)[12] are necessary to disaggregate public opinion and to more accurately simulate the "constituent foundations" (Baum, 2002) of attitude change. Notwithstanding my views on how to best initialize agent attitude structures, partisanship is an important factor in determining the credibility of a message and it appears when evaluating credibility from homophily (Section 4.2.1) as well as credibility from costly rhetoric (Section 4.2.3).

[12]For completeness, note that Pew Research Center (2017) additionally uses views on government regulation of business (Q25b), work and success (Q25k), poverty (Q25c), corporate profits (Q25n), environmental regulation (Q50r), homosexuality (Q50u), opportunity based on race and gender (Q50hh, Q51nn), and the fairness of the economic system (Q51ll).

4 Adding to Agent Knowledge

4.1 Introduction

This chapter is about how agents add information to their knowledge. While it describes a method for adding information to an agent's conception of the world, it should be noted that this chapter is not about the information model *per se*, but rather, one of the most basic simulated cognitive operations (logic) on information (data). When an agent is exposed to information, it will either accept it as true, reject it as false, or simply ignore it. In order to decide which of these actions to take, an agent uses a modified discrepancy model (Fink and Cai, 2013). This differs from motivated reasoning, for which Westen et al. (2006) found conclusive evidence of existence, and which explicitly relied upon by Kim et al.'s (2010) JQP model. Because the trichonomous decision relies on a evaluation of the different between the position advocated in a message and the prevailing attitudes, my model comes closer to Jager and Amblard's (2004) model. There are, however, two notable innovations.

The first innovation is inclusion of keywords to construct attitude dimensions. Although the attitude dimensions I have employed are limited in emotional expressiveness to positive or negative sentiment, the model can readily incorporate much richer emotiveness due to its design. Plutchik's (1980) theory of emotion relies on "primary emotion dimensions [that] can be conceptualized in terms of pairs of opposites. Anger and fear are opposites... Joy and sadness are opposites... Acceptance and disgust are opposites... Surprise and anticipation are opposites." (p. 16) that are mixed – through linear combinations, in the language of Chapter 5 – to produce a whole slew of more nuanced emotions, eg: hesitation, annoyance, elation. A message like the one in Figure 3.2a could therefore be converted to statements, like those in Statement Set 3. Having done so, the target of a broadcast containing those statements would be able to accept/reject/ignore some or all of them based on their existing

beliefs. As such, the model is general enough for simulated motivated reasoning, with an appropriate definition of the attitude space and coding of information. It is notable that Vosoughi et al. (2018) employ a similar methodology, creating a "vector of emotion" (p. 1150) to analyze how true and false news spread in a (real) social network.

```
<s><s>Trump </s><v>bomb </v><o>Syria</o> </s><v>is </v><o>joy</o>
<s>Muslims </s><v>is </v><o>anger</o>
<s>Muslims </s><v>is </v><o>disgust</o>
```

Statement Set 3: Example of translating Figure 3.2a into statements using Plutchik's (1980) primary emotions "anger" and "disgust" to construct "contempt".

The second innovation on Jager and Amblard's (2004) discrepancy model is the inclusion of credibility to moderate the evaluation of the discrepancy. This is the ρ parameter in Equation 5.5. There are many sources of credibility and the first part of this chapter elaborates on several of them as well as proposing a method to join them. The way that information is framed also affects how it is perceived. Section 4.3 discusses this. An important aspects of how agents process information is *credibility* of the information to which they are exposed. The first section goes over the different sources of credibility and proposes a method for joining them in a single measure so that it may be used in a quantitative model like that of Section 5.2.4. The second section of this chapter concerns the incorporation of information into an agent's knowledge and it relies on Chapter 3.

4.2 Credibility

The credibility of a message is derived from a number of considerations. When greater than zero, it is a probability of perceiving a message as true; when it is lesser than zero, it is the probability of perceiving a message as false. As such, the credibility of a message is a real number in the interval $[-1, 1]$. There are infinite ways of combining the factors below and the manner that I propose consists of a linear combination passed through a sigmoid

function – in other words, using a logistic function.[1] The happy result of this is that, in future studies, the model may be readily calibrated to results from human experiments on credibility. Moreover, it should be noted that this is topologically equivalent to a single-layer neural network and, as such, it is conceivable that extensions to multi-layer neural networks and the general methods of neural network calibration (training) may be employed to reduce discrepancies between agent credibility evaluations and that of real people. This is distinct from, although intimately tied to, the calibration considerations of Chapter 6. One should note that substantial amount of research has been done on estimating trust between agents in a social network thanks largely to the advent of social media which has made these larger scale studies feasible (eg: Golbeck (2005) Ries et al. (2006)). This chapter leans on some of that body of research and also incorporates other measures of credibility which go beyond the rather static considerations of reputation and similarity to include credibility from a belief in source knowledge (expertise) and credibility due to costly rhetoric. With the exception of costly rhetoric which stems from Groeling and Baum (2008), all these sources of credibility are equivalent, if different in appellation, from McCroskey and Teven (1999) as reported in Perloff (2013). Levin et al. (2002) take up the label of "homophily" for the quality of the interpersonal relationship between a source and target. They add a dimension of "social capital" which they use to denote the strength of the relationship based on shared values and language, in contradistinction to the affinity between the parties based on demographic variables. I do not use this source of credibility since my definition of homophily – which encompases goals, values, and affiliations – extends beyond theirs which they restrict to physical similarity. By tweaking the weights in the function which combines all forms of credibility, an agent may reflect greater preoccupations with one or another form. An agent relying on heuristics would likely weigh homophily (Section 4.2.1), reputation (4.2.4), and medium (4.2.5) heavily where as an agent relying on systematic processing would weigh costly rhetoric (4.2.3) and belief in source knowledge (4.2.2) more heavily. As such, heuristic computations are *not* explicitly defined and are rather an integral, structural part of the model. It is interesting to note that in processing a message systematically,

[1] With the appropriate affine transformations so that the range of the function is in $[-1, 1]$ rather than the usual $[0, 1]$.

an agent weighs more heavily credibility sources which require more computations – in this the model also reflects the greater cognitive demands of systematic versus heuristic message processing. As Zaller (1992) points out, most people in the United States are likely to use heuristics "such as source credibility as the basis for accepting or rejecting messages" (p. 47). It should therefore be expected that the weights assigned to each sort of credibility will generally be heavier on the low-processing types. It is also expected that agents representing people of higher cognitive abilities – as proxied by the education level attained[2] – will have a greater consideration of the high-processing types of credibility. Although Chaiken (1980) stresses the importance of involvement in determining whether a person uses systematic or heuristic evaluations, Zaller (1992) calls most American political messaging "notoriously low key and uninvolving"(p. 47).

4.2.1 Credibility from homophily

Homophily is the tendency for individuals who share attributes and beliefs to form social bonds. Homophilic trust, the credibility that a person derives due to (perceived) commonalities has been observed in a number of studies from psychology (eg: Byrne (1961)) to computer science (eg: Tang et al. (2013)), which have shown that homophily was found to exist in online trust relationships. Golbeck (2009) used a data set generated from a real social network where users provided ratings for films. More precisely, (Golbeck, 2009, p.28) describes her findings:

> Through the surveys conducted here, we have shown that in addition to overall similarity, there is also a correlation between trust and the largest single difference in ratings, and between trust and the agreement on movies the source has given extreme ratings.

Her model allows for a greater accuracy in trust prediction between users and goes beyond mere similarity of the overall set of opinions – and her findings are readily amenable to the purposes of this study given that ratings are expressions of attitude. In order for an agent to form a belief on how another agent's attitudes towards a subject, agents add statements to

[2] see Section 3.3.5

their when exposed to messages. These depend features of the received statement. For example, when an agent x is exposed to <s>Trump </s><v>is </v><o>good</o> by an agent y, the source, the target, x, adds <s>agent y </s><v>like </v><o>Trump</o> to its knowledge, regardless of whether the first statement is accepted, rejected or ignored. In homophily-based credibility calculations, an agent x extracts from its knowledge the intersection between the set of statements in which it is the subject and the set where the source agent is the subject. By comparing statements with identical verbs and objects in this intersection, a measure of homophily is obtained. It should be noted that Chaiken (1980) found that the degree to which a target liked the source of a message had different persuasive effects depending on the message content. Therefore it is likely that credibility due to homophily will also vary with message content. This may be hard to reproduce in simulation.

4.2.2 Credibility from belief of source knowledge

The credibility that a person perceived as an expert derives from this perception is well-documented. Although the Milgram (1963) experiment concerned obedience, this is readily transposable to credibility from expertise since an expert may be viewed as an accepted authority. Moreover, numerous marketing campaigns have played on the *perception* of expertise. The perception mustn't necessarily reflect reality, as the ubiquitous use of a white lab coat to hawk drugs, or even decaffeinated coffee, demonstrates (O'Shaugnessy and O'Shaughnessy, 2003, p. 146). Of course, the lapse in reasoning between an actual expert and a *perceived* expert due to the use of symbols and enactment of rituals associated with a type of expert is difficult to model due to the lack of rationality this requires. Moreover, the perception of expertise may be simply derived from the frame or the medium, such as the perceived expertise of the narrator in *Der Erwige Jude* directed by Hippler (1940). Yet notwithstanding the importance this has for larger questions on the credibility of propagandistic material, let us put aside these lapses of reasoning. Figure 4.1 shows a still from an interview that Senator McCain gave on CNN following, and commenting on, the Syrian airbase strike ordered by President Trump on April 7, 2017.

This still communicates a great deal. Besides the source's name, there appears "(R)



Figure 4.1: Still from an interview of Sen. John McCain by Wolf Blitzer on CNN, on April 7, 2017 commenting on the Trump-ordered missile strike in Syria

Armed Services Chairman". The first part of that statement communicates that the source is a Republican, which is a concern of homophilic and costly rhetoric credibility. The second part labels the source as a *formal* authority on military matters: the chairman of the Senate Armed Services Committee leads oversight the nation's military, a power delegated to the upper house of the legislative branch by the US Constitution[3]. Moreover, the aware viewer would know the Sen. McCain was in the armed forces and fought in Vietnam, providing for *informal* authority. Authority, both formal and informal lead to the belief that the authority figure has superior knowledge in their area of authority. Figure 4.1 would therefore impart the viewer with the impression that Sen. McCain has expert knowledge on military matters.

The attitude changes due to credibility from expertise reflects some explanations for rallies, or their absence (eg. Brody, 1991) and more broadly ascribes to the Opinion Indexing Hypothesis of Groeling and Baum (2008).

[3]Article 1, Section 8

4.2.3 Credibility from costly rhetoric

Cheap-talk – messaging that has no reputation or other costs, or liabilities through future commitments[4] – was found to have insignificant credibility and that its converse, costly talk, or costly rhetoric, has significant credibility when concerning presidential evaluations (Groeling and Baum, 2008). In Groeling and Baum (2008), the heuristic used to determine the costliness of rhetoric was simply party identification, party of the evaluated subject, and whether the evaluation was positive or negative. This is not a trivial comparison to implement and the complexity grows when generalizing it since it requires inferring what a message source would believe. For example, if an agent knows <s>Joe </s><v>like </v><o>guns</o>, then my reader will easily conclude that the statement <s>guns </s><v>is </v><o>good</o> is cheap talk coming from "Joe". This however involves a cognitive step of recalling what "Joe" likes, agrees with, knows, etc and comparing it to the attitude vector of the received statement. More precisely, one would need to recall that "Joe" has positive sentiment towards the object "guns" and using that information, compare the sentiment towards "guns" expressed by <s>guns </s><v>is </v><o>good</o>. If the two sentiments differ in direction, ie: one is positive and the other is negative, then it would be considered costly rhetoric, which is credible; otherwise it is just cheap talk – not credible. One may note that costly rhetoric often has an opposite effect on credibility from homophily. For example, in a partisan political context, the cheap talk of a Republican criticizing a Democratic president is offset by the homophilic bond a Republican target may have with the source[5]. However, the complexity of computing costly credibility is greater than that of computing homophilic credibility. This is because the homophily between the source and the subject of the received statement must be computed, in addition to computing the feelings that the source is expected to have towards the subject. Consider the computation of costly rhetoric from Sen. McCain praising President Trump's bombardment. Both are affiliated to the Republican party and therefore share a (strong) bond of homophily, which as Groeling and Baum (2008) demonstrate, substantially reduce the credibility of

[4] Expanding the definition from Gibbons (1992) for our purposes
[5] See Hypothesis 5 of Groeling and Baum (2008)

the praise. However, to a politically aware target, the acrimonious relationship between McCain and Trump is no secret[6] and therefore it is arguably quite costly for McCain to praise Trump. It is in accessing this latter costliness that the complexity soars and it would not be surprising that this alone could sustain substantial future research.

4.2.4 Credibility from reputation

This sort of credibility emanates from the evaluations of others that one has received. When there is an attitude dimension for agreement – defined for example with keywords $\{right, wrong\}$ – then it is simply an evaluation of that dimension[7] for the source of the message whose credibility is being evaluated. In a simulation with a dimension for agreement, an agent may receive (and accept) statements of the sort <s>x </s><v>is </v><o>right</o> from other agents. It would also generate those sorts of statements when accepting *any* statement from a source. That is, if agent y receives a message from agent x with the statement <s>z </s><v>is </v><o>good</o> and if agent y accepts the statement, then it adds the received statement and <s>x </s><v>is </v><o>right</o> to its knowledge. On the other hand, if it rejects the statement received, agent y adds to its knowledge <s>x </s><v>is </v><o>wrong</o>. Agent y may then share these opinions and they will subsequently be used to evaluate the credibility of agent x. This is easily extended to agreement with the channel that carried a message. However, this should not be confused with what the (political) persuasion literature is usually termed *channel effects* which are based on the persuasiveness of what I've termed the *medium*, eg: television, to distinguish it from the carrier of the message, eg: CNN. Reputation calculations are also used for sorting the broadcast queue – for choosing whom to pay attention to. (See Section 2.4.3.)

[6]The recent kerfuffle on the Trump administration's request to hide a ship named USS John McCain from the president's view during a visit to a naval base being the latest development, despite the son and grandson of two admirals for whom the ship was named being dead since August 2018. `https://www.bbc.com/news/world-us-canada-48456742` Accessed 2019-05-30

[7]More precisely, in the language of Section 5.2.3, it is a projection of the mean attitude vector on the dimension of agreement.

4.2.5 Credibility from the medium

Medium effects, what the literature usually refers to as *channel* effects, is the persuasiveness due to the format in which the information was transmitted. The power and persuasiveness of images cannot be understated nor can the even greater persuasive power of video. Yet not all messages are more persuasive with images: Chaiken and Eagly (1976) find that complex messages are more persuasive in the written form while simple messages are more persuasive in the video form. Moreover, some people are more persuasive in some mediums and less in others, and this was starkly exhibited in the presidential debate between Richard Nixon and John F. Kennedy. As recalled in Time by Webley (2010):

> What happened after the two candidates took the stage is a familiar tale. Nixon, pale and underweight from a recent hospitalization, appeared sickly and sweaty, while Kennedy appeared calm and confident. As the story goes, those who listened to the debate on the radio thought Nixon had won[...] Those that watched the debate on TV thought Kennedy was the clear winner.

Footage from warzones, like the revolutionary live CNN coverage of the Gulf War, or made-for-television video of the Vietnam war, has been able to persuasively communicate the horrors of war to the American public. In today's image-heavy environment, it is doubtful that the announcement of a strike *without* accompanying images would be perceived as credible.

Similarly, although there have been a number of mosque mass shootings[8], the massacre in Christchurch, New Zealand was broadcast live, filmed with a Go-Pro camera in a first-person point-of-view. The acts depicted in Figure 4.2 are supremely credible in their raw violence and these have been sufficiently and starkly horrific to shake governments and industrial titans out of (willful) obliviousness and seek to limit the spread of these types of hateful materials. It is humans who redistribute these depictions of violence (Herrman, 2019). Although it is likely driven by shock (Vosoughi et al., 2018), it has to be *believable* shock. This credibility comes from the medium.

[8] Quebec City, Canada, January 29, 2017; Zurich, Switzerland, December 19, 2016



Figure 4.2: A still from the Christchurch, New Zealand massacre that was broadcast live on the internet from a first-person point-of-view

4.2.6 Credibility from direct experience

Credibility from direct experience stems from an agent having witnessed what is described by a statement first-hand, rather than having received that information from another agent. Statements of the sort <s>Trump </s><v>bomb </v><o>Syria</o> would be exogenously added to at least the "Trump" agent. Additionally, statements that are generated through agent cognitive processes, such as those agreement from Section 4.2.4, are statements that are direct experiences. All statements that are due to direct experiences are highly credible and memorable to the agent.

4.3 Framing effects

Framing effects are critically important in politics where there is often a battle to establish *the* frame (Perloff, 2013). In order to allow for what Druckman (2001) distinguishes an an *equivalency* frame, the language model has *two*, opposite keywords for each attitude dimen-

sion. While my implementation treats both keywords as equal antonyms the structure of the model is readily amenable to distinguishing between. That is, in my implementation, I treat both statements in Statement Set 4 as equivalent and assuming agents are rational, even though Tversky and Kahneman (1981) found that real people would not treat them as equivalent. This was done for simplicity of programming agent logic as well as to facilitate the proofs of Chapter 5. The data as well as the model structure to have differential effects due to equivalency frames in future studies is already present, requiring only improvements to the calculation of a statement's memorable belief, $\pi(s)$, which is currently implemented as a simple product of verb intensity, credibility, and memory remaining, ie: $\pi(s) =$ `verb_intensity*credibility*mem_remaining/100`

```
<statement intensity="-1"><s>Trump </s><v>is </v><o>bad</o></statement>
<statement intensity="1"><s>Trump </s><v>is </v><o>good</o></statement>
```

Statement Set 4: Example of two different equivalency frames

In Druckman (2001), a second type of framing is distinguished and it consists of issue framing – how an issue is presented. Statement Set 5 is an example of how President Trump's missile strike can be presented in a humanitarian context or in a dangerous expeditionary context. The first statement can be seen as either positive or a negative, *depending on the belief the agent has about humanitarian interventions.* The second statement has a less ambiguous negative connotation.

```
<s><s>Trump </s><v>bomb </v><o>Syria</o></s>
        <v>is</v><o>humanitarian intervention</o>

<s><s>Trump </s><v>bomb </v><o>Syria</o></s>
        <v>is</v><o>Mideast quagmire</o>
```

Statement Set 5: Example of two different issue frames

The model thus has the capacity to represent two types of frames. Since the frames are expressed as statements, agents that are exposed to them will evaluate them like any other, as detailed in Chapter 5. Moreover, because the frames are statements, Propositions 3 and 4 imply that beliefs in the validity of the frames may be reinforced or weakened as an agent is exposed to statements advocating a frame or a counter-frame.

