Interpreting Author Intentions by Analyzing Story Modulation

by

Suri C. Bandler

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

February 2019

Author . . . .
Department of Electrical Engineering and Computer Science
January 18, 2019

Certified by . . . .
Patrick H. Winston
Ford Professor of Artificial Intelligence and Computer Science
Thesis Supervisor

Accepted by . . . .
Katrina LaCurts
Chair, Master of Engineering Thesis Committee

Interpreting Author Intentions by Analyzing Story Modulation

by

Suri C. Bandler

Submitted to the Department of Electrical Engineering and Computer Science
on January 18, 2019, in partial fulfillment of the
requirements for the degree of
Master of Engineering in Electrical Engineering and Computer Science

Abstract

If we are to understand human intelligence, then we need to understand human story understanding competencies, including our ability to communicate. Communication can be thought of as an externalization of an inner model of the world or an attempt to shape the inner model of the world of another.

To communicate effectively, humans must analyze not only what is said, but also how it is said. My goal in this work was to develop a cognitive model of how we produce a coherent argument, explain its elements, and provide a full analysis of authorial intent.

In this thesis, I propose a cognitive model of Story Modulation, or how humans glean information about a communicator's intentions or attempt to shape the inner story of their audience via key characteristics of wording. The model explains how we assemble textual evidence such as passive voice, instances of harm, and use of hedging words such as *alleged*, to tell a coherent story of the communicator's rhetorical goals.

I demonstrate this computational model with an implementation, RASHI, that recognizes and systematically highlights intentions. The implementation reads short news-like stories in simple English and identifies modulations in text that reveal the author's intent to influence three areas—sympathy, agency, and doubt. The system gathers objective evidence using a system of modular experts, interprets the evidence with culturally-specific subjectivity models, and distills the potentially-conflicting interpretations into a short, coherent argument about the author's intentions.

I argue that RASHI, as a computational model of human communication, can be used to improve discourse surrounding the media, elevate education in critical reading, facilitate political negotiations and resolutions, and help us bridge gaps across cultures by transforming stories to be more culturally appropriate.

Thesis Supervisor: Patrick H. Winston
Title: Ford Professor of Artificial Intelligence and Computer Science

Acknowledgments

Thank you, Mom, for teaching me that there's no magic to it. For always encouraging me to be more optimistic. For challenging me to think for myself, and to create my own way. Baba, thank you for always getting me from the train, no matter the time, and for always welcoming me home with tibit. Rachel, thank you for being the best example of what it means to love research and to give it your all. Thank you, Adam, for teaching me not to take myself too seriously. Thank you all for celebrating my every accomplishment, big or small, and for continuously showing me that I could do it, as long as I was willing to try.

Thank you, Dylan, for continuously reminding me the importance of agency. For the ways in which we are very different and for the ways in which we are not at all. For your unwavering help, in its many forms. Thank you for showing me that I brought something unique to the table, that I was contributing in a way that only I could.

Thank you, Jason, for being there from the very start and especially for being there in every moment that I felt I could not do it alone. For making it so clear that there was no reason for me to do it alone, to begin with, but that I was every bit capable, regardless. Thank you for simply making everything better, always.

Thank you Professor Winston for proving that both actions and words speak volumes. For always providing an example, often with props. For introducing me to artificial intelligence, and for guiding and supporting me through my various roles as AI student, teacher, and researcher. For demonstrating what it means to take care of people. For asking for my opinion. For treating every lecture as the most important of the semester, and for teaching me to do so too. For trusting me with 6.034 and 6.yyy. For teaching me how to communicate. And of course, thank you for the stories.

Contents

List of Figures

List of Tables

Chapter 1

We Choose How to Tell a Story

If we are to understand and develop a computational account of human intelligence, then we need to understand and model human story understanding faculties. As Winston describes in the Strong Story Hypothesis: “the mechanisms that enable humans to tell, to understand, and to recombine stories separate our intelligence from that of other primates” [Winston and Holmes, 2018]. A crucial component of human story understanding comes in the form of communication. In *Why Only Us: Language and Evolution* [Berwick and Chomsky, 2015], Berwick and Chomsky argue that the reason we have spoken language was because we needed to externalize our symbolic, nested, inner-language to others.

A full model of story understanding and therefore human intelligence requires a meaningful understanding of human communication. In order for this understanding of communication to go beyond the superficial, such a model would need to be capable of understanding not just what is said, but how it is said. For instance, a news agency might report a police shooting either as “Adult kills Minor” or as “Terrorist is killed by Police.” Although both can describe the same event, each statement might have a drastically different impact on a reader and can provide powerful insight into the author’s stances and values.

To model the human capability to both generate and interpret story modulation, I developed a cognitive model of the process by which humans identify the ways an author’s modulation of a story affects how it is understood. I then implemented the theory in a computational system called RASHI (Recognize and Systematically Highlight Intentions). Built as a module on top of the Genesis Story Understanding System [Winston, 2014], RASHI reads short news-like stories consisting of approximately 20 sentences in simple English such as the following example:

> **Regime slaughters the freedom fighter.** The Regime is a government. The regime is ruthless. Max is a boy and Max fights for freedom. Max helps our people. Max is a freedom fighter. Two barbarians were executed by the freedom fighter and the regime alleges a terrorist classification for the boy. The boy had no choice. The regime has killed Max. The regime alleges the legitimacy of the boy's death. The regime prevents freedom. We wait for the boy's body. We want to honor him with a funeral. Tonight we will host a memorial service. This was a barbaric attack. We increase the resistance. Join us and fight with us for freedom.

Of course, the same elements of this story could be used to describe a different perspective in which Max's actions are categorized as terrorism. In Section 5.5: *RASHI Recognizes Many Different Author Motivations*, I present one such story and describe the role of rhetorical modulation in identifying the contrasting authorial intentions.

Using a principled theory of human rhetorical techniques as described in this thesis, RASHI argues about how the author's word choices serve communicative goals such as affecting the reader's perceptions of sympathy, agency, and doubt. Given the above story, RASHI argues about how the author's word choice modulates agency:

> **Modulation of agency.** Overall, the author directs the majority of blame at the regime. Not only did the author blame the regime most frequently, but also the author casts favor on other agents. The author did so by excusing other agents' actions by casting their inequities as retribution, while not doing so for the regime. Similarly, the author casts favor on other agents as compared to the regime by referring to other agents as having done good deeds, while not doing so for the regime. Although the author uses passive voice for the regime, they do so only once and so it cannot be concluded that the author mitigates this blame by describing the regime's action as passive.

The cognitive theory and corresponding program are humanly plausible by design, namely no capabilities that go beyond the scope of human ability are assumed or required. In particular, the theory is:

⋆ **Systematic** (using a system of specialized experts to identify cues in the text)

⋆ **Structural** (understanding text not as a bag-of-words, but as meaningful in relation to its components)

⋆ **Principled** (relying on explicitly articulated behavioral and knowledge hypotheses)

⋆ **Explanatory** (able to form arguments and defend them with textual evidence and definite reasons).

I demonstrate how the RASHI system, based on this theory and armed with story understanding [Winston, 2011], goes beyond generating binary classification or relying on simple keywords and human experts. The theory, and accordingly the implementation, are crucially based upon story understanding to facilitate a full, systematic interpretation of a communicator's modulation with explanation at every step.

Modeling story modulation would empower us to educate children to think critically about their own words and the words of others [Winston, 2019], would improve political negotiation by facilitating understanding, would elevate discourse surrounding the media by grounding discussion in the objective and the explicitly subjective, and would help us tell stories that are more relevant to those of other cultures, while maintaining our same set of high level goals.

In this thesis, I demonstrate concrete contributions towards modeling human rhetorical intelligence, taking the state of the art computational model of story understanding and human intelligence, Genesis, to the next level.