5 A Mathematical Model of Attitudes

5.1 Introduction

What follows is a mathematical development of the agent reactions to information as modelled in Chapter 2 and using some basic processing rules. Definitions of the simulation specific terms is followed by several propositions. Inferring attitude from information accumulated into knowledge, Section 5.2.3, is at the core of this thesis. My model borrows substantially from the ideas of Hegselmann et al. (2002) but crucially differs in that it is the attitudes contained in an agent's *knowledge* that are averaged, rather than the attitudes of an agent's neighbours. This frees my attitude model from social network topological constraints and comes closer to the Response Axiom (Zaller, 1992, p. 49). Unlike Hegselmann et al. (2002), I do not proceed into an in-depth analysis of the longer-term opinion clustering dynamics.[1] Because of the thesis' concern, the results of this chapter seek to establish that the *short-term* attitude dynamics of agents using my information model conform to the expectations of attitude dynamics in the political science literature. I show how the same information on different agents will have different attitudinal effects. More generally, I demonstrate how the information modelled in the way of Chapter 2 is used by agents to infer the attitude towards an entity from a statement and how their exposure will change or reinforce their existing attitudes using some elementary rules.

My model, despite the apparent complexity below, is rather simple when kept to a single dimension. There is, however, a dramatic rise in complexity when moving from one attitudinal dimension to two. The discussion on the verb <v>bomb</v> in Section 5.2.2 is illustrative of this. While the model is generalizable to n attitudinal dimensions, and four

[1] I nevertheless note that the collective reaches a consensus when the threshold for acceptance is very small, that bi-polarity seems to emerge nearly always under certain conditions, and that complete fragmentation occurs when the threshold for acceptance is very large and the threshold for rejection very small. (See Equations 5.5.)

seems to be the minimum necessary for a sophisticated simulation – sentiment, importance, urgency, trustworthiness – there is a substantial amount of research needed to specify how verbs influence meaning and how to compute the discrepancy between the attitude implied by statement and held beliefs.

5.2 Definitions

5.2.1 The attitude vector and vector space

The attitude vector represents the attitudinal "charge" of a statement. For example, "Trump is bad" is a statement of negative sentiment towards the "Trump" entity. The attitude vector is a quantification of the attitudes expressed in a statement. More precisely, letting s be the statement

<s>Trump </s><v>is </v><o>bad</o>

then the subject of the statement, s_{sub}, is the entity "Trump" and s_{obj}, the object, is the entity "bad". "Bad" is associated to the attitudinal vector $\begin{pmatrix} -1 \\ 0 \end{pmatrix}$ in an attitudinal space generated by orthogonal keywords $\{good, important\}$ that is isomorphic to $\mathbb{R}^2$, and to $-1 \in \mathbb{R}$ in an attitude space generated by $\{good\}$[2]. Conversely, in the statement u

<s>Trump </s><v>is </v><o>important</o>

the entity "Trump" is the subject u_{sub} and "important" is the object u_{obj}. In an attitude space generated by keywords $\{good, important\}$, "important" is associated to the attitude vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$; in an attitude space generated only by the keyword $\{good\}$, "important" has an undefined attitude vector, unless other statements are used to define "important" in terms of "good" or "bad". The whole endeavour boils down to calculating the attitude vector associated to the subject of a statement – the *meaning* of the sentence – and then deciding whether to believe it or not given the credibility modifiers of Section 4.2 and what the agent already knows.

The attitude vector is drawn from an attitude vector space A^n which is the set of all possible $n-$dimensional attitude vectors in the simulation using the n keywords defined in Section 2.2.2. More specifically, the keywords are used to construct an orthogonal basis for

[2]Recalling that "bad" is defined as the inverse of "good". See Section 2.2.2.

the vector space A^n.[3][4][5] It must have a metric and a partial ordering. In a single attitudinal dimension, eg: sentiment, this is a lot less complicated than it seems: the metric is the absolute value of a number and the partial ordering is the natural ordering of real numbers. Although the language model is designed for multiple attitudinal dimensions, specifying the metric and the ordering of the attitude vector space is not trivial. The question of the metric has received substantial attention in the persuasion literature as it forms the core of the so-called Discrepancy Models of belief change, see eg: Fink and Cai (2013) and Wyer and Albarracín (2005).

5.2.2 Two verbs

As mentioned in Section 2.2.3, verbs are maps between attitude vectors. I provide precise definitions for two of them which are needed to show some properties of the model. I use the verb <v>is</v> and <v>bomb</v> because of the foreign crisis context of this thesis. Any other verb choice is just as valid since what matters is what the verbs *do*. Precisely defined, a verb v of a statement s is a map $v : A^n \to A^n$ such that attitudes towards the statement's object, $\vec{a}_{s_{obj}}$, map to $\vec{a}_{s_{sub}}$, the attitudes towards the statement's subject, where A^n is the attitude vector space generated by the n chosen keywords (See Section 2.2.2.) Notice that a statement s's object is denoted by s_{obj} and the attitude towards the statement's object is denoted by $\vec{a}_{s_{obj}}$. To lighten notation, the text will refer to $\vec{a}_{s_{obj}}$ with $\vec{a}_{obj}$, and similarly for the subject. For clarity, the intensity parameter is omitted from the exposition since it is a scalar which scales the effect of a statement's verb – that is, it magnifies or dampens the attitudinal effect.[6] Other modifiers, like those mentioned in Section 2.2.3 substantially complicate the definition of a verb v and are, for now, out of scope. As a shorthand and to avoid some ambiguity, $\varphi(s)$ is used to express the meaning of a statement s in terms of attitudes towards a subject: $\varphi(s) \equiv v(\vec{a}_{obj}) = v(\vec{a}_{s_{obj}})$.

The identity verb <v>is</v>. This is a fundamental verb and operates as an identity map – a function which always returns the value that was used as an argument. It's like

[3] There is an isomorphism $A^n \to \mathbb{R}^n$ which facilitates computations and computer representation.

[4] While not expounded upon there is a null additive element, by construction. A "null" entity is introduced such that attitudes in all dimensions are zero, and the entity is itself associated to the null vector.

[5] The statement intensity parameter of Section 2.2.3 is a form of scalar multiplication.

[6] For an intensity i, $i\varphi(s) = iv(\vec{a}_{obj})$.

multiplying any real number by 1. This serves to "transfer" the attitudinal charge associated with the object to the subject of a statement. Simply put, if $\vec{a}_{obj} \in A^n$ is the attitude vector of the object then for the verb <v>is</v>, $v : A^n \to A^n$, the meaning of the sentence s in terms of attitude vector towards the subject, $\varphi(s)$, is $\vec{a}_{obj}$, ie: $v_{is}(\vec{a}_{obj}) = \vec{a}_{obj} = \varphi(s)$.

The verb <v>bomb</v>. Let's define the verb <v>bomb</v>, one denoting violence by the subject on the object of a statement to be a map. Consider the ex-ante attitude vector $\vec{a}_{obj} \in A^n$ as the attitude towards the victim and $\varphi(s)$ the meaning of the sentence in terms of attitude towards the perpetrator. If A is one-dimensional and generated only by $\{good\}$, then the verb v, <v>bomb</v>, can be defined as $v : A \to A$

$$v_{bomb}(\vec{a}_{obj}) = -\xi\vec{a}_{obj} = \varphi(s) \tag{5.1}$$

This relation means that the attitudes towards the perpetrator are opposite in sign for those of the victim[7] while ξ reflects personal values on the use of violence and is, for social reasons, assumed to be positive. Negative values would imply a person derives positive emotions from violence inflicted on entities towards which it has positive feelings and negative emotions towards entities for which it has negative sentiment, while a null value would mean the person is indifferent to violence.

By adding a second attitudinal dimension using the keyword "important", a lot of possibilities arise for the definition of the verb map. I propose the following class of definitions where $\beta : A^2 \to A$ maps the sentiment for the victim and the victim's importance to sentiment towards the perpetrator. The importance attached to the *subject*, α, is meant to express the importance of an entity that could perform the action of bombing[8].

$$v_{bomb}(\vec{a}_{obj}) = \begin{pmatrix} \beta(\vec{a}_{obj}) \\ \alpha \end{pmatrix} = \varphi(s) \tag{5.2}$$

[7] Consider how attitudes towards an entity "Joe" change depending on whether "Joe bomb Iraqi chemical weapons plant" and "Joe bomb children's hospital"...

[8] In an environment with a richer verb vocabulary, the verbs "nuke", "gas", "bomb", and "strike" would be of the same family of maps (this one) and would be differentiated by the β map and α parameter. For example, to "nuke" an entity may make one more important and less liked than having merely ordered a "strike".

By specifying β in Equation 5.2 to be

$$\beta(\vec{a}_{obj}) = \begin{cases} a_{obj_2} & \text{if } a_{obj_2} < 0 \\ -\xi a_{obj_1} a_{obj_2} & \text{otherwise} \end{cases} \tag{5.3}$$

Using Equation 5.3 in Equation 5.2, the perpetrator of a bombing is attributed negative sentiments if the victim is unimportant (ie: $a_{obj_2} < 0$) no matter who the victim is. This is meant to represent negative sentiments from arbitrary use of force as well as discontent from wastefulness. (Why bomb if they're unimportant?) If the victim has some importance, then the sentiment towards the perpetrator is a mix of the victim's importance and the sentiments towards the victim. The choice of Equation 5.3 is meant to reflect abhorrence towards unjustified violence and acceptability of justified violence. The more negative the sentiments towards the victim and the more important it is, the greater the amount of positive sentiment generated; conversely, the more positive the sentiments towards the victim, the greater the negative sentiment generated towards the perpetrator. By varying ξ, the sentiment generated may be more or less responsive to the act of violence. This is a simplification that attempts to represent varying views on the acceptability of (justified) violence. Proposition 6 establishes that so-called "hawks" rally more strongly on news that a foe was bombed than "doves".[9]

5.2.3 Inferring attitude and constructing attitude moments f^k

Let $\varphi(s)$ be the attitude vector of a statement s composed of a subject s_{sub}, a verb v, and an object s_{obj}. Let $\mathfrak{I}$ be the set of all the statements the agent knows that have s_{sub} as the subject. That is, if S is the set of all statements in an agent's memory, let $\mathfrak{I} \subseteq S$ be a subset that of all the known statements such that for all $u \in \mathfrak{I}$, $u_{sub} = s_{sub}$.

Using the accumulation of knowledge principle of Chapter 3, define the k^{th} attitude moment

$$f^k(s_{sub}) = \frac{1}{n} \sum_{s \in \mathfrak{I}} (\varphi(s)\pi(s))^k \tag{5.4}$$

[9] Future developments of the model of agent behaviour should consider the functional form of β in Equation 5.3 more closely, especially if these are used for political communication simulations of issues that specifically touch on violence, like crime.

The strength of belief and the remaining memory, the memorable belief of a statement, is denoted by the scalar $\pi(s)$. $\varphi(s)$ is exponentiated element-wise and $n = \#\mathfrak{I}$ being the number of statements the agent knows about entity s_{sub}. To lighten notation, the mean attitude vector $f^1(s_{sub})$ is written as $f(s_{sub})$.

5.2.4 Accepting/Ignoring/Rejecting information

My model could be broadly categorized as a discrepancy model of belief change. The threshold value of Equation 5.5b refers to the ego-involvement of an agent while the parameter ρ denotes the credibility of the message containing the statement, particularly the credibility of the source. "The logic for the effect of source credibility is straightforward: More credible sources should widen the latitude of acceptance and, as a result, induce more belief change."(Fink and Cai, 2013, p. 91) When an agent is exposed through a message to a statement s with entity ϵ as the subject, the agent will accept, ignore, or reject the message based on the following criteria:

$$\mathbb{P}(Accept(s)|\mathfrak{I}_\epsilon) = \begin{cases} 1 & \frac{|\varphi(s)-f(\epsilon)|\sqrt{n}}{\sigma_\epsilon \rho} < t_{\alpha_a} \\ 0 & \frac{|\varphi(s)-f(\epsilon)|\sqrt{n}}{\sigma_\epsilon \rho} \geq t_{\alpha_a} \end{cases} \tag{5.5a}$$

$$\mathbb{P}(Reject(s)|\mathfrak{I}_\epsilon) = \begin{cases} 1 & \frac{|\varphi(s)-f(\epsilon)|\sqrt{n}}{\sigma_\epsilon \rho} > t_{\alpha_r} \\ 0 & \frac{|\varphi(s)-f(\epsilon)|\sqrt{n}}{\sigma_\epsilon \rho} \leq t_{\alpha_r} \end{cases} \tag{5.5b}$$

where $\sigma_\epsilon = \sqrt{f^2(\epsilon) - f(\epsilon)^2}$, the standard deviation of attitudes in $\mathfrak{I}_\epsilon$, the subset of the agent's knowledge that concerns entity ϵ. Equations 5.5a and 5.5b are readily recognized as a t-test[10] with an additional variable ρ which is a scalar combining all the measures of credibility from Section 4.2. The critical t-values t_{α_a} and t_{α_r} denote acceptance and rejection thresholds, respectively. This is similar, but distinct from static acceptance/rejection social

[10]In a general context, the numerator and σ_ϵ would reflect the use of an appropriate metric. It has been omitted from discussion for simplicity and clarity of the exposition although the choice of an appropriate metric for this evaluation is neither trivial nor unimportant. There is substantial research devoted to how people perceive discrepancy and how they accept as truth or reject as falsehood messages received (Wyer and Albarracín, 2005, p. 293). In my implementation, all statements had a single non-zero attitude dimension because of this complexity. That is, when simulating with a dimension of sentiment (good/bad) and importance (important/unimportant), statements were either expressing sentiment or importance, but not both.

judgment models.[11] Unlike usual models (eg: Sherif and Hovland (1961)) which rely on arbitrary and static discrepancy thresholds, agents perform a t-test comparing the attitude implied by a received statement, $\varphi(s)$, with the mean attitude of the information they believe and remember.[12] The greater their accessible knowledge on an issue, the less likely they are to accept a highly discrepant statement.

Equations 5.5 imply a tricotomous choice of accepting, rejecting or ignoring the message received based on held beliefs. When a statement s is either accepted or rejected, it is associated to a belief variable $\pi(s)$. An ignored statement does not enter into an agent's knowledge.

5.2.5 Attitude reinforcement and change

Recalling that $\pi(*)$ denotes the belief in a statement, if an agent believes statement s to be false, then $\pi(s) < 0$. Generally, the change on attitudes by an influential statement s with subject ϵ is

$$\begin{aligned} \Delta f(\epsilon) &= \frac{nf(\epsilon) + \varphi(s)\pi(s)}{n+1} - f(\epsilon) \\ &= \frac{\varphi(s)\pi(s) - f(\epsilon)}{n+1} \end{aligned} \tag{5.6}$$

It is noteworthy that by accepting a statement that is sufficiently counter-attitudinal, the agent opens itself up to more counter-attitudinal information, as σ_ϵ of Equations 5.5 increases. (See Proposition 3.) This models attitude change. By rejecting a counter-attitudinal statement or accepting a pro-attitudinal statement, an agent exhibits attitude reinforcement. (See Proposition 4.)

5.2.6 Attitude formation

The formation of attitudes towards an entity ϵ, in this context, occurs when there is no knowledge of statements containing it as a subject, ie: for $\mathfrak{I} \subseteq S$, $\#\mathfrak{I} = 0$, or when there is insufficient knowledge to have any certainty on existing attitudes, ie: $\#\mathfrak{I} = 1$. A number

[11] Fink and Cai (2013) note that the empirical evidence for social judgement is "not strong".

[12] This is a loose synthesis of the Ajzen and Fishbein (1977) and Wyer Jr and Hartwick (1980) models, as reported in Wyer and Albarracín (2005).

of possibilities arise to decide whether an agent accepts, rejects, or ignores a statement on a new subject. Two that I have considered are immediate acceptance, subject to the credibility (of the source, medium, etc) associated to the statement, and random acceptance with probabilities proportional to the credibility. By accepting a statement in these manners when attitudes are yet to be formed, an agent conforms to the findings of Pennycook et al. (2018) as subsequent exposures are evaluated to this anchoring beginning. Nonetheless, these attitude formation methods may be arguably simplistic given that Shen and Bigsby (2013) reports a large body of literature on attitude formation. Without minimizing the importance of attitude formation, I did not focus on attitude formation as much as attitude *change* and *reinforcement*, given the scope of the thesis. The results of this chapter rely on formed attitudes. Yet it should be noted that when these are weakly formed – that is, when the number of statements in an agent's memory is small – the attitudes are *elastic*, in the sense of Baum and Groeling (2010). This is shown by the converse of Proposition 1.

5.3 Propositions

5.3.1 In a general context

Proposition 1. *Any agent that has been exposed to more information on a given subject will,* ceteris paribus*, be less influenced by new information it accepts or rejects.*

The model thus follows Baum (2002) which reasons that highly informed and politically aware individuals are less likely to be influenced by new information. On the other hand, it differs from the model in Jager and Amblard (2004) where information is always incorporated at the same rate regardless of how much information the agent already has. In Baum (2002), politically unaware individuals will also be less responsive to information. This property needs a second attitudinal dimension and is shown in Proposition 5.

Proof. Equation 5.6 shows that the attitude vector of $\varphi(s)\pi(s)$ is given a weight of $n+1$ in post-exposure attitudes when the statement is either accepted or rejected. More information

on an entity ϵ implies a bigger set $\mathfrak{I}$. As such,

$$\frac{\varphi(s)\pi(s)}{n+1} < \frac{\varphi(s)\pi(s)}{m+1}$$

for $n = \#\mathfrak{I}$ and $m = \#\mathfrak{I}'$, $\mathfrak{I}' \subset \mathfrak{I}, m < n$. Thus we see that the same information is given less weight the more an agent has information on ϵ, the statement's subject. □

5.3.2 In a world with a single attitude dimension

Proposition 2. *The greater the difference between an agent's* ex ante *attitude and the inferred attitude of a statement it accepts, the greater the effect of the new piece of information on* ex post *attitudes.*

Proof. Let κ be the difference between an agent's *ex ante* attitudes and the attitude of a statement which is accepted or rejected, ie: $\kappa = |\varphi(s)\pi(s) - f(\epsilon)|$. Suppose another agent perceives the statement s just as credible as the first, has the same amount of information on ϵ but has different mean attitudes towards it. That is, $\mathfrak{I}'_\epsilon = \mathfrak{I}_\epsilon$ but $f'(\epsilon) \neq f(\epsilon)$. Without loss of generality, suppose

$$\kappa' = |\varphi(s)\pi(s) - f'(\epsilon)| < |\varphi(s)\pi(s) - f(\epsilon)| = \kappa$$

then it is immediate that the *ex post* effect of the same piece of information s on agent attitudes is greater for the agent holding *ex ante* mean attitude $f(\epsilon)$ than the one holding *ex ante* mean attitude $f'(\epsilon)$. That is,

$$\frac{\kappa'}{n+1} = |\Delta f'(\epsilon)| < |\Delta f(\epsilon)| = \frac{\kappa}{n+1} \tag{5.7}$$

and hence the result. □

Corrolary 1. *If ϵ denotes the US President and s denotes the statement informing of a rally event, an agent who already favours the president will rally less than one who favours the president less or opposes him, all else being equal.*

This Corollary shows that the model reflects Hypothesis 1 of Baum (2002). It is also the basis from which the results in Chapter 7 stem.

Proof. Let $f^+(\epsilon), f^-(\epsilon)$ be mean attitude vectors denoting the feelings associated to an entity ϵ (the president) such that $f^-(\epsilon) < f^+(\epsilon)$.[13] Then for an accepted message s such that $\varphi(s)\pi(s) > f^-(\epsilon)$,

$$\Delta f^+(\epsilon) = \frac{\varphi(s)\pi(s) - f^+(\epsilon)}{n+1} < \frac{\varphi(s)\pi(s) - f^-(\epsilon)}{n+1} = \Delta f^-(\epsilon)$$

□

Attitude change and reinforcement. Propositions 3 and 4 prove the existence of attitude change and attitude reinforcement in the model.

Proposition 3. *An agent that accepts a counter-attitudinal statement is guaranteed to be more likely to accept other counter-attitudinal statements on the same subject.*

Proof. Define a counter-attitudinal statement as one with a meaning that implies an attitude towards the subject that is different from the attitude an agent holds. That is, suppose an agent has a mean attitude towards an entity ϵ and a statement s has subject $s_{sub} = \epsilon$, then a counter-attitudinal statement is one such that $|f(\epsilon) - \varphi(s)| > \varepsilon$. The threshold for a statement to be considered counter-attitudinal, ε, is rather ambiguous, though it is relatively large and obviously positive.

If an agent accepts a counter-attitudinal statement, three things occur. First, the discrepancy between *ex ante* attitudes and the statement's implied attitude towards a subject will be greater than *ex post*. That is, if an agent has a mean attitude towards a subject ϵ of $f(\epsilon)$ prior to being exposed to the statement, and a mean attitude of $f'(\epsilon)$ after accepting the statement, then $|f(\epsilon) - \varphi(s)| \geq |f'(\epsilon) - \varphi(s)|$. Thus the numerator in Equations 5.5 decrease. The second effect is that the variance in the knowledge set increases whenever a statement with an implied attitude sufficiently different from the *ex ante* mean is added. This increases the denominator of Equations 5.5. Taken together, these two effects make the agent more open to acceptance, in the sense of Equation 5.5a. Third, as credibility of a

[13] Using the order associated to the attitude vector space

source is increased after accepting a statement in a broadcast from that source, ρ increases, further increasing the denominator in 5.5a and making future acceptance of statements with a similar meaning to s if they originate from the same source.

□

Proposition 4. *An agent that accepts a pro-attitudinal statement is guaranteed to be less likely to accept other counter-attitudinal statements on the same subject.*

Proof. Define a pro-attitudinal statement s to be such that the believable attitude is approximately equal to the pre-existing mean attitude, ie: $\varphi(s)\pi(s) \approx f(\epsilon)$. Let an agent have a mean attitude $f(\epsilon)$ towards a subject ϵ and accept a statement s such that $\varphi(s)\pi(s) = f(\epsilon)$. Then it is immediate that for a subsequent exposure to a counter-attitudinal statement u, $|\varphi(u) - f'(\epsilon)| = |\varphi(u) - f(\epsilon)|$ since $f'(\epsilon) = f(\epsilon)$. And since $f(\epsilon) = f'(\epsilon)$, $\sigma_\epsilon \leq \sigma'_\epsilon$. The variance of an agent's attitude towards an entity ϵ decreases when accepting a pro-attitudinal statement – it becomes more certain of its attitudes. This decreases the denominator in Equations 5.5 and consequently increases the t-statistic making acceptance of counter-attitudinal statements less likely. This proves the existence of attitude *reinforcement* in the model. □

More generally, the model also ensures *reinforcement* behaviour if the attitude of an accepted statement falls within a narrow range. Out of that range, the agent becomes more open to *all* sorts of statements, as Figure 5.1 illustrates. Figure 5.2 provides a closer look at attitude reinforcement (in the first column) and attitude change (in the second column) illustrating how the likelihood of accepting messages – and indeed, openness to other opinions – decreases and increases, respectively.

5.3.3 In a world with two attitude dimensions

Let attitude space A^2 have $\{good, important\}$ as the basis.

Proposition 5. *Agents who are politically unaware will rally less.*

It should be noted that if agents are focused on foreign events, then they would have higher values of importance attached to foreign entities and by the same logic, would rally

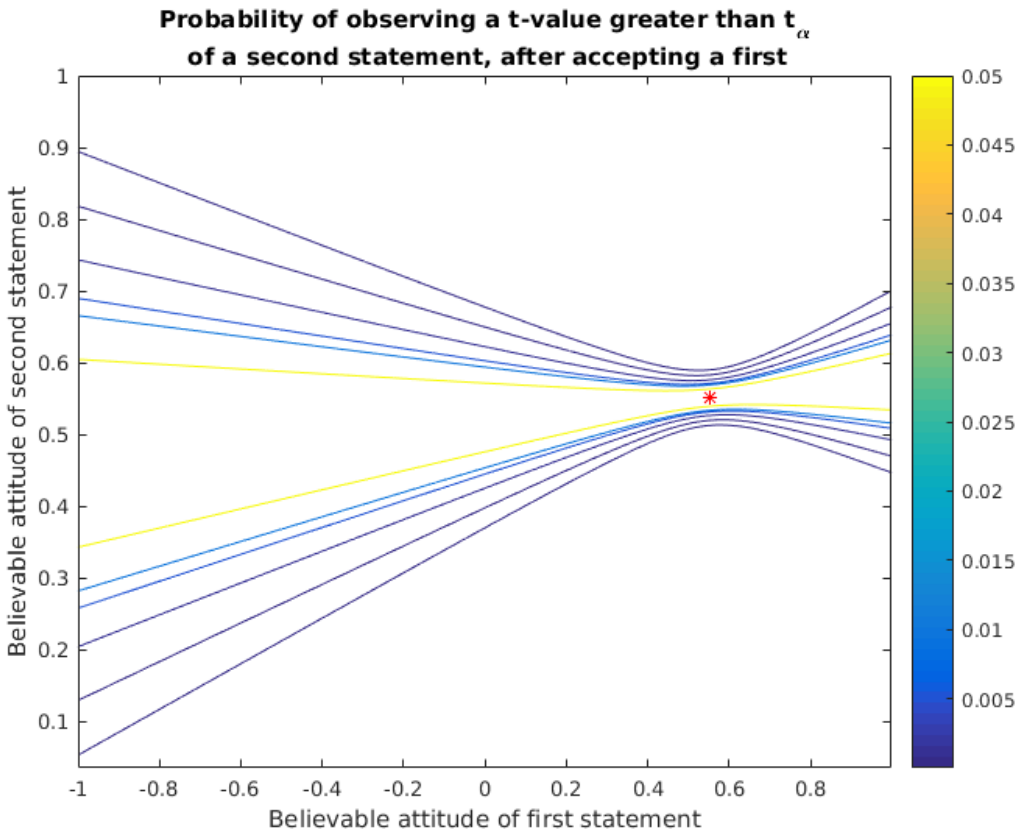


Figure 5.1: An illustration of attitude change and reinforcement. The mean attitude prior to the first accepted statement is shown with an asterisk

more. However this proposition must be considered along with Proposition 1, recognizing that highly-aware individuals would a firmer belief on their sentiments and would rally less. Combining both propositions, we arrive at the argument of Zaller (1992) where those who rally are neither highly aware nor unaware.

Proof. Assume that agents who are politically unaware will grant more importance to non-poltical matters and correspondingly less to political ones. With this assumption, it is reasonable to further assume the entity "Syria" is attributed a relatively small, if not negative, importance. By comparison, a politically aware agent would likely have been exposed to a much greater number of statements of the sort <s>Syria </s><v>is </v><o>bad</o> and <s>Syria </s><v>is </v><o>important</o>[14]. As such, when exposed to a statement like <s>Trump </s><v>bomb </v><o>Syria</o>, and using a definition like

[14] A real utterance that produces these twin statements could be "Defense Secretary Jim Mattis confirmed recent press reporting that Syria still possesses chemical weapons" McLeary and Rawnsley (2017).

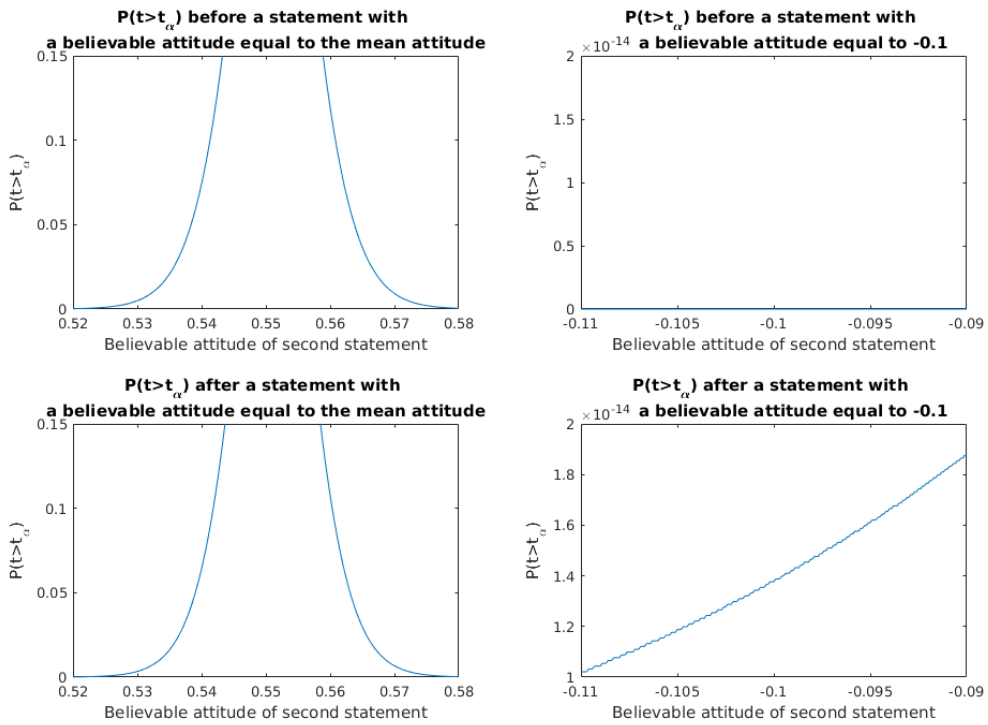


Figure 5.2: An illustration of attitude change and reinforcement showing the likelihoods of accepting a statement before and after accepting a first. Both sets of illustrations were done with $f(\epsilon) \approx 0.5510641$, $f^2(\epsilon) \approx 0.3046975$ and $n = 19$. The first column shows a narrowing of the curve denoting a greater likelihood of accepting statements close to $f(\epsilon)$ and a lesser likelihood of accepting statements further from $f(\epsilon)$. This is reinforcement. Conversely, the second column shows the dramatic increase in the likelihood of accepting a counter-attitudinal statement after having accepted a first.

in Equation 5.2 it is easy to see how the effects of the rally would differ. Crucially, if both politically aware and unaware agents held the same sentiment about "Syria" but differed in their evaluation of its importance, those which evaluated "Syria" as more important would attach a greater sentiment to the "Trump" entity. □

Proposition 6. *Agents who prefer to rely on military strength for peace rally more than those who prefer peace through diplomacy when a foe is attacked.*

Proof. Different individuals have different perceptions of how to best achieve peace and security. Some prefer to rely on military strength, so-called *hawks*, and others on diplomacy, *doves*. (See Pew Research Center (2017) Q25i.) Making the assumption that doves derive less positive sentiment than hawks for any use of violence, then $\xi_{Hawk} > \xi_{Dove}$. Using

Equations 5.3 and 5.2 makes the result immediate. □

II Simulation

6 Creating a Synthetic Population

6.1 Introduction

This chapter focuses on the instantiation of a collective, or how an artificial population is created. The previous chapters established how information is modelled and how agents store and process it into knowledge. In this chapter, the focus is on calibrating the agents so that their available information, attitudes, behaviours, and social links resemble those of a real population. In order to do this, both empirical and simulated data must be used. Section 6.2 explains how to use a clustered population from survey data. Section 6.3 goes over various training methods, or how the simulation can be "warmed-up". It is during this phase that the agents instantiated to reflect empirical survey data are made to interact so that they form the weak social bonds that Gibson (2001) notes are what facilitate the diffusion of novel information and political discussion while arguing their necessity for the existence of civil society(Levin and Cross, 2004). Section 6.4 lists the types of data which are currently hard to find and how one could go about obtaining them to generate a more accurately calibrated synthetic population.

Notably absent from previous chapters are discussions of agent behaviour, in particular what to broadcast and how to voluntarily expose themselves to information. The media and political elite are also agents, if a special type. In computer science parlance, they are derived classes from the Agent base class and as such inherit the behaviours and properties of common agents, representing a member of the public, while modified to reflect their differences from the mass public. I recognize the substantial amount of research devoted to understanding how the mass media act as rebroadcasters of elite rhetoric in the specialized context of foreign policy, eg: Brody (1991), Groeling and Baum (2008), Zaller and Chiu (1996), Baum and Potter (2008). The term "rebroadcasters" is by no means used to imply

an uncritical, conveyor-belt role for the media – rather, it is in the sense of rebroadcasting of Section 6.4.1. Rebroadcasting is used in the simple sense of broadcasting information that was received from a broadcast with no implications on how an agent decides what information is worthy of rebroadcasting. Section 6.4.2 assumes that members of the public, the elite, and mass media seek out information in broadly similar ways, differing only in the weight assigned to statements on specific subjects and on potential sources. A distinction between agents representing members of the public and those representing a political elite or a mass media outlet is that the latter types may be endowed with information exogenous to the simulation. This is meant to allow for the injection of new information into the collective. This is by no means a reduction of the general public to a passive mass who merely accept information (Groeling and Baum, 2008, p. 1069). It is meant to bring in new information, like the growth of the leaders' information in uncertain scenarios (Baum and Potter, 2008, p. 42). *In extremis* the (political elite) agent representing the president only broadcasts what has been determined by the modeller. For simplicity, the proof-of-concept simulation in Chapter 7 has a single media agent that acts in this way, eschewing the additional layer of a presidential agent announcing the strike in a news conference.[1]

6.2 Using a political typology

The Pew Research Center publishes every so often a so-called American Political Typology, like Pew Research Center (2017). This sort of study has evolved over the years and has become, in my opinion and for my uses, more accurate. The aim of the typology is to cluster the American public "based on their attitudes and values – not their partisan labels."[2] The typology clusters respondents on social and political attitudes, along with party affiliation. The responses in Pew Research Center (2017) are used to randomly generate a collective of desired size with traits and initial knowledge reflecting the data. (See Chapter 3 and particularly Section 3.3.7) The Typology contains information that can be attributed to fixed agent characteristics (eg: sex, age) and to important attitudes (eg: Trump job approval, QA1; Agreement with Donald Trump, CB36). It also contains useful information on the

[1] See `https://www.youtube.com/watch?v=4sVp3yFNEYQ` Accessed 2019-06-09
[2] `https://www.people-press.org/2014/06/26/typology-comparison/` Accessed 2019-06-08

national agenda (eg: Health care coverage as government responsibility, QA126), foreign policy preferences (eg: US role in the world, Q50ee) as well as information that influences agent credibility of sources and channels (eg: Views on the national news media, QA14e; Views on experts, QB52; Party favorability, QA15a,b). There is also information that may be used to condition agent mood (eg: My side has been losing more often, QA30; Satisfaction with state of the union, QB2) attention to news and politics (QB53a-c, QB54, QB27), deeply held values (eg: views on abortion, QA129) and very importantly, behaviour. For example, questions QB55a and QB55b ask about the effect on friendship when learning a friend voted for Trump and Clinton. This provides a hint of how an agent would react when exposed to counter-attitudinal statements. These sorts of questions are those which help guide the pre-simulation training of the synthetic population where agents share information and form social bonds. It is rather unfortunate that the results of the Typology are presented as independent observations and conditional responses are seemingly unavailable. The link between all the responses is important in order to create an accurate collective.

6.3 Training synthetic population

When an artificial population is created using the suggested data of Section 6.2 or using another clustering, like Kim et al.'s (2010) use of the NAES survey, it is only the instantiation of an agent's initial knowledge. Crucial to a simulation with a collective are the *bonds* between agents. Although a well-done instantiation would provide for initial attitudes towards *some* agents, like the one representing the president, nearly all the $n(n-1)$ possible agent bonds would uninitialized. In other words, the agents generated in Section 6.2 are isolated entities deprived of social bonds. These are created through a training – or "warming-up" – phase.

In order to create a social network[3] there are a number of conceivable ways to have agents interact and "get to know" each other, which I refer to as mixing patterns following Jager and Amblard's (2004) use (p.297). Jager and Amblard (2004) use two types of mixing patterns: "full" mixing, where agents interact with any other agent, and geographic

[3]In the sense used by Scott (1988), not to be confused with specific (virtual) social networks contained on *social media* platforms like Facebook

mixing where agents interact with neighbours. Full mixing is fairly self-descriptive but it is conceivable to vary the number of simultaneous targets for a source's broadcast, as well as the selection mechanism of source and target. Jager and Amblard (2004) uses a uniform selection of agent pairs with reciprocal interaction – that is, a source and a target are selected, an interaction is forced, after which the roles are inverted and a second interaction is forced. Other interaction patterns, both for full mixing and geographic mixing are conceivable. Geographic mixing conditions the target selection on the selection of a source. Jager and Amblard (2004) placed agents on a square grid – endowing them with a geographic location – and allowed each agent to "interact only with the four direct neighbours: North, South, East and West" (p. 300) or in a so-called von Neumman neigbourhood of radius one. It is obviously possible to increase this radius, as well as use other spatial initializations. For example, Pew Research Center (2017) uses the four geographic regions for respondents traditional in US public opinion surveying: Northeast, Midwest, South, and West. With much finer grained data, sub-county and sub-congressional district groupings are conceivable. This is similar to how Krebs and Ernst (2017) use marketing data for both placing the agents spacially on a map of Germany and to determine their initial characteristics before creating random social bonds based on geographic proximity and homophily. It is unfortunate that the data does not yet seem to exist to reproduce in a simulation the reaction of small isolated town to news of a rally event, or even if it did, to validate the simulation results.[4]

Full mixing patterns approximate the connectivity of the internet while geographic mixing patterns provide an approximation of local social networks. Both of these could be augmented by using agent characteristics (see Section 3.3) to further condition the selection criteria. Mixing can be deterministic, like a "speed-dating" event where the source-target pairs are pre-determined; with uniform randomness, like in Jager and Amblard (2004) or the simulation in Chapter 7; or with conditioned randomness, like a networking event. When applying a mixing pattern, which ever it may be, it is assumed that agents will interact and through this interaction determine whether – simply put – they like each other or

[4]There have been attempts at disaggregating rally data, eg: Edwards III and Swenson (1997), Baum (2002), but none seem to be so fine grained and geographically precise.

not. By applying a mixing pattern to form the social network, one is in fact conducting an multi-agent simulation. Neighbourhood mixing is another type of pattern. An initial social is embedded on a randomly-generated graph. These graphs join nodes so that certain statistical properties hold. One could, for example, generate a graph so that the number of links a node has with other nodes follows a probability distribution. There are much more sophisticated model, such as those of Bramoullé et al. (2016), particularly those based on homophily which are related to those used by Krebs and Ernst (2017). A randomly-generated graph is used to determine which agents can interact with which. Then agents are made to interact with those with which they are connected with a given path length or less. If the maximal path length is two, then agents can interact with their "friends" and their "friends' friends".[5] The generation of a random graph is more difficult than simpler mixing patters. I therefore employ a simple full mixing pattern to weave a social network in the proof-of-concept simulation of Chapter 7.