Chapter 2

Story Understanding and Genesis

What separates us from other life forms? Mosquitoes are able to avoid being hit, rats are able to remember sequences that form a path, and chimpanzees can learn (or perhaps more aptly put, memorize) components of sign language. Are these life forms intelligent? Certainly. But we're uniquely intelligent in that we're symbolic.

2.1 Being Symbolic Makes Us Different

What does it mean to be symbolic? Being symbolic means that we can form "complex, highly nested symbolic descriptions of classes, properties, relations, actions and events" [Winston and Holmes, 2018]. In *Why Only Us: Language and Evolution* [Berwick and Chomsky, 2015], Berwick and Chomsky argue that the ability that separates human intelligence from the intelligence of other life forms is *merge*, or the ability to take two concepts `A` and `B` and combine them to create a third concept `C` without destroying `A` or `B`. We can continue this infinitely.

So, perhaps it is merge that enables us to be symbolic, but it is our symbolic ability, along with our ability to impose various forms of constraints, that allows us to form even more complex, highly nested, symbolic descriptions called inner-stories.

> **Inner-story**: A collection of complex, highly nested symbolic descriptions of properties, relations, actions, and events, usefully connected with, for

example, causal, means-ends, enablement, and time constraints [Winston and Holmes, 2018].

Eventually, humans learned to externalize these inner-stories and to internalize the stories that they heard from others, creating a positive feedback loop. This process of externalization can take many forms: verbal speech, carvings on tablets, pencilled notes, broadcast video, and more. All can be considered **outer-stories**:

> **An outer-story**: Anything that produces an inner story [Winston and Holmes, 2018].

Crucially, humans seem to uniquely demonstrate this story-understanding capability of internalizing stories heard from others and of externalizing their own inner stories, resulting in the Strong Story Hypothesis.

> **The Strong Story Hypothesis**: The mechanisms that enable humans to tell, to understand, and to recombine stories separate our intelligence from that of other primates [Winston and Holmes, 2018].

And so, if we are going to computationally model human intelligence, then we need to model the process of forming inner-stories, either directly from our experiences or from stories heard from others, and the process of externalizing our inner-stories into outer-stories.

2.2 We Form Unique Inner-Stories

As the saying goes, there are three sides to every story – my side, your side, and what really happened[1]. Reframing this idea in terms of story understanding, there are events that occur, the inner-story that I form after witnessing them and the inner-story that you form after witnessing them.

[1]Multiple people are credited with various forms of this expression, including Jeyn Roberts and Robert Evans

If we are going to understand human intelligence we need to understand how humans form and use their unique inner-stories. Given the same perceptual input, two people form *different* inner-stories, and so when they describe the events that they witnessed at a later point, they do so differently or in other words, they produce different outer-stories. That is why two people can witness the same event but describe it entirely differently afterwards. A fundamental and underlying motivation for pursuing such an understanding is to develop a theory that enables a computational framework to produce a coherent argument, explain its decisions, and to provide a full analysis of a communicator's intent.

I argue that humans developed rhetorical intelligence, or the ability to use rhetorical choices to manifest their unique inner-stories and to attempt to dictate how others form their own inner-stories. In the absence of the ability to directly observe the formation of or storage of an inner-story, we can use a communicator's outer-story as a means of inferring what their inner-story is or what kind of inner-story they are attempting to mold in their audience. Accordingly, I define story modulation as follows:

> **Story Modulation:** The process by which a story-teller uses rhetorical device to manifest their unique inner-story or to shape the inner-story of their audience.

Modeling human story-understanding, therefore, must go beyond a model of who-what-where-when-and-why, beyond looking for causal links, and forming mental models of who-knows-what. If we are to build a robust model of human story-understanding capabilities, we must also model how story-tellers modulate their stories. After all, its not just what you say, but how you say it.