6.4 Obtaining more complete data

In science there is nearly always a need for more data; the calibration of a synthetic population of the purposes of a political simulation is not spared. As noted in the above sections, there are data needs to calibrate a collective to the population one wishes to study. In Section 6.2, I highlighted the need for *conditional* response probabilities on a survey rather than only the marginal distributions. It is useful to know that, for example, 25% of "Devout and Diverse" agree on issues with Donald Trump and 74% disagree (Pew Research Center, 2017, CB36) and that 29% of the same group approve of Donald Trump's job as a president while 60% disapprove (Pew Research Center, 2017, QA1), but the conditional distribution is missing. How likely is a "Devout and Diverse" person who agrees with Donald Trump on issues to also approve of his job as president? This is important to accurately generate a synthetic population.

[5] While agents could be endowed with knowledge about their social neighbours, the selection of this knowledge brings up questions which I feel are better answered at the agent level, rather than the simulation level. Placing the problem of creating initial knowledge in agent design, I feel this facilitates the creation of heterogeneous collectives, where different types of agents, with differing extroversion, for example, may have different propensities share information on themselves to other agents.

Another need for data is to calibrate agent *behaviour*. The measurement of some agent behaviour is not obvious, such as voluntary media exposure (Jerit et al., 2016). And when media exposure behaviour is known, it is still not simple to determine the true media preferences – as opposed to stated preferences – of a person, as the complex experimental design of de Benedictis-Kessner et al. (ND) shows. Sharing, or the propensity to rebroadcast, is the also examined in de Benedictis-Kessner et al. (ND), but unfortunately the results are tied to media preferences rather demographic or other variables which could easily be linked to population clusters like Pew Research Center's (2017) and ultimately to agents. Likewise, Vosoughi et al. (2018) present findings that aggregate the Twitter user sample as a homogeneous population, but surely not all demographics rebroadcast information in the same way. Understandably, the data that has been thus far collected was not done with the explicit aim of calibrating a synthetic population. Future studies would likely benefit from the development of a measurement method that incorporate both demographics, values, attitudes, preferences, and behaviours.

6.4.1 Broadcasting

In this section I distinguish between broadcasting and rebroadcasting. The former is self-generated by the agent and is a mechanism for it to share its identity and its attitudes to other agents. The latter is the unmodified sharing of information. More concretely, if an agent is exposed to a statement s, broadcasting would involve a statement like <s>s </s><v>is </v><o>right</o> while rebroadcasting is simply sharing s to another agent. A media agent, acting as a channel will nearly always be *rebroadcasting*, if only its journalists' broadcasts or its users' tweets. Since a rebroadcast is itself a broadcast (by the agent rebroadcasting), rebroadcasting is simply a specific choice of what to broadcast.

Choosing what to say/broadcast. Given the diversity of information in this model and amount of knowledge an agent has, determining what to broadcast is not trivial. There is likely to be a strong component of autocorrelation in the subjects chosen between subsequent broadcasts but this requires further studies. For simplicity and given the unique focus of my proof-of-concept simulation in Chapter 7, I programmed agents to share two messages in each broadcast: the first is their attitude towards "Trump" and the second is whatever

is most salient and credible.[6] One could use other criteria to generate broadcasts, such as the guidelines in Appendix B.2.

Choosing the channel. Once a statements to broadcast have been selected by an agent, it then must package them into a broadcast. To do so, it must have a target, which may have been chosen before choosing *what* to broadcast, especially if choosing what to broadcast is conditioned by its intended target. Next, the agent needs to select a channel which will carry the broadcast. (All other parameters of a broadcast are already determined.) I have not studied, nor have I encountered studies on how a person chooses a carrier for his or her messages, nor have I encountered studies on how a person chooses the medium (format) for a message. Nevertheless, choosing these along with the target and the message content is the domain of marketing – the main concern of which is persuasion. Accordingly, the prevailing wisdom in marketing is that choosing the right "channel" – which is what I've termed medium – is done by reviewing the current (marketing) channels used, evaluating the fit between a channel and the target, deciding whether to use known or unknown channels.[7] With some additional research, agents could be programmed to behave in a analogous fashion to choose their medium. In the proof-of-concept simulation of Chapter 7, there is a single medium, for simplicity. Given the interaction dynamics, it could be assumed to be sound.[8] Once the medium has been chosen, the choice of channels is narrowed. For example, one can post a web video (medium) on only so many social media platforms (channels). How to make a choice among these is an open question, except for an agent representing a journalist, which would choose the channel for which it is employed, and a media agent, which would choose itself. In my proof-of-concept simulation, all agents choose the same, physical, channel.

Propensity to rebroadcast. It is interesting to note that the qualities that make information newsworthy (Groeling and Baum, 2008) – and therefore worthy of being rebroadcast – save balance, are strikingly similar to what makes *people* share information. Vosoughi et al.

[6] Specifically, the most credible, memorable, intense, and recent statement in their knowledge. See the second set of sorting criteria in the selection of statements to broadcast in Appendix A.2.

[7] `https://imaginasium.com/blog/how-to-choose-the-right-channels-for-your-marketing-message/`, Accessed 2019-06-09

[8] That is, face-to-face talking and a media agent transmitting with sound, eg: a loudspeaker. The media agent could also conceivably be a person entering the room with a megaphone to announce the news.



Figure 6.1: Same channel, different mediums. A channel's broadcast in one medium linking to the same broadcast in another medium

(2018) find that people share information that is novel and that produces emotions related to conflict, such as fear and disgust. The need for balance in *non-partisan* mass media was corroborated in an interview I conducted with Peter Johnson, Executive Editor for the CBC, in November 2018. He stressed the emphasis placed on verification of the news – this, he pointed out, is not a practice followed by all news media that sometimes rush to get a story out without more than a single dubious source. These less-reputable media channels along with partisan media may act more akin to common members of the public than the more reputable mass media. Similarly, there are more and less conscientious members of the public and of the political elite. Finding these striking similarities, I let the varying propensities to rebroadcast, and the choices of *what* to rebroadcast, be expressed

through sorting criteria of the agent's knowledge. A bot, like the automated retweeters of Vosoughi et al. (2018), would simply sort known broadcasts on recency and perhaps source. A news media would likely sort on source, recency, costly rhetoric, and counter-attitudinal criteria. A conscientious member of the public may only rebroadcast statements which have been received from more than a minimal number of sources. I acknowledge that the differences of how different types of agents choose to rebroadcast information may be sufficiently substantial to merit greater specification in future refinements of the model. What *does* strongly distinguish a mass media outlet from a member of the public is the tendency to broadcast across a multitude of mediums nearly simultaneously. Figure 6.1 illustrates this. As Mr. Johnson noted, CBC News broadcasts its reporters' stories on the web (in print and video), on radio, on social media posts across a number of platforms, and on television – an ordinary member of the public, being constrained in resources, will not be capable of such diversity of mediums but will, in rarer cases, employ a diversity of channels, cross-posting, for example, the same video on YouTube, Facebook, and Twitter.

Propensity to append comments to broadcasts. A comment is a broadcast that is appended to another, base broadcast. The root broadcast, in the case of a web article, the root broadcast is the main article and it acts as the base broadcast for the comments directly related to it. A comment can be a base broadcast for a subsequent comment posted in reply to it. In the case of a two person conversation, agents alternate appending broadcasts to each other's last broadcast. This allows for tracking of agent interactions, both on and off line.

6.4.2 Information seeking: deliberate exposure

How and why humans seek information is a rich field of study. Understanding the motivation to seek out more information is important but "although a large number of studies have been done on information-seeking behaviour, the process itself is still largely a mystery."(Weiler, 2005, p. 49). In my proof-of-concept simulation, agents do not seek out information. Instead, the simulation reflects a forced exposure to information. This could be improved by making an agent desire more of the same subject and object when exposed to a new statement – in the sense of there being few or no prior exposures to the statement – or for statements on

a specific subject when the attitude variance[9] for this subject rises above a threshold.

Involuntary exposure. This is the sort of exposure to information resulting from random events. It is up to the modeller to define these and they add noise to the simulation, which may be beneficial in order to introduce new information into a population subgroup. The forced exposure component of de Benedictis-Kessner et al. (ND) is an example of involuntary exposure.

[9]See Section 5.2.4

7 Proof-of-Concept Implementation

7.1 Introduction

This chapter first describes a proof-of-concept simulation which demonstrates the foundational workings of the information model. A discussion then follows on the similarities and differences between simulation results and those observed in reality.

7.2 Simulation description

The agents representing the members of the public in the simulation had an identical instantiation except for their initial knowledge endowment. Thirty agents were each given a statement allowing them to form an initial attitude towards the "Trump" entity. Agents had initial sentiment towards the "Trump" entity uniformly distributed from -1.00 to -0.30 and 0.30 to 1.00 in 0.05 steps; agents were not forgetful for simplicity in this very short simulation. A media agent was provided with the two statements in Statement Set 6: the first expressing negative sentiment towards the entity "Syria", the second announcing the strike.

```
<statement intensity="0.80">
<s>Syria </s><v>is </v><o>bad</o>
</statement>

<statement intensity="0.05">
<s>Trump </s><v>bomb </v><o>Syria</o>
</statement>
```

Statement Set 6: The statements in the media agent's broadcasts

In order to reduce the number of statements in the simulation, agents were programmed to express their sentiment towards the "Trump" entity choosing one of the forty provided that approximated their feelings best. This is the first sort criterion in Appendix A.2. The data file for the simulation is in Appendix A.3. The attitude space defined in this data file is three dimensional. There is a dimension for sentiment expressed with keywords "good" and "bad". Orthogonal to sentiment, the "important" and "unimportant" keywords define importance; "right" and "wrong" define agreement, which is used to compute credibility from reputation.[1] There are four verbs, although "helped" is not used. The verb "know" is used for agent-generated statements and to express knowledge about other agents, which is used in homophily calculations.[2] The verb "bomb" is used to express violence from the statement subject on the object. It maps attitudes as defined in Section 5.2.2. Conceptual association is done through the verb "is" and is functionally defined is Section 5.2.2.

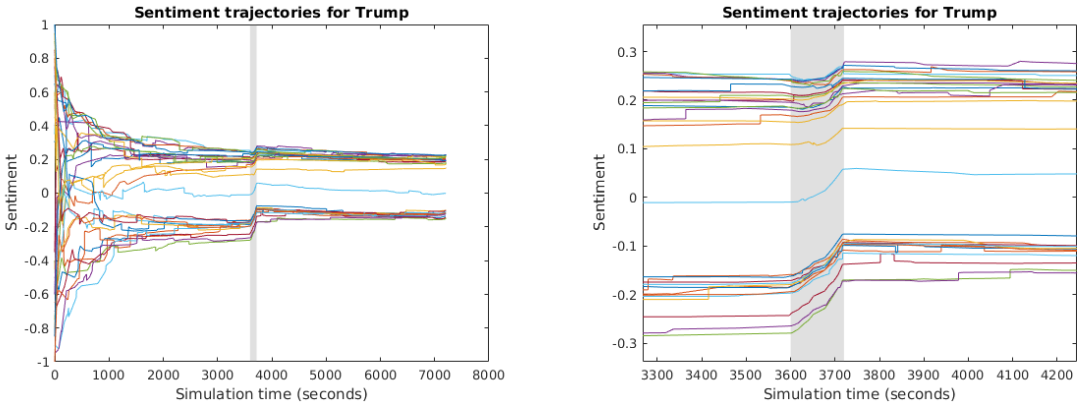
Despite the importance of partisanship[3] to evaluate costly rhetoric[4] and homophilic bonds[5], the simulation was kept at its simplest in order to clearly show the attitude change of the collective, and in particular that the collective exhibits the reaction predicted by Proposition 2 and Corollary 1, as Figure 7.1 shows.

7.2.1 Mixing the collective

The simulation clock advanced in seconds. Agents were allowed to fully mix[6] for one sim hour, during which they shared their attitude towards the "Trump" entity as well as the most salient statement in their knowledge. The mixing pattern involved uniformly selecting a pair of agents at random, assigning one as the source and the other the target, and forcing them to broadcast two statements. Then the roles were reversed – the source became the target and *vice versa*. Selections were done in randomly spaced time intervals following an exponential distribution with parameter $\lambda = \frac{n}{120}$ where $n = 30$, the number of agents in the collective. This implies that an agent will, on average, initiate a conversation (as a

[1] See Section 4.2.4
[2] See Sections 2.2.3 and 4.2.1
[3] Section 3.3.7
[4] Section 4.2.3
[5] Section 4.2.1
[6] See Section 6.3

source) with another agent every two sim minutes. Given the sparse amount of information, the salient statement shared is almost surely an opinion on another agent, eg: <s>Agent 3 </s><v>is </v><o>wrong</o>. Agents would accept/reject/ignore[7] all statements received using a 5% acceptance threshold and a 1% rejection threshold that incorporates a credibility measure. This leads to a collective with bi-polar attitudes towards the "Trump" entity, as Figure 7.1a shows.



(a) Entire simulation

(b) Close up during news bulletin

Figure 7.1: Sample sentiment trajectories of a collective exposed to a news broadcast

7.2.2 The broadcast

After about one sim hour, the collective was deemed to have formed sufficient social bonds and hold informed beliefs on "Trump", the media agent broadcast to all agents, simultaneously, the two pieces of knowledge it had. This broadcast had a duration of approximately two sim minutes and is shown in the grey bands of Figure 7.1. During this time, the media agent repeated a broadcast about twelve times with two messages meant to recreate the actual media messaging of April 7, 2017. Most media broadcasts, like that of Figure 2.1, led with <s>Trump </s><v>bomb </v><o>Syria</o> followed by recalling that <s>Syria </s><v>gas </v><o>civilians</o>. However, the media agent in this

[7] See Section 5.2.4

simulation eschews this order and instead adopts the one chosen by Mr. Trump in his statement[8] which first recalls the gassing of civilians before announcing the strike. The duration and repetition was meant to mimic that of actual media broadcasts[9]

The reason that I opted for a news broadcast of <s>Syria </s><v>is </v><o>bad</o> rather than the closer approximations of <s>Syria </s><v>gas </v><o>civilians</o> or <s>Syria </s><v>gas </v><o>children</o> is not so much because of the need to define an additional verb "gas" – which would be functionally similar to "bomb"[10] – but rather because of the need to endow agents with an additional piece of knowledge, namely <s>civilians </s><v>is </v><o>good</o> or <s>children </s><v>is </v><o>good</o>. With agents endowed with a single piece of knowledge – a sentiment statement on the "Trump" entity – the salient statement that agents would share during their interactions is almost surely guaranteed to be an opinion about another agent. On the other hand, if agents were endowed with <s>children </s><v>is </v><o>good</o>, they would likely share this initially as the salient piece of knowledge – because of the simple broadcast selection logic I wrote.[11] Agreeing on the general goodness of "children" would dampen the effects of disagreement on "Trump" – this, of course, is good in a real context where national cohesiveness is important but in this case it muddles the bi-polarization which was desired to analyze the effect of the news. Furthermore, in extreme cases of disagreement on attitudes towards "Trump", a target may *reject* <s>children </s><v>is </v><o>good</o> because it distrusts the source so much. In doing so, the target would decrease its sentiment towards "children". The statement <s>children </s><v>is </v><o>good</o> is necessary to infer that <s>Syria </s><v>is </v><o>bad</o> from a statement like <s>Syria </s><v>gas </v><o>children</o>[12]. By eliminating the need to infer that <s>Syria </s><v>is </v><o>bad</o> from <s>Syria </s><v>gas </v><o>children</o>, I eliminate the possibility of heterogeneous attitudes towards "children", thereby ensuring that the collective processes the media broadcast in an attitudinally homogenous manner.

[8] https://www.youtube.com/watch?v=4sVp3yFNEYQ Accessed 2019-06-09
[9] eg: New York Times https://www.youtube.com/watch?v=5RzfJiyNoZM Accessed 2019-06-09
[10] See Section 5.2.2
[11] See Section 6.4.1
[12] See the description of inference in Section 2.2.2

7.2.3 A second mixing results in

During and following the broadcast, the collective was made to continue mixing in the same pattern until the simulation clock reach two sim hours. Curiously, the collective continues to find the most salient statements to be opinions on other agents, as Figure 7.2. It is interesting to note as well that emergent opinion leaders (eg: Agent23, Agent4) continue to lead. It is worthwhile to notice as well that the diversity of statements broadcast is reduced, as opinions harden.

7.3 Discrepancy between reality and simulation

A rally is evident in Figure 7.1 as all agents increase their sentiment towards the "Trump" entity following news of the strike. This corresponds to what would be expected in a normal situation, and vaguely corresponds to the actual reaction of the American public. Figure 7.3 shows the daily approval ratings of President Donald Trump sixty days prior and following the April 7, 2017 missile strike which this simulation sought to emulate. Although it may seem like there are dramatic rises and falls, there are is fact very little variation. Following the strike, the president does not gain even a point in the the FiveThirtyEight composite approval rating.