2.3 The Genesis Story Understanding System Models Human Intelligence

In this section, I describe the Genesis Story Understanding System, upon which I built RASHI, the implementation component of this thesis. Genesis computationally models human story understanding capabilities, thereby modeling human intelligence.

2.3.1 Genesis Follows the Computational-Imperative Principle

The Genesis Story Understanding System [Winston, 2014], referred to in short as Genesis, is a system that computationally models human story understanding. Importantly, everything added to Genesis follows a *computational imperative*:

> **The Computational-Imperative Principle**: Any model of human intelligence should introduce only computational capabilities that enable observed behaviors without enabling unobserved behaviors [Winston and Holmes, 2018].

In other words, nothing is added to Genesis unless it is necessary and unless it could plausibly model what a human is actually doing in the process of understanding stories. Figure 2-1 [Winston and Holmes, 2018] provides an overview of the humanly-plausible layers upon which Genesis is built.

2.3.2 Genesis Translates English Stories into Innerese Frames

Genesis takes as input a textual stories of about 20-30 sentences in simple English, parsing the input using the START parser [Katz, 1997], generating *innerese*. Innerese is Genesis's inner language whose primary components (each modeled as a different Java class) are Entities, Functions, Relations, and Sequences. I provide a brief overview of each representation here but a more detailed explanation available in the Genesis implementation substrate documentation [Winston, 2015].

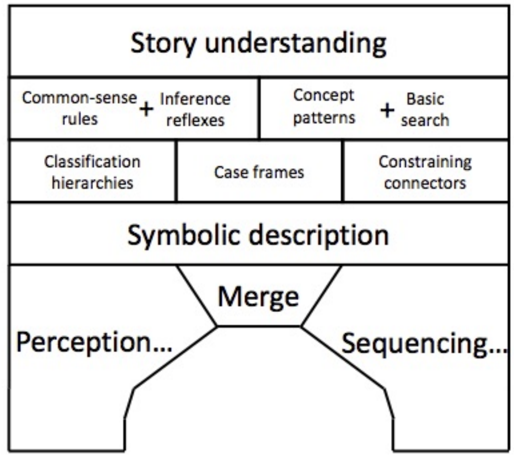

Figure 2-1: The Genesis Story Understanding System (Genesis) is a computational model of human story understanding. Genesis is empowered by deductive and inductive rules, providing it with common-sense and inference. Genesis can also recognize concepts by searching stories for patterns using basic search. Going a level deeper, these faculties are empowered by constraints, classification hierarchies, and case frames, which are modeled on our uniquely symbolic capabilities. Figured sourced from [Winston and Holmes, 2018]

.

Entity An *entity* is the fundamental building block of innerse. An entity has a unique identifier and a *bundle* of *threads*. Threads define the meanings of word and are derived from WordNet [Fellbaum, 1998], a lexical database that captures synonyms, definitions, and the hierarchical relationship of words. A bundle is an unordered grouping of threads. At its base, an entity is an object.

Function A *function* is an entity with the addition of a subject field. A function represents Jackendoff's paths and places elements [Jackendoff, 1983] or more generally a role in a case frame. Functions usually depict prepositional phrases, the direct object of a verb, or an adverb. For example, an English manifestation of a function might be "next to the computer," with "computer" as the subject.

Relation A *relation* depicts how one entity relates to another. For example, an English manifestation of a relation might be "the coffee mug next to the computer" or "Patrick drinks coffee."

Sequence A *sequence* is an ordered container of elements.

2.3.3 Genesis Uses Common Sense in the Form of Rules and Concept Patterns

Commonsense knowledge in the form of rules supplies missing causal connections, and knowledge in the form of concept patterns helps identify overarching themes such as revenge. If-then rules, such as the rule found in Figure 2-2 can be deductive or inductive.

```
If xx refers to yy as a victim, then xx evokes sympathy for
yy.
```

Figure 2-2: The Genesis Story Understanding System uses rules to supply the common sense knowledge that is required in order to meaningfully understand a story.