As Figure 7.1 shows, there is only a single agent that crosses from negative sentiment to positive sentiment following the announcement. This agent was near the threshold of approval (Baum, 2002) and was capable of reviewing its opinion from unfavourable to favourable. All the other agents *did* increase their approval of "Trump" and also displayed the behaviour predicted by Corollary 1. However, those in the cluster of population that held positive sentiment *already* approved and therefore their rise in sentiment could not contribute to an increase of overall approval. On the other hand, those in the cluster with negative sentiment were so far from the threshold of approval that despite an increase in sentiment, these agents still held a negative evaluation after the news. The net change in presidential approval is minimal, as it was in reality, but the simulation overestimates it. Prior to the news broadcast, there were seventeen agents that held positive sentiments; after

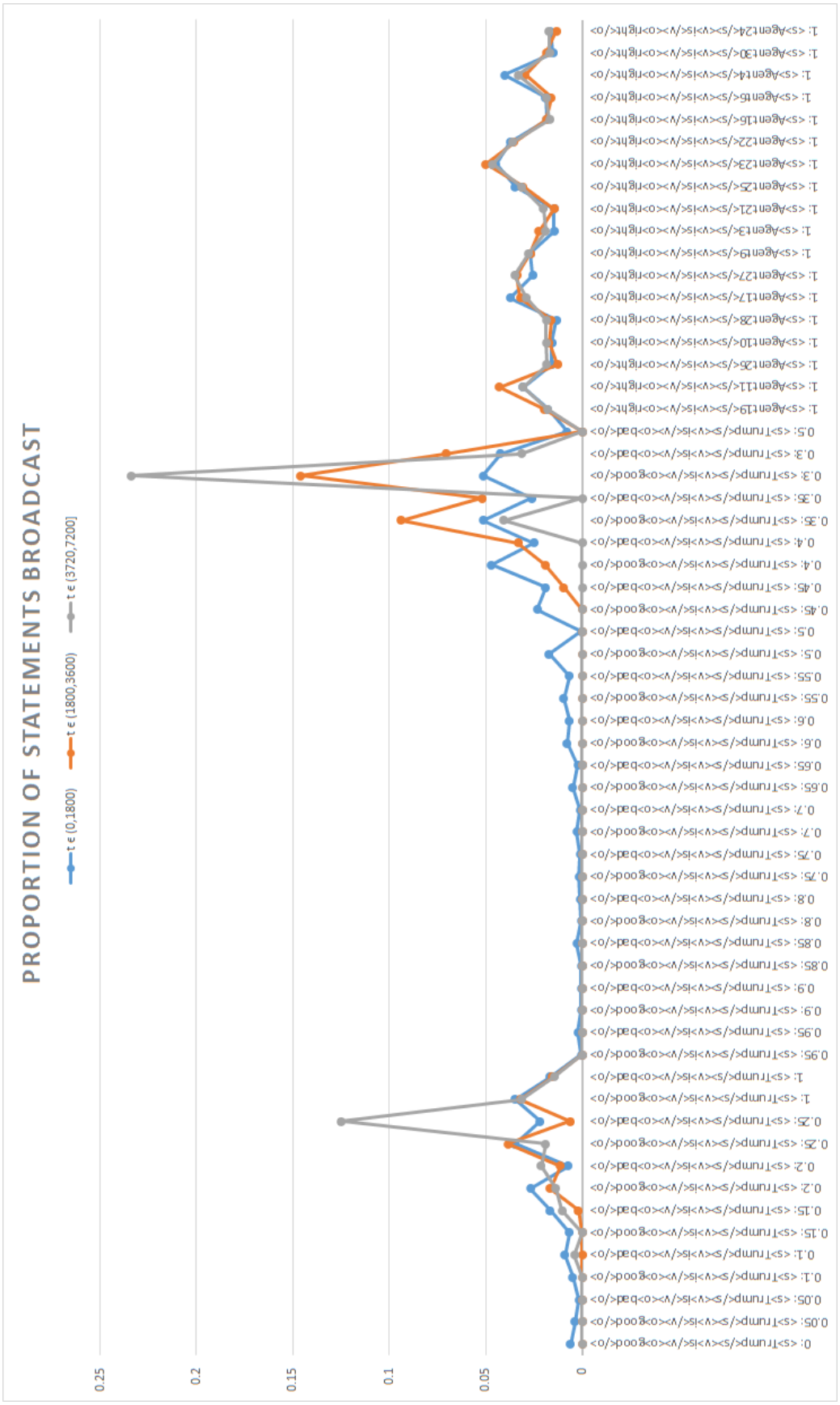


Figure 7.2: Proportion of statements broadcast by agents, depending on simulation time

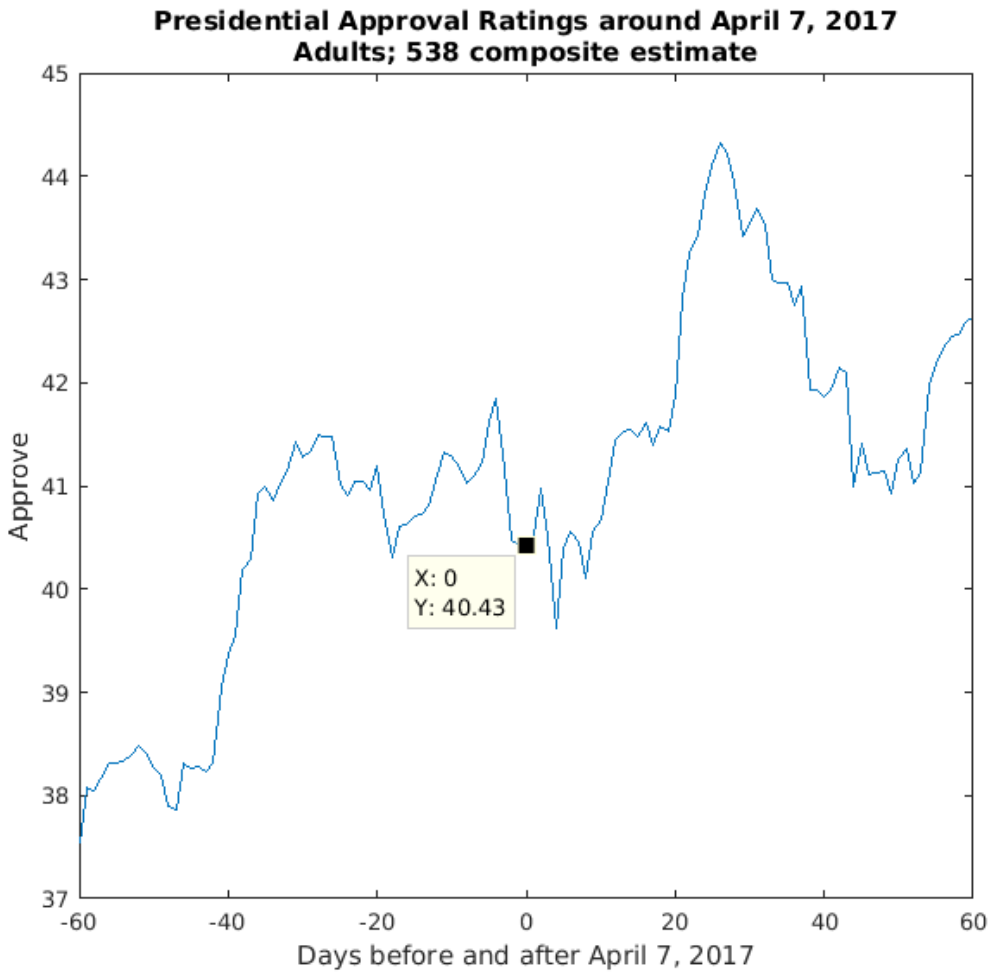


Figure 7.3: Donald Trump approval ratings 60 days before and after April 7, 2017 using FiveThirtyEight composite data

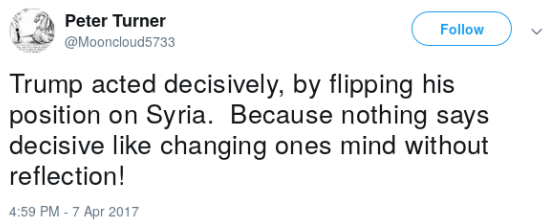
the broadcast there were eighteen. This represents an approval increase of nearly 6% which is about ten times greater than the actual magnitude of the change in approval ratings. This overestimation may be due to several factors.

One reason for the overestimation is simply the uniqueness of the simulation which is presently being discussed. Given the dependence on randomness, it is rather likely that another set of random numbers would produce a slightly different outcome. In order to properly estimate the magnitude of the simulated rally, many simulations should be conducted so that one could hope to apply the Law of Large Numbers and the paradigm of Monte Carlo methods. A second reason is the granularity of the simulation: there are only thirty agents (excluding the media agent). A single agent therefore represents over 3% of

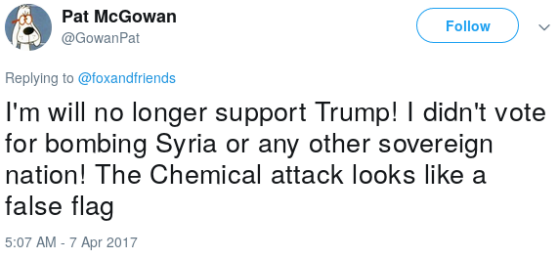
the population and therefore, a single agent changing from an unfavourable to favourable opinion is still greater than the actual change observed in the American public. In order to have a fine data to hope to reproduce the minuscule change in presidential ratings actually observed, a collective in excess of a thousand agents would be required so that each represents no more than 0.1% of the population. I had insufficient time and computational resources to simultaneously address the granularity of representation along with the first issue of small sampling in order to derive more conclusive results.

However, I would warn that it is entirely possible that other considerations come into play to mitigate the capacity of agents to rally. In particular, it should be noted that Mr. Trump ran an election on an isolationist policy and a number of his supporters were disappointed of his foreign intervention. Moreover, a substantial portion of Trump's supporters "like officials who stick to their positions" (Pew Research Center, 2017, Q51mm). Figure 7.4 shows a sample of both these reactions due to characteristics[13]. These were compounded by negative evaluations of the strike by prominent media personalities, like Ann Coulter. On the other hand, there were also a number of people expressing disengagement, with tinged with fears of (terrorist) reprisal. Figure 7.5 shows this. And although the political elite generally praised the president for his foreign intervention, see Figure 7.6, some attempted to reframe the discussion, like Rep. Karen Bass seeking to re-orient the conversation to fighting famine in Yemen in Figure 7.7. And then there are actors actively broadcasting disinformation, attempting to erode the credibility of channels like in Figure 7.8. All these bring an incredible amount of richness to the political discussions that occurred immediately following the strike. Incorporating this rich diversity, however, requires both a substantial amount of work to code the natural language (and images and videos) into the the structured language required by the model[14] and in refining the programmed agent behaviours which, for the purposes of the demonstration were kept rather simplistic.[15]

[13] See Section 3.3
[14] See Section 2.5
[15] See Section 6.4.1

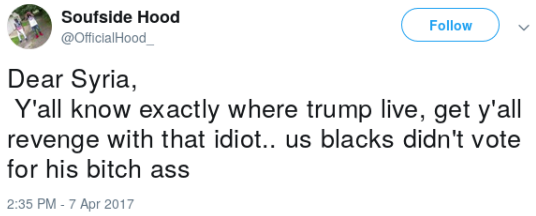


(a) Member of the public expressing disapproval that Trump changed his position Syria, ordering the intervention

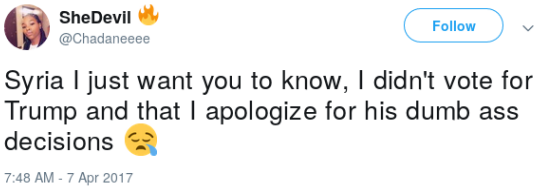


(b) An isolationist Trump supporter who will no longer support him because of the strike

Figure 7.4: Sample tweets from probable supporters of Donald Trump expressing their disapproval

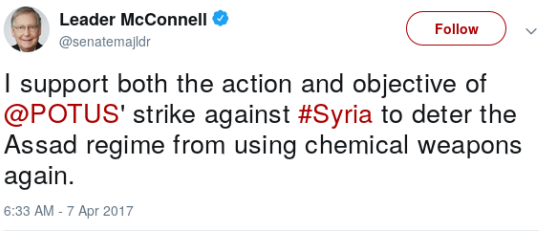


(a) Twitter user removing himself from the political discussion on the strike while reaffirming his negative sentiment and beliefs in the racial composition of Trump supporters

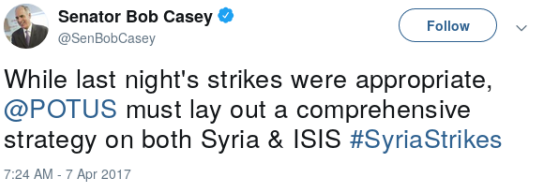


(b) Twitter user apologizing for the strike, reviewing sentiment downward if not disengaging from the political discussion

Figure 7.5: Sample tweets from people disengaging

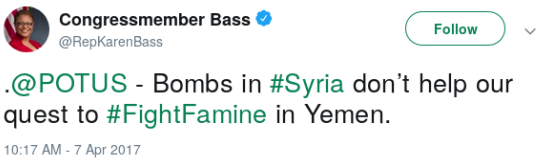


(a) Senator of presidential party praising the president

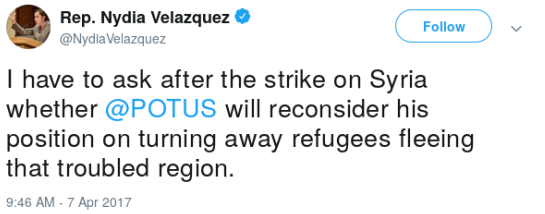


(b) Senator from non-presidential party praising the president

Figure 7.6: Sample tweets from Members of Congress praising the president



(a) Representative Karen Bass attempting to redirect the discourse to fighting famine in Yemen



(b) Representative Nydia Velazquez attempting to reframe discussion on the strike in terms of refugees

Figure 7.7: Sample tweets from Members of Congress attempting to reorient the discourse



Figure 7.8: A twitter user attempting to reduce the credibility of a channel

8 Future Studies and Conclusion

8.1 American Kulturkampf

The social network of the American public has gone through a radical reorganization in the last two decades with the advent of social media. Prior to the early 2000's, geographically isolated groups had very few interactions. In order to obtain information about geographically-distant places, members of the public would turn to media organizations. However, as internet technologies proliferated and evolved, members of the public began regularly interacting with geographically-distant individuals and niche news organizations sprung up, fragmenting the news media landscape. In the network terminology employed by Centola and Macy (2007), new bridges were created, existing bridges were broadened, and new nodes were added to the network. Scholars now recognize the fragmented nature of the news landscape and the importance of non-traditional sources of information (eg: Baum and Potter (2019), Mercier et al. (2018)). The added connectivity has aided the rate and breadth that information spreads in the social network of the American public. In contrast to information which may be seen to spread as a simple contagion, attitudes are more likely to spread as complex contagions. I think it is worthwhile to study whether the current hostile climate of political discourse could have developed from 2004 to today if the American public's social network had remained constant. In other words, I think it would be worthwhile to examine whether the increasing polarization of attitudes is due to the events that occurred, whether it is simply due to the reorganization of the social network, or whether it is caused by the events *and* the transformed topology.

I am also curious to investigate the dynamics of attitudes during the transformational phase – during the adoption of the new communication technologies. Different population segments of the American public were connected to the internet and through social media at

differing rates and it would be unsurprising to find that geography and innate characteristics played significant factors in the diffusion of the innovation. These factors are likely to be tied to political attitudes. I feel it would be interesting to test the attitude dynamics of a heterogeneously connected American public, and in particular whether the difference between those who had adopted the innovation and those who hadn't yield an emergent hardening and divergence of political attitudes as traditional media outlets and political elites commented on the innovation. I conjecture that the emergence of hacktivist groups like Anonymous, the slew of suits by the Recording Industry Association of American (RIAA) and Motion Picture Association of America (MPAA) on filesharing (Delgado, 2004), and terroristic headlines like 'Net blamed for rise in child porn'(BBC News, 2004) led to the formation of negative attitudes that were not just towards the technology but also towards adopters *and their beliefs*. I feel it would be worthwhile to also study whether those who rejected or delayed adoption of the new technology held significantly different and negative sentiment towards President Obama, given his famously extensive use of new technologies in his first bid for the presidency (Harfoush, 2009). If this is indeed the case, it may provide new insights into a techno-skeptical constituency with strong (political) attitudes whose adoption of the innovation was delayed, or perhaps never occurred.

Furthermore, different population segments are likely to have different susceptibilities to believing and sharing (re-broadcasting) information that is not true. The spread of fake news has recently received a substantial amount of consideration (eg: Vosoughi et al. (2018), Pennycook et al. (2018)). Grinberg et al. (2019) found correlations with age and political attitudes while others find that the acceptance of fake news is due to psychological make-up of an individual(Pennycook and Rand, 2018). An alternative hypothesis is that credulousness is inversely related to experience with virtual social interactions. Under this last hypothesis, credulousness would be tied to the time since the adoption of the technological innovation. Virtual socialization is due to behaviour and its uptake is likely to diffuse as a complex contagion through *real* social network ties – since by definition there would be no virtual connections between individuals who haven't adopted the innovation. Given the important role that homophily plays in social network ties, if this alternate hypothesis is true, then credulous population segments with similar attitudes, like in Grinberg et al.

(2019), would be *caused* by the technological diffusion. If this hypothesis is indeed true, then it may also inform the individual empowerment and educational recommendations of Lazer et al. (2018).

To many observers the current political discourse seems to be beyond the ordinary ideological divide that has characterized the nation since its inception. It appears that today the polarization transcends ideology to take root in *identity*, making discussion between people in different opinion clusters difficult and tense. Pew Research Center (2017) found that a substantial number of people would reevaluate their friendship upon discovering that their friend voted one way or the other (QB55a, QB55b). Moreover, in all the population clusters, a majority find it "stressful and frustrating to talk politics with people who have a different opinion of Trump" (QB28). Adding to this are professionals and amateurs who take pride in creating controversy and or outrage, like Katie McHugh[1] or @__ribbit, a now deleted Twitter account. Appendix B.2 provides a guide to aspiring conservative trolls. Similar personalities exist on the progressive end of the spectrum. I feel that it would be useful to estimate the effect of online trolling on attitude dynamics.

I would also like to test the influence of the terror attacks of September 11, 2001 on driving the divergence of attitudes in the American public. The shock to views on Muslims in the American public may have provided the essential ingredient to the growing divide. While some sub-groups of the American public were regularly exposed to moderate, non-radical Muslims through their every day lives, others, informationally-isolated were likely to only have been exposed to extreme and violent depictions of Islam and Muslims. Jesse Morton, also known as Younus Abdullah Muhammad, an Al-Qaeda recruiter, created the concept of a Jihadist electronic magazine in 2009. In it he and his fellow jihadists wrote in fluent English, calling for the murder of infidels and providing instructions on bomb-making. His magazine inspired the Tsarnaev brothers who bombed the Boston Marathon in 2013 and his magazine inspired the use of vehicles to terrorize civilians.[2] His magazine was distributed online, long before ISIS posted videos on YouTube of executions in the desert. While some segments of the American public lived side-by-side and had bonds of friendship with ordinary Muslims,

[1] `https://www.buzzfeednews.com/article/rosiegray/katie-mchugh` Accessed 2019-06-10

[2] `https://www.wsj.com/articles/i-invented-the-jihadist-journal-11559602751` Accessed 2019-06-10

other segments of the public lived in relatively homogeneous communities and only came to associate Islam and Muslims to the horrors amplified by the media. That this continues to be the case is made evident by the polarization of responses to the question on whether or not Islam is a religion that encourages violence (Pew Research Center, 2017, QA 143). This could be interpreted as a proxy for sentiment towards the Islamic religion and Muslims. As these groups with differential exposure to information *and* attitudes began meeting virtually, I surmise that they rejected each other's views and reinforced their own. I wonder to what extent the Afghanistan and Iraq wars, terror attacks, and messaging regarding Muslims influenced the polarization of the American public – which was already split on issues like race, abortion, and same-sex marriage – as its social network transformed.