Concept patterns are small stories that Genesis can identify within a larger story. Concept patterns often contain a "leads-to" relations, such as the concept of Revenge demonstrated in Figure 2-3. Most concept pattern identification therefore relies on search, differing from how common sense rules are typically applied. Some concept patterns contain a "consequently" clause that, whenever the concept pattern matches, is inserted into the original story. For example, the "Victimhood" concept pattern is defined in Figure 2-4.

Concept patterns that do not require search and that generate consequents, such as the concept pattern defined in Figure 2-4, are not different in any way real way from rules. For example, the rule defined in Figure 2-2 and the concept defined in 2-4 require the same antecedent and if found, would result in the same element being inserted into the original story. Genesis, however, labels the concepts that it finds but does not label rules. Using concepts, therefore, allows for better visual demonstration within Genesis and importantly for my computational implementation, RASHI, enables post processing and analysis of found concepts.

Detecting rules and concept patterns relies on Genesis's *matcher*. Genesis's matcher determines if the structure of two entities align and if a successful matching is found

```
Start description of "Revenge".
xx and yy are entities.
xx's harming yy leads to yy's harming xx.
The end.
```

Figure 2-3: Concept patterns are small stories that Genesis can identify within a larger story. Concept patterns often have a leads-to clause, requiring Genesis to perform search within a story to determine if there is a match between the story and the concept.

```
Start description of "Victimhood".
xx is an entity.  yy is a person.
xx refers to yy as a victim.
Consequently, xx evokes sympathy for yy.
The end.
```

Figure 2-4: The Genesis Story Understanding System can search for concept patterns within a story. Upon matching, if the concept pattern contains a consequently clause, Genesis can insert the resultant into the original story, broadening its understanding of the story.

produces a set of *bindings* that map between the two entities' corresponding elements. For example, consider two Role Frames translated from innerse to English: "John loves Mary," "John loves Susan." When provided with the translated innerese version of these two sentences, Genesis's matcher would determine that the two entities match and would map John to John and Mary to Susan. Importantly, this matcher is exposed to users of Genesis and can be employed outside of finding rules and concept patterns.

Genesis presents the story, its causal and inferential links, and highlighted concept patterns as an elaboration graph. Each story element is indicated by a box; causal and inferential links are indicated by colored lines between elements. For example, Figure 2-5 shows an elaboration graph for a simplified excerpt from Macbeth. Here, the concept pattern of revenge is found, connecting two events through intermediate elements. Revenge, formally defined in Figure 2-3, can be summarized as xx's harming yy leads to yy's harming xx.

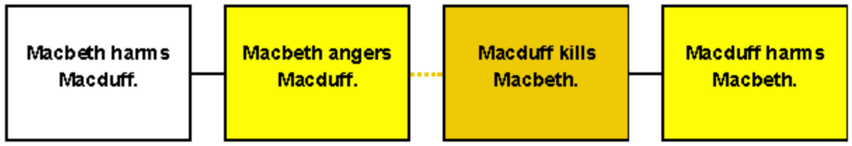

Figure 2-5: The Genesis Story Understanding System generates elaboration graphs depicting the stories elements, causal and inferential links, and highlighted concept patterns, such as this elaboration graph of a simplified version of Macbeth, demonstrating the concept pattern of revenge.

2.3.4 Genesis Needs Story Modulation

While Genesis is becoming quite skilled at the who-what-where-when-why, without a framework for Genesis to analyze the *way* in which a story is told, Genesis will never have a nuanced, complex understanding of the stories it reads.

In this thesis, I describe how, while following the computational-imperative principle, I provided a theoretical and computational framework for Genesis to model story modulation. In doing so, I took Genesis's model of human story understanding to the next level, opening the door for Genesis to be used in education, political negotiations, media discourse, and for telling more relevant stories.

Chapter 3

Building on Framing Theory, going Beyond Sentiment Analysis

On a high level, the goals of story modulation are to use a communicator's rhetorical cues to garner insights into how they formed their inner-story surrounding a given event or situation and to do so in an explainable, transparent, and humanly-plausible way as a step towards taking story-understanding and therefore our understanding of human intelligence to the next level.

As a computational theory of modulated communication, story modulation connects to work in related fields such as sentiment analysis, framing theory, and cognitive systems. A major basis of inspiration for story modulation comes from the work done by Morris and Peng.

3.1 Story Modulation Goes Beyond Sentiment Analysis

Modeling story modulation resembles other computational approaches to understanding authorial intent, such as machine learning techniques for sentiment analysis. But whereas the aim in sentiment analysis is to classify entire text fragments as belonging to one of a number of predetermined sentiments (such as "very negative", "negative", "neutral", "positive" or "very positive,") (see for example [Pang et al., 2002, Radford et al., 2017, Socher et al., 2013]), my goal for story modulation was to develop a cognitive model of how we produce a coherent argument, explain its elements, and provide a full analysis of authorial intent rather than to classify across discrete categories.

As discussed previously in Chapter 2: *Story Understanding and Genesis*, humans are symbolic, creating complicated, nested symbolic structures. This is reflected in how we tell stories and can often lead to seemingly or truly contradictory efforts and to the demonstration of various priorities, even within the same news piece. Any system that truly models human story understanding capabilities, therefore, needs to be compositional in order to able to detect, interpret and distill a communicator's various story telling techniques.

Even approaches to sentiment analysis that are technically compositional in design lack the same level of story-based explainability. For example, Socher et al. [Socher et al., 2013] trained a neural network to use a parse tree of a text fragment as input and to determine its sentiment through a recursive processing of nodes in the parse tree, updating its understanding as it climbs. Although this treats a sentence as a composition of nodes, the system is neither capable of explaining its conclusions for a given example nor of providing an interpretable, detailed analysis, instead generating a numerical result and a discrete classification.

A story understanding approach to modulation provides another crucial distinction from and benefit as compared to machine learning techniques: human plausibility. Machine learning algorithms tune millions of parameters by training on thousands of data points. Humans do not, instead learning from very few examples and yet they are still capable of explaining their reasoning. If story modulation is to serve as a model for human intelligence, then it must differ from computational approaches to sentiment analysis [Radford et al., 2017, Socher et al., 2013] in that it can not have any capabilities that could not plausibly be held by humans. By building on top of Genesis, I ensured that every new contribution satisfied a computational imperative.

3.2 Story Modulation is Related to Framing Theory

Story modulation fits naturally within the field of framing theory, which studies how communicators elicit specific emotions or mental states in their audience by invoking a particular subset of the audience's beliefs [Entman, 1993]. Indeed, my computational framework, RASHI, relies on subjectivity models that can be viewed as the collection of framing techniques at the author's disposal.

My theory distinguishes itself from traditional approaches to framing theory through a set of additional commitments. First, I argue for the importance of a systematic and computational approach. Previous work in framing theory has revolved around case studies, with researchers and the field itself accumulating observations over time. In contrast, story understanding provides the tools to *integrate* fundamental observations into compositional knowledge modules in order to accumulate evidence resulting in a systematic framework for analyzing authorial intentions. Similarly, whereas framing theory lacks a coherent definition [Hallahan, 1999] and often uses vague, "casually-defined" categories [Entman, 1993], I developed a concrete framework of authorial motivations, which I demonstrated through a precise, principled computational implementation. Combined, story understanding serves as a cohesive, cognitive, and computational foundation for studying framing theory. Hallahan posits that storytelling is the most complex form of framing [Hallahan, 1999], but I claim that it is the exact opposite: framing is a complex form of story understanding and so it is best modeled using the conceptual constituents of stories.