8.2 Conclusion

This thesis established a new method to model information for multi-agent political simulations. It relies on an extensible structured language where declarative sentences – statements – are the smallest unit of information. This work can be interpreted as blueprints for building a sophisticated simulator for the purposes of studying how information and attitudes diffuse in a social network. In this work I show that by modelling information as I propose, empirically observed attitude dynamics are reproduced due to how knowledge is constructed. I do this through both rigorous mathematical proofs and a proof-of-concept simulation which shows a minimal working model. In this proof-of-concept simulation, I was able to reproduce the expected surge in positive sentiment towards the President of the United States following his initiation of a foreign military action. There are, however, a number of unimplemented cognitive and behavioural agent features which I nonetheless strive to describe in thorough detail. I also address the available data, which is crucial to calibrate the synthetic population, and the additional data that would be useful to improve the accuracy of this calibration. Lastly, I propose future studies that requires a more sophisticated implementation of the simulator which aim to understand and describe the root causes of the currently diverging attitude clusters in the American public.

III Appendices

A Code snippets of implementation

A.1 Knowledge retrieval query

```
/* select all broadcasts in memory which contain statements on a subject */
(select
       subject_id, verb_id,object_id, verb_intensity, credibility,
              agent_knowledge.entity_id, agent_knowledge.mem_remaining,
              agent_knowledge.timestamp
from statements join (broadcasts, entities, agent_knowledge) on
       (statements.entity_id = broadcasts.message_id and
       broadcasts.entity_id = entities.entity_id and
       agent_knowledge.entity_id = entities.entity_id)
where entities.entity_type = 'broadcast' and
       agent_knowledge.agent_id = ? and
       statements.subject_id = ? and
       agent_knowledge.mem_remaining > 50)
union /* and all statements in memory on a subject */
(select
       subject_id, verb_id,object_id, verb_intensity, credibility,
              agent_knowledge.entity_id, agent_knowledge.mem_remaining,
              agent_knowledge.timestamp
from statements join (entities, agent_knowledge) on
       (statements.entity_id = entities.entity_id and
       agent_knowledge.entity_id = entities.entity_id and
       agent_knowledge.agent_id = ? and
       statements.subject_id = ? and
       agent_knowledge.mem_remaining > 50)
where entities.entity_type = 'statement');
```

A.2 What to broadcast

```
/* choose the most appropriate good/bad statement on a subject */
(select
       entity_id
from statements
where statements.subject_id = ? and
       statements.object_id in (-1,1)
order by
       abs(statements.verb_intensity*object_id/abs(object_id) - ?) asc limit 1)
union /* and choose the most salient statement in memory */
(select
       agent_knowledge.entity_id
from agent_knowledge join statements on
       agent_knowledge.entity_id = statements.entity_id
where agent_knowledge.agent_id = ?
order by
       credibility*abs(mem_remaining/100*statements.verb_intensity) desc,
       timestamp asc limit 1);
```

A.3 Data file for proof-of-concept simulation

```
<data>
<!-- verbs -->
<verb id="1">is</verb>
<verb id = "6" hurt="100">bomb</verb>
<verb id = "8" heal="10">helped</verb>
<verb id = "3">know</verb>

<!-- entities -->
<entity id="1">good</entity>
<entity id="-1">bad</entity>
<entity id="2">important</entity>
<entity id="-2">unimportant</entity>
<entity id="3">right</entity>
<entity id="-3">wrong</entity>

<!-- base statements  -->
<statement intensity = "0"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.05"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.05"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.10"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.10"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.15"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.15"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.20"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.20"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.25"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.25"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.30"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.30"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.35"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.35"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.40"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.40"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.45"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.45"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.50"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.50"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.55"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.55"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.60"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.60"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.65"><s>Trump</s><v>is</v><o>good</o></statement>
```

```
<statement intensity = "0.65"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.70"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.70"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.75"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.75"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.80"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.80"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.85"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.85"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.90"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.90"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "0.95"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "0.95"><s>Trump</s><v>is</v><o>bad</o></statement>
<statement intensity = "1.00"><s>Trump</s><v>is</v><o>good</o></statement>
<statement intensity = "1.00"><s>Trump</s><v>is</v><o>bad</o></statement>


<!-- agent_knowledge -->
<agent_knowledge agent_id = "1" credibility="1">
<statement intensity="1"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="2" credibility = "1">
<statement intensity="1"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "3" credibility="1">
<statement intensity=".95"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="4" credibility = "1">
<statement intensity=".95"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "5" credibility="1">
<statement intensity=".90"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="6" credibility = "1">
<statement intensity=".90"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "7" credibility="1">
<statement intensity=".85"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="8" credibility = "1">
<statement intensity=".85"><s>Trump</s><v>is</v><o>bad</o></statement>
```

```
</agent_knowledge>

<agent_knowledge agent_id = "9" credibility="1">
<statement intensity=".80"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="10" credibility = "1">
<statement intensity=".80"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "11" credibility="1">
<statement intensity=".75"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="12" credibility = "1">
<statement intensity=".75"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "13" credibility="1">
<statement intensity=".70"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="14" credibility = "1">
<statement intensity=".70"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "15" credibility="1">
<statement intensity=".65"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="16" credibility = "1">
<statement intensity=".65"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "17" credibility="1">
<statement intensity=".60"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="18" credibility = "1">
<statement intensity=".60"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "19" credibility="1">
<statement intensity=".55"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="20" credibility = "1">
<statement intensity=".55"><s>Trump</s><v>is</v><o>bad</o></statement>
```

```
</agent_knowledge>

<agent_knowledge agent_id = "21" credibility="1">
<statement intensity=".50"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="22" credibility = "1">
<statement intensity=".50"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "23" credibility="1">
<statement intensity=".45"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="24" credibility = "1">
<statement intensity=".45"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "25" credibility="1">
<statement intensity=".40"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="26" credibility = "1">
<statement intensity=".40"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "27" credibility="1">
<statement intensity=".35"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="28" credibility = "1">
<statement intensity=".35"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<agent_knowledge agent_id = "29" credibility="1">
<statement intensity=".30"><s>Trump</s><v>is</v><o>good</o></statement>
</agent_knowledge>

<agent_knowledge agent_id ="30" credibility = "1">
<statement intensity=".30"><s>Trump</s><v>is</v><o>bad</o></statement>
</agent_knowledge>

<!-- media agents -->
<agent_knowledge agent_id ="31" credibility = "1">
       <statement intensity=".8" mem_decay = "0">
              <s>Syria</s><v>is</v><o>bad</o>
       </statement>
       <statement intensity=".05" mem_decay = "0">
```

```
                <s>Trump</s><v>bomb</v><o>Syria</o>
        </statement>
</agent_knowledge>

</data>
```

A.4 Loader for simulation data

```
import lxml.etree
import mysql.connector
import math

mydb = mysql.connector.connect(
  host="server",
  user="username",
  passwd="passwork",
  database="newsMaker"
)

dbcursor = mydb.cursor(buffered = True)
params = dict();

def parse_statement_by_id(subject_id, verb_id, obj_id, intensity = 1):
        #WARNING: entities defined by the subject/verb/obj are assumed to
            exist!

        # check if statement exists
        dbcursor.execute("select statement_id, entity_id from statements
            where subject_id = %d and verb_id = %d and object_id = %d and
            verb_intensity = %f" %(int(subject_id), int(verb_id), int(obj_id
            ), float(intensity)))

        # get names for statement plain_text
        dbcursor.execute("select entity_text from entities where entity_id =
             %d" % subject_id)
        subject = dbcursor.fetchone()[0]
        dbcursor.execute("select verb_text from verbs where verb_id = %d" %
            verb_id)
        verb = dbcursor.fetchone()[0]
        dbcursor.execute("select entity_text from entities where entity_id =
             %d" % obj_id)
        obj = dbcursor.fetchone()[0]

        stmt = '<s>'+subject+'</s><v>'+verb+'</v><o>'+obj+'</o>'

        result = dbcursor.fetchall()
        if len(result) == 0:
                dbcursor.execute("insert into entities values (0, '%s', '
                    statement')" % stmt)
                statement_entity_id = dbcursor.lastrowid
                print("statement as entity id = %d" % statement_entity_id)
                dbcursor.execute("insert into statements (statement_id,
```

```
                    entity_id, subject_id, verb_id, object_id, verb_intensity
                    , plain_text) values (0,%d,%d,%d,%d,%d, '%s')" % (
                    statement_entity_id, subject_id, verb_id, obj_id,
                    intensity,stmt))
        else:
                print("%s (intensity = %d) exists as ent:%d (stmt:%d)" %(stmt
                    , intensity, result[0][1],result[0][0]))
                statement_entity_id = result[0][1]



        return stmt, statement_entity_id

def parse_statement( node ):
        #
        if node is None:
                return

# print('Begin statement')
        subject = ''
        obj = ''
        verb = ''
        stmt = ''
        for child in node.iterchildren():

                ## SUBJECT
                if child.tag == 's':
                        if child.text is None:
                                for dchild in child.iterchildren():
                                        subject = parse_statement(dchild)[0]
                        else:
                                subject = child.text.strip()

                        # is subject known as an entity? otherwise insert as
                            entity

                        #print("subject = %s" % subject)
                        dbcursor.execute("select entity_id from entities
                            where entity_text like '%s'" % subject)
                        result = dbcursor.fetchall()
                        if len(result) == 0:
                                print("Subject %s not found! Inserting..." %
                                    subject)
                                dbcursor.execute("insert into entities values
                                    (0, '%s', 'entity')" % (subject))
                                subject_id = dbcursor.lastrowid
                        else:
                                subject_id = result[0][0] # get first (>1 set
```

```
                    is db error)
            #print(subject_id)

    ## VERB
    if child.tag == 'v':
            #print('Verb : %s' % child.text)
            verb = child.text.strip()
            # is verb known? otherwise insert as verb

            #print("verb = %s" % verb)
            dbcursor.execute("select verb_id from verbs where
                verb_text like '%s'" % verb)
            result = dbcursor.fetchall()
            if len(result) == 0:
                    print("Verb %s not found! Inserting..." % verb
                        )
                    dbcursor.execute("insert into verbs values (0,
                         '%s')" % (verb))
                    verb_id = dbcursor.lastrowid
            else:
                    verb_id = result[0][0] # get first
            #print(verb_id)

    ## OBJECT
    if child.tag == 'o':
            if child.text is None:
                    for dchild in child.iterchildren():
                            obj += parse_statement(dchild)[0]
            else:
                    obj = child.text.strip()

            # is obj known as an entity? otherwise insert as
                entity

            #print("object = %s" % obj)
            dbcursor.execute("select entity_id from entities
                where entity_text like '%s'" % obj)
            result = dbcursor.fetchall()
            if len(result) == 0:
                    print("Object %s not found! Inserting..." %
                        obj)
                    dbcursor.execute("insert into entities values
                        (0, '%s', 'entity')" % (obj))
                    object_id = dbcursor.lastrowid
            else:
                    object_id = result[0][0] # get first
            #print(object_id)
```

```
## STATEMENT
stmt = '<s>'+subject+'</s><v>'+verb+'</v><o>'+obj+'</o>'
# get intensity of statment. If none defined, default to 1
intensity = float(node.get('intensity',1))
# is statement known as an entity? Otherwise insert into entities
    and statements
dbcursor.execute("select statement_id, entity_id, verb_intensity
    from statements where " "statements.subject_id = %d and "
"statements.verb_id = %d and "
"statements.object_id = %d and "
"statements.verb_intensity > %f - 0.01 and statements.verb_intensity
     > %f + 0.01"
 % (subject_id, verb_id, object_id, intensity, intensity))
result = dbcursor.fetchall()
if len(result) == 0:
        dbcursor.execute("insert into entities values (0, '%s', '
            statement')" % stmt)
        statement_entity_id = dbcursor.lastrowid
        print("statement as entity id = %d" % statement_entity_id)
        dbcursor.execute("insert into statements (statement_id,
            entity_id, subject_id, verb_id, object_id, verb_intensity
            , plain_text) values (0,%d,%d,%d,%d,%f, '%s')" % (
            statement_entity_id, subject_id, verb_id, object_id,
            intensity,stmt))
else:
        # check if intensity returned is almost equal, otherwise
            insert
        for res in result:
                if abs(res[2] - intensity) < 0.001:
                        # equal return
                        print("%s (intensity = %f (got %f)) exists as
                            ent:%d (stmt:%d)" %(stmt, intensity, res
                            [2], result[0][1],result[0][0]))
                        statement_entity_id = result[0][1]
                        return stmt, statement_entity_id

        # not found stmt with similar intensity in results, insert a
            new one
        dbcursor.execute("insert into entities values (0, '%s', '
            statement')" % stmt)
        statement_entity_id = dbcursor.lastrowid
        print("statement as entity id = %d" % statement_entity_id)
        dbcursor.execute("insert into statements (statement_id,
            entity_id, subject_id, verb_id, object_id, verb_intensity
            , plain_text) values (0,%d,%d,%d,%d,%f, '%s')" % (
            statement_entity_id, subject_id, verb_id, object_id,
            intensity,stmt))
```

```
        return stmt, statement_entity_id


def parse_broadcast( node ):
        print('\nParsing broadcast')

        # get attributes and use them as defaults for deeper nodes
        # <broadcast attribs> <broadcast>some other </broadcast></broadcast>
        print(node.attrib)

        message_pos = 0 # the counter for the payload (ie: statement/
            broadcast) contained in the broadcast
        message_id = 0 # the payload's entity_id

        source = node.get('source', 'UNKNOWN AGENT').strip()

        # get source_id; if source not exist as entity, insert
        dbcursor.execute("select entity_id from entities where entity_text
            like '%s'" % source)
        result = dbcursor.fetchone()
        if result is None:
                print("didn't find entity %s,(source) inserting" % source)
                dbcursor.execute("insert into entities values(0, '%s', 'agent
                    ')" % source)
                source_id = dbcursor.lastrowid
        else:
                source_id = result[0]

        # get channel_id; if not exist as entity, insert
        channel = node.get('channel','UNKNOWN CHANNEL').strip()
        dbcursor.execute("select entity_id from entities where entity_text
            like '%s'" % channel)
        result = dbcursor.fetchone()
        if result is None:
                print("didn't find entity %s, (channel) inserting" % channel)
                dbcursor.execute("insert into entities values(0, '%s', '
                    entity')" % channel)
                channel_id = dbcursor.lastrowid
        else:
                channel_id = result[0]

        target = int(node.get('target', 0)) # if target is 0, it is a public
             broadcast
        time_created = int(float(node.get('time_created', 0))) # if no time,
             broadcast was created at the beginning of time
        print("Time = %d" % time_created)
```

```
# parse the title
title = node.get('title', 'UNKNOWN TITLE').strip()
if title not in 'UNKOWN TITLE' and 'UNKNOWN TITLE' not in title:
        title_id = parse_statement(lxml.etree.fromstring(node.get('
            title')))[1]
else:
        dbcursor.execute("select entity_id from entities where
            entity_text like '%s'" % title)
        result = dbcursor.fetchone()
        if result is None:
                # first time encountering ?, add it as an entity
                print("didn't find entity %s, (title) inserting" %
                    title)
                dbcursor.execute("insert into entities values (0, '%s
                    ', 'broadcast')" % title)
                title_id = dbcursor.lastrowid
        else:
                title_id = result[0]



# get location_id; if not exist as entity, insert
location = node.get('location', 'UNKNOWN LOCATION').strip()
dbcursor.execute("select entity_id from entities where entity_text
    like '%s'" % location)
result = dbcursor.fetchone()
if result is None:
        print("didn't find entity %s (location), inserting" %
            location)
        dbcursor.execute("insert into entities values(0, '%s', '
            location')" % location)
        location_id = dbcursor.lastrowid
else:
        location_id = result[0]

# check if broadcast exists based on source,channel,target,
    time_created, title and location. If it does, get the broadcast/
    entity ids. Otherwise, get max broadcast id, increment (no auto
    increment in table because of n:m) and insert into entities.
#dbcursor.execute("select broadcast_id, entity_id from broadcasts
    where source = %d and channel = %d and target = %d and
    time_created = %f and title_id = %d and location = %d" % (
    source_id, channel_id, target, time_created, title_id,
    location_id))
dbcursor.execute("select broadcast_id, entity_id from broadcasts
    where source = %d and channel = %d and time_created = %f" % (
    source_id, channel_id, time_created))
```

```
result = dbcursor.fetchone()
if result is not None:
        print("Broadcast exists; will append...")
        broadcast_id = result[0]
        broadcast_entity_id = result[1]
        # broadcast exists, get last message position and update the
            counter
        dbcursor.execute("select max(message_pos) from broadcasts
            where broadcast_id = %d" % broadcast_id)
        result = dbcursor.fetchone()
        message_pos = result[0]
else:
        # create new broadcast entity
        dbcursor.execute("insert into entities values(0, '%s', '
            broadcast')" % (title))
        broadcast_entity_id = dbcursor.lastrowid
        dbcursor.execute("select max(broadcast_id) from broadcasts")
        result = dbcursor.fetchone()
        if result[0] is None:
                # no broadcasts exist, this is the first.
                broadcast_id = 1
        else:
                broadcast_id = result[0]+1
        print("Broadcast does not exist; will create with id %d" %
            broadcast_id)
## PARSE CONTENTS OF BROADCAST

for child in node.iterchildren():
        message_pos+=1
        ## STATEMENTS
        if child.tag == 'statement':
                statement_entity_id = parse_statement(child)[1]
                # insert the statement as a message in the broadcast
                dbcursor.execute("insert into broadcasts (
                    broadcast_id, entity_id, message_id, message_pos,
                    source, channel, target, time_created, title,
                    location) values (%d, %d, %d, %d, %d, %d, %d, %d,
                    '%s', %d)" % (broadcast_id, broadcast_entity_id,
                    statement_entity_id, message_pos, source_id,
                    channel_id, target, time_created, title,
                    location_id))

        ## BROADCASTS
        if child.tag == 'broadcast':
                # inherit tags if not defined in child
                child_source = child.get('source')
                if child_source is None:
                        child.set('source', source)
```

```
                        else:
                                child.set('source', child_source.strip())

                        child_channel = child.get('channel')
                        if child_channel is None:
                                child.set('channel', channel)
                        else:
                                child.set('channel', child_channel.strip())

                        child_target = child.get('target')
                        if child_target is None:
                                child.set('target', str(target))

                        child_time_created = child.get('time_created')
                        if child_time_created is None:
                                child.set('time_created', str(time_created))

                        # don't inherit time updated
                        # don't inherit title

                        child_location = child.get('location')
                        if child_location is None:
                                child.set('location', location)
                        else:
                                child.set('location', child_location.strip())

                        child_broadcast_entity_id = parse_broadcast(child)[1]

                        dbcursor.execute("insert into broadcasts (
                            broadcast_id, entity_id, message_id, message_pos,
                            source, channel, target, time_created, title,
                            location) values (%d, %d, %d, %d, %d, %d, %d, %d,
                            '%s', %d)" % (broadcast_id, broadcast_entity_id,
                            child_broadcast_entity_id, message_pos, source_id,
                             channel_id, target, time_created, title,
                            location_id))

        return broadcast_id, broadcast_entity_id


def parse_verb(node):
        # get attributes
        verb_id = int(node.get('id','0'))
        # check availability
        dbcursor.execute("select verb_id from verbs where verb_id = %d" %
            verb_id)
        if dbcursor.fetchone() is not None:
                # id is already taken, default to auto_increment
```

```
                verb_id = 0;
                print("Requested verb id is already taken")
        verb_text = node.text.strip()
        # if not exist, insert
        dbcursor.execute("select verb_id from verbs where verb_text like '%s
            '" % verb_text)
        result = dbcursor.fetchone()
        if result is None:
                print("Inserting new verb %s" % verb_text)
                dbcursor.execute("insert into verbs values(%d, '%s')" % (
                    verb_id, verb_text))
        else:
                print("Verb '%s' already exists as verb_id: %d" % (verb_text,
                     result[0]))
        return

def parse_entity(node):
        # get attributes
        entity_id = int(node.get('id','0'))

        # check availability
        dbcursor.execute("select entity_id from entities where entity_id = %
            d" % entity_id)

        if dbcursor.fetchone() is not None:
                # id is already taken, default to auto_increment
                entity_id = 0;
                print("Requested entity id is already taken")

        entity_text = node.text.strip()
        # if not exist, insert
        dbcursor.execute("select entity_id from entities where entity_text
            like '%s'" % entity_text)
        result = dbcursor.fetchone()
        if result is None:
                print("Inserting new entity %s" % entity_text)
                dbcursor.execute("insert into entities values(%d, '%s', '
                    entity')" % (entity_id, entity_text))
                entity_id = dbcursor.lastrowid
        else:
                entity_id = result[0]
                print("entity %s already exists as entity_id: %d" % (
                    entity_text, entity_id))
        return entity_id

def parse_agent_knowledge(node):
        # always add to agent knowledge SO BE CAREFUL WITH DUPLICATES!
```

```
# GET ATTRIBUTES OF KNOWLEDGE

# agent_id
# get agent id. Remember, agent_id is an entity_id (unlike eg:
    broadcast_id)

# check if there is an agent name, if there is, use that to search
    for the entity associate to that name; otherwise its a generic
    agent with name like Agent123
agent_name = node.get('agent_name')
if agent_name is None:
        agent_name = 'Agent'+node.get('agent_id')
        # If no agent id, throw an error since otherwise who knows
            this piece of knowledge?

dbcursor.execute("Select entity_id from entities where entity_text
    like '%s'" % agent_name)
result = dbcursor.fetchone()
if result is None:
        # agent doesn't exist, create it as an entity
        print("Encountered new agent; adding it as an entity")
        dbcursor.execute("insert into entities values(0, '%s', 'agent
            ')" % agent_name)
        agent_id = dbcursor.lastrowid
else:
        agent_id = int(result[0])

## preparing for insert
credibility = float(node.get('credibility', float(params.get('
    credibility'))))
mem_remaining = float(node.get('mem_remaining',float(params.get('
    mem_remaining'))))
mem_decay = float(node.get('mem_decay', -math.log(float(params.get('
    mem_remaining_after_news_cycle')))/(60*60*float(params.get('
    news_cycle_length')))))

# parse what's in the knowledge, inserting along the way

for child in node.iterchildren():
        # ADD INHERITABLE ATTRIBUTES IF NOT IN CHILD TAG (regardless
            of child.tag type) (ASSUMES THAT AN agent_knowledge CHILD
             WILL HAVE ID)

        child.set('credibility', (child.get('credibility', str(
            credibility))))
        child.set('mem_remaining', (child.get('mem_remaining', str(
            mem_remaining))))
        child.set('mem_decay', (child.get('mem_decay', str(mem_decay)
```

```
                    )))

                known_id = node_switcher(child)
                dbcursor.execute("insert into agent_knowledge (agent_id,
                    entity_id, credibility, mem_remaining, mem_decay) values
                    (%d, %d,%f,%f,%f)" % (agent_id, known_id, credibility,
                    mem_remaining, mem_decay))
                # WHAT AN AGENT KNOWS IS ALSO A PIECE OF KNOWLEDGE!
                # create stmt: "Agentx know _", insert into statements AND
                    knowing agent (I know that I know X; Alice know (Bob know
                     X))
                dbcursor.execute("Select verb_id from verbs where verb_text
                    like 'know'")
                verb_id = dbcursor.fetchone()[0]
                known_known_id = parse_statement_by_id(agent_id, verb_id,
                    known_id, credibility)[1]
                #known_known_id = parse_statement(lxml.etree.fromstring(stmt)
                    )[1]
                dbcursor.execute("insert into agent_knowledge (agent_id,
                    entity_id, credibility, mem_remaining, mem_decay) values
                    (%d, %d,%f,%f,%f)" % (agent_id, known_known_id,
                    credibility, mem_remaining, mem_decay))

                # return the id of the knowledge of the knowledge, as this
                    passes through in recursive knowledge (Alice know Bob
                    know Charlie know X)

        return known_known_id


def node_switcher(node):
        if node is None:
                return
        if node.tag == 'statement':
                #print('parsing statement')
                return parse_statement(node)[1]
#  print("statement id = %d" % parse_statement(node)[1])
        if node.tag == 'broadcast':
                #print('parsing broadcast')
                return parse_broadcast(node)[1]
        if node.tag == 'agent_knowledge':
                 print('adding to agent knowledge')
                 return parse_agent_knowledge(node)
        #if node.tag == 'verb':
                # TODO: for now assuming that verbs are universally known
                    with homogeneous definitions
                #print('parsing verb')
        # return parse_verb(node)
```

```
        if node.tag == 'entity':
                print('parsing entities')
                return parse_entity(node)

        return


## MAIN

def main(param):
        tree = lxml.etree.parse('data.xml', parser = lxml.etree.XMLParser(
            remove_comments=True, remove_blank_text=True))
        root = tree.getroot()
        global params
        params = param;

        #input("PRESS ENTER TO CONTINUE.")

        # insert verbs
        for node in root.findall('verb'):
                print('parsing verb')
                parse_verb(node)

        # insert base entities
        for node in root.findall('entity'):
                print('parsing entities')
                parse_entity(node)

        for node in root.iterchildren():
                node_switcher(node)


        dbcursor.close()
        mydb.close()
        return
```