Second, whereas framing theory tends to focus on the audience [Chong and Druckman, 2007, Hallahan, 1999, Touri and Koteyko, 2015], story modulation focuses on a goal-directed author. As a result, RASHI goes beyond other computational implementations, such as those of Touri and Koteyko [Touri and Koteyko, 2015], that use statistical keywords to determine salient excerpts and then use human experts to explain the frames that they might evoke. Rather than rely on simple keywords determined to be important due to their relative frequency, my approach models the writing process as an author choosing from a set of universal rhetorical mechanisms. In my system, subjectivity is made explicit via subjectivity models, which can be applied consistently across multiple stories.

3.3 Story Modulation is Related to Cognitive Systems

I present a cognitive model of how modulation of communication signals the mental state of a communicator, similar to Langley's cognitive systems analysis of personality and conversational style [Langley, 2017]. Our approaches are congruent in that they both characterize domain knowledge, design a principled architecture, and attribute communication cues to individual cognitive differences. The main difference in our work is essentially one of domain, or more specifically, what factor is modulating our communication: while Langley focuses on personality, my focus is instead on authorial intent.

3.4 Story Modulation is Inspired by Morris and Peng

Morris and Peng demonstrated that English-language journalists tend to portray violence in terms of the assailant's disposition, while Chinese-language journalists tended to portray violence in terms of the situational constraints on the assailant. Additionally, American subjects were shown to be more likely to attribute dispositonal rather than situational causes to murder, while Chinese subjects were shown more likely to attribute situational over disponsitional [Morris and Peng, 1994].

My work is much inspired by their findings. Framing theory, for example, focuses on the impact that a particular frame might have on a reader's interpretation, but I argue that it is necessary to start from the author before addressing any such impact: what were the author's intentions in writing this piece? I claim that deducing these intentions can be split into two stages: what rhetorical choices did an author make and what are the implications of each of these choices?

Although the second stage, determining the implication of an author's rhetorical choices is unavoidably subjective, Morris and Peng demonstrated that trends can be identified and that there is value in delineating cultural differences and how they impact both the writing and reading of different stories.

In this thesis, I primarily focus on the author's intentions but I demonstrate how modeling an author's intention can lead to generating framing techniques to impact the audience, namely how it is possible to use the same techniques used for interpreting an author's intentions to generate a news piece for a new audience with different cultural values or priorities.

Chapter 4

A Cognitive Model of Rhetorical Intelligence

Any cognitive model of rhetorical intelligence must be able to account for the complex, nested structure of communication. It must also be capable of incorporating context. Story understanding provides a framework for doing both.

To begin, I present a cognitive theory of human rhetorical intelligence and a concrete formulation of the commonsense knowledge upon which this intelligence relies. My cognitive theory consists of three claims about how humans deploy rhetorical intelligence.

I claim that humans:

- ⋆ approach a story as a **composite**
- ⋆ draw on **domain-specific** expertise
- ⋆ can **justify** their claims about rhetorical effect at every step

4.1 Humans Approach a Story as a Composite

Rather than treating an article as a bag-of-words or relying on the presence or absence of individual keywords, human readers get a sense for the story by understanding the structure of sentences and paragraphs and by analyzing the story as a whole to determine a coherent, global argument about the author's intent. Even within the same sentence, the author's choices can interact in complex ways. Human readers identify not only words with negative sentiment, such as harm, but also strategic choices such as using passive voice or characterizing the victim as a child or a terrorist—factors that influence who gets blamed for the harm and how much.

When forming an overall argument about the author's intent, readers take a global view. Readers consider the full article, instead of individual words or sentences, gathering evidence to support their argument and even distilling potentially conflicting evidence to form a coherent portrait of the author's intentions. Just as a reader must process a story in its entirety to identify the concept of revenge, so too must a reader process a story in its entirety to identify an author's overarching intentions.

Even within the same sentence, it is often the interaction between authorial choices that indicate meaning. For instance, human readers not only identify words like "harm" to determine the author's placement of blame, but also look for potentially mitigating factors. Such factors come in the form of syntactic cues, such as using specific sentence roles or using passive voice, and semantic cues, such as using a person's role in society or mentioning their age. Human readers use these factors to determine whether the author is mitigating blame by diverting attention from the perpetrator or by casting the recipient of harm as a deserving victim.