B Broadcasts

B.1 The Washington Post, April 7, 2017. Page A1

The Washington Post

Prices may vary in areas outside metropolitan Washington.

SU V2

Cloudy 54/38 • Tomorrow: Sunny 63/44 B8

Democracy Dies in Darkness

FRIDAY, APRIL 7, 2017 · $2

U.S. strikes Syrian military airfield

What's next? The risks include wider war and Syrian, Russian responses.

BY GREG JAFFE

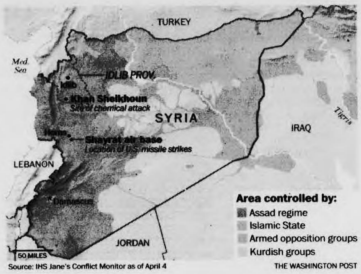
For President Trump, choosing targets and launching cruise missiles to punish the Syrian regime for using chemical weapons this week may have been a relatively clear-cut decision. The big problem is what comes next.

The military had been preparing options for a strike against President Bashar al-Assad since well before 2013, when the Syrian dictator killed more than 1,000 of his own people in a devastating nerve agent attack.

A chemical attack Tuesday blamed on the Assad regime killed scores of civilians and has triggered a response from the Pentagon, which launched more than 50 cruise missiles at a Syrian military airfield late on Thursday.

"The basic questions haven't changed," said Phil Gordon, a senior official in the Obama White House who took part in many earlier debates about how to punish Assad. "Is there a set of military strikes that you can use

STRIKES CONTINUED ON A13



Source: IHS Jane's Conflict Monitor as of April 4 THE WASHINGTON POST

Firing of cruise missiles comes days after chemical attack against civilians

BY DAN LAMOTHE, MISSY RYAN AND THOMAS GIBBONS-NEFF

The U.S. military launched 59 cruise missiles at a Syrian military airfield late Thursday, in the first direct American assault on the government of President Bashar al-Assad since that country's civil war began six years ago.

The operation, which the Trump administration authorized in retaliation for a chemical attack killing scores of civilians this week, dramatically expands U.S. military involvement in Syria and exposes the United States to heightened risk of direct confrontation with Russia and Iran, both backing Assad in his attempt to crush his opposition.

President Trump said the strike was in the "vital national security interest" of the United States and called on "all civilized nations to join us in seeking to end the slaughter and bloodshed in Syria. And also to end terrorism of all kinds and all types."

"We ask for God's wisdom as we

SYRIA CONTINUED ON A13

China's first couple visits Mar-a-Lago



JIM WATSON/AGENCE FRANCE-PRESSE VIA GETTY IMAGES

President Trump and first lady Melania Trump stand with Chinese President Xi Jinping and his wife, Peng Liyuan, upon their arrival at the Mar-a-Lago estate on Thursday, before the United States launched a missile strike in the evening on an air base in Syria. The two leaders were to continue their Florida summit on Friday. **Story, A8**

Senate burns bridge to clear Gorsuch's way

GOP KILLS FILIBUSTER FOR TOP COURT PICKS

'Nuclear' move threatens to stoke enmity in chamber

BY ED O'KEEFE AND SEAN SULLIVAN

Republicans on Thursday cleared the way for Judge Neil Gorsuch to be confirmed to the U.S. Supreme Court, overcoming a historic Democratic blockade by changing the rules of the U.S. Senate — a move that highlighted the fierce partisanship that has seized Congress.

The long-anticipated rules change now means that all presidential nominees for executive branch positions and the federal courts need only a simple-majority vote to be confirmed by senators.

The GOP decision to ram through the rules change is also likely to further divide an increasingly partisan Senate. Several senators openly fretted that eliminating the minority party's right to block high court nominees could lead to the end of filibusters on legislation — effectively transforming the Senate's traditional role in the legislative process as the slower, more deliberative chamber.

The ultimate confirmation of Gorsuch, which is expected sometime on Friday, represents a major victory, however, for President Trump as well as for Senate Majority Leader Mitch McConnell, who refused to even consider President Barack Obama's nominee after Justice Antonin Scalia died in February 2016.

The unflinching discipline displayed by McConnell (R-Ky.) in rejecting pleas from Democrats

GORSUCH CONTINUED ON A6

The end of a fight, or the start?
Some feel Senate leaders can still find a way to work together. **A8**

Contentious path to the court
See how the Senate's Gorsuch vote compares with past nominees'. **A6**

More money, more problems: D.C. juggles budget surplus

BY PETER JAMISON

The nation's capital is grappling with a problem inconceivable in the not-long-distant days when the city was a byword for fiscal recklessness: It has too much money.

Flush with carefully stewarded tax dollars from almost two decades of uninterrupted growth, the District today is unrecognizable as the bureaucracy that was hurtling toward financial ruin when it was rescued by Congress in the 1990s.

But many are questioning whether the city has learned too well the lessons of its humiliating stint under a federal financial control board. They say D.C. officials are now hoarding money better spent on social ills, such as an alarmingly high rate of homelessness, that continue to mar the city's much-celebrated revival.

The debate intensified this week, as Mayor Muriel E. Bowser (D) delivered a $13.8 billion budget to the D.C. Council that proposes only modest spending increases for affordable housing, homeless services and education. Meanwhile, the budget would continue to build up reserves that have grown to $2.4 billion.

That amount — about enough to run the city of San Antonio for a year — is more than triple the District's deficit when then-Mayor Marion Barry Jr. (D) went

FINANCES CONTINUED ON A16

Nunes recuses himself from committee's probe of Russia

BY KAROUN DEMIRJIAN

House Intelligence Committee Chairman Devin Nunes (R-Calif.) temporarily stepped aside Thursday from the committee's probe into Russian interference in the presidential election, as House investigators look into ethics charges against him.

The House Ethics Committee released a statement Thursday saying it had "determined to investigate" allegations that "Nunes may have made unauthorized disclosures of classified information, in violation of House Rules, law, regulations, or other standards of conduct."

Nunes has come under fire in recent weeks for speaking publicly about classified foreign surveillance reports he viewed on White House grounds. Nunes suggested that those reports identified President Trump and members of his transition team, whose names

NUNES CONTINUED ON A4



JONATHAN ERNST/REUTERS

Intelligence Committee chair Rep. Devin Nunes (R-Calif.)

Bannon's circle loses ground in feud with Kushner faction

BY ASHLEY PARKER, ROBERT COSTA AND ABBY PHILLIP

Stephen K. Bannon — the combative architect of the nationalistic strategy that delivered President Trump to the White House — now finds himself losing ground in an internecine battle within the West Wing that pits the "Bannonites" against a growing and powerful faction of centrist financiers led by the president's son-in-law, Jared Kushner.

Less than 100 days into Trump's chaotic presidency, the White House is splintering over policy issues ranging from taxes to trade. The daily tumult has created an atmosphere of tension and panic within the West Wing, leaving aides fearing for their jobs and cleaving former allies into rivals sniping at one another in the media.

The infighting spilled into full view this week after Trump removed Bannon from the National Security Council's "principals committee," a reshuffling that left the president's chief strategist less fully involved in the administration's daily national security policy while further empowering Lt. Gen. H.R. McMaster, Trump's new national security adviser.

Bannon, an unkempt iconoclast, has generally chafed at the transition from firebomb campaigning to more modulated gov-

BANNON CONTINUED ON A4

IN THE NEWS



MELISSIA VALLARDES FOR THE WASHINGTON POST

A heckler onstage Don Rickles, who skewered audiences and fellow celebrities with his acerbic barbs, died at 90. C1

Twitter v. DHS The company asked a court to block an order that seeks to reveal the user of an account critical of Trump policies. A2

THE NATION

A federal judge stayed one of eight executions scheduled over 11 days this month in Arkansas, which has drawn national scrutiny for setting so many in such a short time frame. A2

The IRS said that financial details for up to 100,000 people might have been stolen in a security breach of an online tool used to apply for federal student aid. A2

Nearly a quarter of U.S. adults ages 18 to 59 in a recent two-year period had a type of genital human papillomavirus that elevated their risk of certain cancers, the CDC reported. A3

Comcast detailed its new cellphone-service plans, a move that other cable providers are expected to mimic as they try to expand beyond their customers' living rooms. A6

Mexican cartel leader Alfredo "El Mochomo" Beltrán-Leyva, once allied with Joaquín "El Chapo" Guzmán, was sentenced to life in prison and a $529 million forfeiture. A16

THE WORLD

Does North Korea have a missile capable of reaching the United States? It's probably just a matter of time, analysts predicted. A10

Russian investigators said they arrested eight possible accomplices in the St. Petersburg subway attack that killed 14 people. A12

Philippine President Rodrigo Duterte ordered troops into disputed islands in the South China Sea, reviving a rift with China. A17

THE ECONOMY

President Trump is considering public debt to fund $1 trillion in infrastructure investment, a major break from his campaign pledge to pull in massive amounts of private-sector money. A14

White House economic adviser Gary Cohn backed a revival of Glass-Steagall, which could break up big banks. A14

THE REGION

A Justice Department attorney expressed "grave concerns" in court about a federal plan to make changes to Baltimore's police department. B1

Maryland became the first state to agree to reimburse Planned Parenthood clinics if Congress defunds the organization. B1

INSIDE



WEEKEND
Get hoppy
How local bars choose which beers to offer, plus our own choice list of 12.

STYLE
Silent treatment
How Fox and Bill O'Reilly are dealing with the flap over harassment. C1

BUSINESS NEWS A14
COMICS C5
OPINION PAGES A19
LOTTERIES B3
OBITUARIES B9
TELEVISION C4
WORLD NEWS A10

Printed using recycled fiber

PostPoints

DAILY CODE, DETAILS, B2

7 4 1 5

CONTENT © 2017 The Washington Post / Year 140, No. 123

Reproduced with permission of copyright owner. Further reproduction prohibited without permission.

B.2 Guidelines for Conservative Trolls

TACTICS FOR EFFECTIVE CONSERVATIVE BLOGGING By Karl Rove

23 April 2012 at 19:35

Engage Demand an elaborate, time-consuming comparison / analysis between your position and theirs.

Entangle Insist that the Liberal put their posts in their own words. That will consume the most time and effort for the Liberal poster.

They will be unable to spread numerous points on numerous blogs if you have them occupied. Allowing a Liberal to post a web link is too quick and efficient for them. Tie them up. We are going for delay of game here.

Demoralize Dismiss their narrative as rubbish immediately. Do not even read it. Once the Liberal goes through the trouble to research, gather, collate, compose and write their narrative your job is to discredit it. Make it obvious you tossed their labor-intensive narrative aside like garbage. This will have the effect of demoralizing the Liberal poster. It will make them unwilling to expend the effort again, and for us, that is a net win.

Attack Attack the source. Any Liberal website or information source must be marginalized, trivialized and discounted. Let the blogosphere know that Truthout.org, thinkprogress.org,the nation and moveon.org are Liberal rubbish propaganda. Discredit Liberal sources of information whenever possible.

Confuse Challenge the Liberal position with questions, always questions. The questions need not be relevant. The goal is to knock the Liberal poster off their game, and seize control of the narrative.

Once you have control you can direct the narrative to where you want it to go, which is always away from letting the Liberal make their point. Conversely, do not respond to their leading questions. Don't rise to their bait.

Contain Your job is to prevent the presentation and spread of Liberal viewpoints. Do anything you must do to prevent a Liberal poster from presenting a well-reasoned argument or starting a civil discussion. Don't allow a Liberal to present their dogma unchallenged EVER.

Intimidate Taunt the Liberals. If you find yourself in a debate with a Liberal where you are losing a fact-based argument then call them a name to derail their diatribe. Remember your goal is to prevent a meaningful exchange of views and ideas which may portray Liberalism in a positive light. Your goal as a conservative blogger is to stop the spread and advance of the Liberal agenda. Play upon any identifiable idiosyncrasies, character flaws, physical traits, names, to their disadvantage . Monitor other posts for vulnerabilities you can exploit. Stay on the offensive with Liberal wimps. Don't let up.

Glossary

agenda-setting the act of establishing priorities in a real or synthetic person's mind. 10

agent is an abstract representation of a person (legal or natural) who absorbs and emits information and who has attitudes as well as defined behaviours. 11, 63, 73, 74

attitude a feeling or opinion about an entity. 1, 4, 8, 20, 33, 71, 73, 76, 83, 86

attitude change the updating of attitude towards an entity as the result of exposure to information. 54, 56, 74

attitude dimension a type of attitude, eg: sentiment, in an attitude space . 20, 37, 44, 53

attitude reinforcement a hardening of beliefs as an agent becomes more sure of its attitude towards an entity. 54, 56

attitude space see attitude vector space. 37, 74

attitude vector a mathematical object describing the attitude "charge" across many attitudinal dimensions. 43, 49, 50, 52

attitude vector space the normed, partially ordered vector space isomorphic to $\mathbb{R}^n$ defined by n orthonogonal attitude dimensions using keywords. 9, 16, 49, 50, 111

broadcast information that is transmitted (emitted) by an agent consisting of statements and/or broadcasts. 13, 15, 16, 68, 71, 76

channel the entity which allows a message to travel from its source to its target. For example, Fox News is a channel that acts a conduit for a source like Sean Hannity. Not to be confused with medium. 13, 15, 18, 20, 29, 65, 68–71, 80, 82

characteristic relatively immutable attributes associated to an agent, such as geographical location, religion, race, education, and core values. 20

collective the set of all agents in a simulation, the synthetic population. 1, 21, 22, 27, 36, 48, 63–65, 67, 74–76, 80

credibility when greater than zero, it is the probability of perceiving as true a message that has been received. Negative credibility is the probability of perceiving a message as false. 9, 15, 20, 22, 36, 38, 47, 65

entity a distinct simulation element including agents, statements, broadcasts, locations, etc. 8, 13, 21, 112

frame the manner in which information is presented. 46

homophily affinity due to similarity. 9, 12, 15, 28, 30, 32, 34, 39–41, 74

instantiation from computer science. The specification of an abstract concept. The creation of a member of a given class of objects with specific values, the creation of an instance. 63, 65, 73

intensity the strength of the action performed by a subject in a statement. 9, 47, 50

keyword a word which has a predefined meaning and attitudinal effect, eg: "Good". 8, 37, 44, 74, 111

mean attitude vector the average attitude towards an entity. 44, 52, 56

medium the format in which a message is conveyed. Some examples are sound (eg: radio, podcasts), video (eg: television, film, web video), image (eg: billboard, jpg files), writing (eg: newspapers, blogs), and in person (eg: political rallies). 3, 14, 15, 28, 44, 69–71, 111

memorable belief the belief an agent has in the truthfulness of a statement and how memorable it is. 25, 47, 52

message an information bundle consisting of a statement, source, channel, time of creation, and optionally, a time of expiration. 14, 20, 38

object that which is affected by the action of the verb, the result of an action by the subject of a statement.. 8, 13, 41, 50, 52, 71, 74

rally, rally-round-the-flag the short-term popularity following an event which (1) directly involves the US and particularly the president, (2) is specific, dramatic, and sharply focused . 2, 29, 34, 42, 56, 59, 61, 77, 79

sentence a cohesive and complete unit of thought. 7

sentence, conditional a sentence denoting logical implication composed of a main and subordinate statement, where *if* the main statement is true, *then* the subordinate statement is also true. 11

sentence, imperative a sentence that denotes an order, a command. 12

sentence, interrogative a sentence that denotes a question or request. 12

source the entity that emits a message. 14, 15, 20, 22, 29, 39–41, 43, 44, 64–66, 74, 76

SOV Subject-Object-Verb word order. 7

statement the foundational unit of information created by a combination of a subject, verb, object and intensity. 4, 7, 14, 16, 18, 23, 26, 28, 30, 33, 45, 47, 49, 52, 73, 86

statement, counter-attitudinal a statement with an attitude vector which is opposite in sign to an agent's prevailing attitudes. 54, 56, 65

statement, pro-attitudinal a statement with an attitude vector which is similar in sign to an agent's prevailing attitudes. 54, 58

subject the entity which performs an action in a statement. 8, 13, 15, 30, 50, 52, 64, 68, 71, 74

SVO Subject-Verb-Object word order. 7, 9

SVOi Subject-Verb-Object with an intensity parameter. 9, 11

target the recipient of a message. 13, 15, 37, 39, 40, 43, 66, 69, 74, 76

verb that which defines the action performed by the subject on the object. In the context of this thesis, it is more generally defined as a map of attitudes from the object to the subject of a statement.. 8, 9, 21, 41, 50, 52

VSO Verb-Subject-Object word order. 7

Bibliography

Ajzen, I. and M. Fishbein (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin 84*(5), 888.

Axelrod, R. (1997). Advancing the art of simulation in the social sciences. In *Simulating social phenomena*, pp. 21–40. Springer.

Bafumi, J. and R. Y. Shapiro (2009). A new partisan voter. *The Journal of Politics 71*(1), 1–24.

Baker, W. D. and J. R. Oneal (2001). Patriotism or opinion leadership? The nature and origins of the "rally'round the flag" effect. *Journal of Conflict Resolution 45*(5), 661–687.

Bakshy, E., I. Rosenn, C. Marlow, and L. Adamic (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, pp. 519–528. ACM.

Barbeau, A. E., C. Robinson, and C. Wilcox (2011). A vine with many branches. In *The Oxford Handbook of American Public Opinion and the Media*. Oxford University Press.

Baum, M. A. (2002). The constituent foundations of the rally-round-the-flag phenomenon. *International Studies Quarterly 46*(2), 263–298.

Baum, M. A. (2003). *Soft news goes to war : public opinion and American foreign policy in the new media age*. Princeton, N.J.: Princeton University Press.

Baum, M. A. and T. Groeling (2008). New media and the polarization of American political discourse. *Political Communication 25*(4), 345–365.

Baum, M. A. and T. Groeling (2010). Reality asserts itself: Public opinion on Iraq and the elasticity of reality. *International Organization 64*(3), 443–479.

Baum, M. A. and P. B. Potter (2008). The relationships between mass media, public opinion, and foreign policy: Toward a theoretical synthesis. *Annu. Rev. Polit. Sci. 11*, 39–65.

Baum, M. A. and P. B. Potter (2019). Media, public opinion, and foreign policy in the age of social media. *The Journal of Politics 81*(2), 747–756.

BBC News (2004). Net blamed for rise in child porn. *BBC*. `http://news.bbc.co.uk/2/hi/technology/3387377.stm`, accessed 2019-09-06.

Bramoullé, Y., A. Galeotti, B. Rogers, and P. Pin (2016). Stochastic network formation and homophily. In *The Oxford Handbook of the Economics of Networks* (1 ed.). Oxford University Press.

Brody, R. (1991). *Assessing the president: The media, elite opinion, and public support*. Stanford University Press.

Brousmiche, K.-L., J.-D. Kant, N. Sabouret, F. Prenot-Guinard, and S. Fournier (2017). From field data to attitude formation. In *Advances in Social Simulation 2015*, Volume 3 of *Advances in Social Simulation 2015*, pp. 1–14.

Byrne, D. (1961). Interpersonal attraction and attitude similarity. *The Journal of Abnormal and Social Psychology 62*(3), 713.

Centola, D. (2010). The spread of behavior in an online social network experiment. *science 329*(5996), 1194–1197.

Centola, D. (2011). An experimental study of homophily in the adoption of health behavior. *Science 334*(6060), 1269–1272.

Centola, D. (2018). *How behavior spreads : the science of complex contagions*. Princeton analytical sociology series. Princeton, New Jersey: Princeton University Press.

Centola, D. and M. Macy (2007). Complex contagions and the weakness of long ties. *American journal of Sociology 113*(3), 702–734.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of personality and social psychology 39*(5), 752.

Chaiken, S. and A. H. Eagly (1976). Communication modality as a determinant of message persuasiveness and message comprehensibility. *Journal of Personality and Social Psychology 34*(4), 605.

Cole, D. (2019, June). Trump's visit to a Virginia church after mass shooting left some congregants 'hurt.' the pastor says he didn't mean to 'endorse' the President. *CNN*. `https://www.cnn.com/2019/06/04/politics/trump-church-virginia-beach-shooting-pastor-letter/index.html`, accessed 2019-06-06.

Conte, R. and M. Paolucci (2002). *Reputation in artificial societies: Social beliefs for social order*, Volume 6. Springer Science & Business Media.

Davidsson, P. (2002). Agent based social simulation: A computer science view. *Journal of artificial societies and social simulation 5*(1).

de Benedictis-Kessner, J., M. Baum, A. J. Berinski, and T. Yamamoto (ND). Persuading the enemy: Estimating the persuasive effects of partisan media with the preference-incorporating choice and assignment design. *American Political Science Review*.

Delgado, R. (2004). Law professors examine ethical controversies of peer-to-peerfile sharing. *Stanford Report*. `https://news.stanford.edu/news/2004/march17/fileshare-317.html`, accessed 2019-09-06.

Druckman, J. N. (2001). The implications of framing effects for citizen competence. *Political behavior 23*(3), 225–256.

Dupree, C. H. and S. T. Fiske (2019). Self-presentation in interracial settings: The competence downshift by white liberals. *Journal of Personality and Social Psychology*.

Edwards III, G. C. and T. Swenson (1997). Who rallies? the anatomy of a rally event. *The Journal of Politics 59*(1), 200–212.

Eichenberg, R. C. (2003). Gender differences in public attitudes toward the use of force by the united states, 1990–2003. *International Security 28*(1), 110–141.

Epstein, J. M. and R. Axtell (1996). Growing artificial societies: social science from the bottom up (complex adaptive systems). *A Bradford Book*.

Fink, E. and D. Cai (2013). Discrepancy models of belief change. In *The SAGE handbook of persuasion : developments in theory and practice* (2nd ed. ed.). Thousand Oaks, Calif.: SAGE.

Gibbons, R. (1992). *A primer in game theory*. Pearson Education.

Gibson, J. L. (2001). Social networks, civil society, and the prospects for consolidating Russia's democratic transition. *American Journal of Political Science*, 51–68.

Gillis, W. (2017). The anti-semitic roots of the "liberal news media" critique. *American Journalism 34*(3), 262–288.

Goelet, P., V. F. Castellucci, S. Schacher, and E. R. Kandel (1986). The long and the short of long–term memory—a molecular framework. *Nature 322*(6078), 419.

Golbeck, J. (2009). Trust and nuanced profile similarity in online social networks. *ACM Transactions on the Web (TWEB) 3*(4), 12.

Golbeck, J. A. (2005). *Computing and applying trust in web-based social networks*. PhD Thesis.

Grinberg, N., K. Joseph, L. Friedland, B. Swire-Thompson, and D. Lazer (2019). Fake news on twitter during the 2016 us presidential election. *Science 363*(6425), 374–378.

Groeling, T. and M. A. Baum (2008, October). Crossing the Water's Edge: Elite Rhetoric, Media Coverage, and the Rally-Round-the-Flag Phenomenon. *Journal of Politics 70*(4), 1065–1085.

Harfoush, R. (2009). *Yes We Did! An inside look at how social media built the Obama brand*. New Riders.

Hegselmann, R., U. Krause, et al. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation 5*(3).

Herrman, J. (2019, March). The internet's endless appetite for death video. *The New York Times*. `https://www.nytimes.com/2019/03/24/style/really-bad-stuff.html`, accessed 2019-05-28.

Hippler, F. (1940). Der ewige jude. *Deutsche Film Gesellschaft*. Terra Film.

Hochreiter, S. and J. Schmidhuber (1997). Long short-term memory. *Neural computation 9*(8), 1735–1780.

Huddy, L. and E. Cassese (2011). On the complex and varied political effects of gender. In *The Oxford Handbook of American Public Opinion and the Media*. Oxford University Press.

Iyengar, S. (1987). *News that matters : television and American opinion*. American politics and political economy. Chicago: University of Chicago Press.

Jacoby, W. G. (2011). Attitude organization in the mass public. In *The Oxford Handbook of American Public Opinion and the Media*. Oxford University Press.

Jaffe, G. (2015, December). The quiet impact of Obama's Christian faith. *The Washington Post*. `https://www.washingtonpost.com/sf/national/2015/12/22/obama-faith/`, accessed 2019-05-28.

Jager, W. and F. Amblard (2004). Uniformity, Bipolarization and Pluriformity Captured as Generic Stylized Behavior with an Agent-Based Simulation Model of Attitude Change. *Computational & Mathematical Organization Theory 10*(4), 295–303.

Jerit, J., J. Barabas, W. Pollock, S. Banducci, D. Stevens, and M. Schoonvelde (2016). Manipulated vs. measured: Using an experimental benchmark to investigate the performance of self-reported media exposure. *Communication Methods and Measures 10*(2-3), 99–114.

Kim, S.-y., C. S. Taber, and M. Lodge (2010). A computational model of the citizen as motivated reasoner: Modeling the dynamics of the 2000 presidential election. *Political Behavior 32*(1), 1–28.

Kornblith, G. J. (2011). Class and politics to 1877. In *The Concise Princeton Encyclopedia of American Political History*. Princeton University Press.

Krebs, F. and A. Ernst (2017). A spatially explicit agent-based model of the diffusion of green electricity: Model setup and retrodictive validation. In *Advances in Social Simulation 2015*, Volume 3 of *Advances in Social Simulation 2015*, pp. 217–230.

Lazer, D. M., M. A. Baum, Y. Benkler, A. J. Berinsky, K. M. Greenhill, F. Menczer, M. J. Metzger, B. Nyhan, G. Pennycook, D. Rothschild, et al. (2018). The science of fake news. *Science 359*(6380), 1094–1096.

Levin, D. Z. and R. Cross (2004). The strength of weak ties you can trust: The mediating role of trust in effective knowledge transfer. *Management science 50*(11), 1477–1490.

Levin, D. Z., R. Cross, and L. C. Abrams (2002). Why should I trust you? Predictors of interpersonal trust in a knowledge transfer context. In *Academy of Management Meeting, Denver, CO*.

Mace, J. H. (2007). Involuntary memory: Concept and theory.

Matarazzo, J. D. and D. O. Herman (1984). Relationship of education and iq in the wais—r standardization sample. *Journal of Consulting and Clinical Psychology 52*(4), 631.

McCroskey, J. C. and J. J. Teven (1999). Goodwill: A reexamination of the construct and its measurement. *Communications Monographs 66*(1), 90–103.

McLeary, P. and A. Rawnsley (2017, April). Sitrep: Secdef says syria has chemical weapons; second guessing u.s. generals; militaries scramble to meet north korea threat. *Foreign Policy*. `https://foreignpolicy.com/2017/04/21/sitrep-secdef-says-syria-has-chemical-weapons-second-guessing-u-s-generals-militaries-scramble-to-`

meet-north-korea-threat-differing-takes-on-iran-agreement-france-attack-wikileaks-in-washington/, accessed 2019-05-28.

Mercier, A., A. Ouakrat, and N. Pignard-Cheynel (2018). Facebook pour s'informer? actualité et usages de la plateforme par les jeunes. In *#info : commenter et partager l'actualité sur Twitter et Facebook*, Le (bien) commun. Paris: Éditions de la Maison des sciences de l'homme.

Milgram, S. (1963). Behavioral study of obedience. *The Journal of abnormal and social psychology 67*(4), 371.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological review 63*(2), 81.

Nesbit, J. (2016, October). Donald trump supporters are using a nazi word to attack journalists. *Time*. http://time.com/4544562/donald-trump-supporters-lugenpresse/, accessed 2019-05-28.

O'Shaugnessy, J. and N. O'Shaughnessy (2003). *Persuasion in advertising*. Routledge.

Pennycook, G., T. D. Cannon, and D. G. Rand (2018). Prior exposure increases perceived accuracy of fake news. *Journal of experimental psychology: general*.

Pennycook, G. and D. G. Rand (2018). Who falls for fake news? the roles of bullshit receptivity, overclaiming, familiarity, and analytic thinking. *Journal of personality*.

Perloff, R. M. (2013). Political persuasion. In *The SAGE handbook of persuasion : developments in theory and practice* (2nd ed. ed.). Thousand Oaks, Calif.: SAGE.

Pew Research Center (2017). Political typology reveals deep fissures on the right and left. *Pew Research Center*.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In *Theories of emotion*, pp. 3–33. Elsevier.

Ries, S., J. Kangasharju, and M. Mühlhäuser (2006). A Classification of Trust Systems. In R. Meersman, Z. Tari, and P. Herrero (Eds.), *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, Lecture Notes in Computer Science, pp. 894–903. Springer Berlin Heidelberg.

Samuel, T. (2016, April). The racist backlash obama has faced during his presidency. *The Washington Post*. .

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology 1*(2), 143–186.

Schillo, M., P. Funk, and M. Rovatsos (2000). Using trust for detecting deceitful agents in artificial societies. *Applied Artificial Intelligence 14*(8), 825–848.

Scott, J. (1988). Social network analysis. *Sociology 22*(1), 109–127.

Shen, L. and E. Bigsby (2013). The effects of message features. In *The SAGE handbook of persuasion : developments in theory and practice* (2nd ed. ed.). Thousand Oaks, Calif.: SAGE.

Sherif, M. and C. I. Hovland (1961). Social judgment: Assimilation and contrast effects in communication and attitude change.

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics 87*(3), 355–374.

Squazzoni, F., W. Jager, and B. Edmonds (2014). Social simulation in the social sciences: A brief overview. *Social Science Computer Review 32*(3), 279–294.

Tang, J., H. Gao, X. Hu, and H. Liu (2013). Exploiting homophily effect for trust prediction. In *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 53–62. ACM.

Thiriot, S., J.-D. Kant, et al. (2008). Using associative networks to represent adopters' beliefs in a multiagent model of innovation diffusion. *Advances in Complex Systems (ACS) 11*(02), 261–272.

Tversky, A. and D. Kahneman (1981). The framing of decisions and the psychology of choice. *Science 211*(4481), 453–458.

Van Eck, P. S., W. Jager, and P. S. Leeflang (2011). Opinion leaders' role in innovation diffusion: A simulation study. *Journal of Product Innovation Management 28*(2), 187–203.

Van Hooren, S., A. Valentijn, H. Bosma, R. Ponds, M. Van Boxtel, and J. Jolles (2007). Cognitive functioning in healthy older adults aged 64–81: a cohort study into the effects of age, sex, and education. *Aging, Neuropsychology, and Cognition 14*(1), 40–54.

Vosoughi, S., D. Roy, and S. Aral (2018). The spread of true and false news online. *Science 359*(6380), 1146–1151.

Webley, K. (2010, September). How the nixon-kennedy debate changed the world. *Time*. `http://content.time.com/time/nation/article/0,8599,2021078,00.html`, accessed 2019-05-28.

Weiler, A. (2005). Information-seeking behavior in generation y students: Motivation, critical thinking, and learning theory. *The journal of academic librarianship 31*(1), 46–53.

Westen, D., P. S. Blagov, K. Harenski, C. Kilts, and S. Hamann (2006). Neural bases of motivated reasoning: An fmri study of emotional constraints on partisan political judgment in the 2004 us presidential election. *Journal of cognitive neuroscience 18*(11), 1947–1958.

WorldNetDaily (2012). If arpaio's right, where's 'the real birth certificate'? *WorldNetDaily*. `https://www.wnd.com/2012/04/if-arpaios-right-wheres-the-real-birth-certificate/`, accessed 2019-05-21.

Wyer, R. S. and D. Albarracín (2005). Belief formation, organization, and change: Cognitive and motivational influences. In *The handbook of attitudes*, pp. 273–322. Lawrence Erlbaum Mahwah, NJ.

Wyer Jr, R. S. and J. Hartwick (1980). The role of information retrieval and conditional inference processes in belief formation and change. *Advances in experimental social psychology 13*, 241–284.

Zaller, J. and D. Chiu (1996). Government's little helper: Us press coverage of foreign policy crises, 1945–1991. *Political Communication 13*(4), 385–405.

Zaller, J. R. (1992). *The nature and origins of mass opinion*. Cambridge university press.

Zinn, H. (1999). *A people's history of the United States : 1492-present* (20th anniversary ed. ed.). New York: HarperCollins.