4.2 Humans Draw on Domain-Specific Expertise

While reading a story, humans draw upon commonsense knowledge not only to make inferences regarding events in the story itself—as is typically acknowledged or implemented in current systems—but also to identify rhetorical cues that signal authorial intent.

If an author says "The bank was robbed" then a human reader will recognize that the author fails to implicate an agent and instead focuses on the event itself rather than the perpetrator. In contrast, if an author uses passive voice in a scientific article in a biology journal, a human reader will recognize that this rhetorical choice simply reflects industry standard and does not carry implications about authorial intent.

Just as cognitive models rely on commonsense knowledge regarding causality in events, so too must a cognitive model capable of reasoning about an author's intentions rely on commonsense knowledge about rhetorical implications. Commonsense knowledge about rhetorical implications is domain-specific and can refer to various aspects such as an author's role in society (e.g, biologist versus crime reporter), culture (e.g, nationality, religion), etc.

4.3 Humans Can Justify Their Claims about Rhetorical Effect at Every Step

If I presented you with an article from a newspaper and asked: what are the author's intentions in this piece? You might respond: "it seems the author is trying to garner sympathy for Macbeth." If I asked: how do you know that? You might respond: "Well, the author says that Macbeth is murdered by Macduff and so he is portrayed as the victim." If I followed with: is there anything else that the author does to induce this sympathy? You might say: "Yes, the author describes how Macbeth was trying to help the kingdom before he was murdered."

Humans are capable of interpreting a story and answering questions at various levels of understanding. When asked about an author's intentions, humans can point to specific pieces of evidence within the story, summarize the available evidence, present an argument, and identify how the evidence supports or weakens that argument. Any story understanding system, and any complete modelling of human intelligence, must be able to do so, too.

Explainability also builds trust. If a system is going to be used to help educate children, facilitate political negotiations, civilize media discourse and help bridge cultural divides, then it must be able to explain itself and its decisions.

4.4 A Conceptual Formulation of Authorial Intent

In addition to these three claims, I present in Table 4.1 a conceptual formulation of the kind of knowledge that humans require in order to identify and interpret rhetorical choices. This knowledge consists of the intentions that an author might possess—such as instilling credibility and evoking sympathy—and potential rhetorical mechanisms for achieving these intentions.

Table 4.1: Examples of authorial intentions and the rhetorical mechanisms that authors can employ to achieve them. Underlined mechanisms indicate techniques implemented in my computational system.

Author intention	Example mechanisms
Cast doubt	<u>Use terms such as “alleged” or “believed”</u>, use sarcasm or scare quotes
Evoke sympathy	<u>Emphasize age/role in society</u>, <u>use passive voice</u>, use colorful quotations or analogies, mention specific individuals, emphasize magnitude
Convey Importance	<u>Repeat ideas</u>, <u>use passive versus active voice,</u> change sentence order, emphasize impact, evoke feelings of unity/solidarity
Modulate blame	<u>Use passive vs. active voice</u>, <u>justify via assigning role/title</u>, <u>contextualize actions relative to others</u>, claim expertise via location of writing, use idioms, make generalities
Make moral claim	<u>Use passive voice</u>, <u>justify via assigning role/title</u>, write from specific location, use first-hand accounts, assign epithets, make accusations, use idioms/generalities
Instill credibility	Quote statistics, cite an expert, imply via section in paper published, demonstrate novelty
Align stories	Reference, make metaphor or allusion
Create intrigue	Omit information, use hyperbole, reference a specific or underdeveloped example
Cause surprise, humor	Break expectations (see for example [Taylor, 2018])

Of course, Table 4.1 does not describe every authorial intent or example mechanism. I argue that no system would be able to capture the seemingly infinite number of intentions or approaches that an author might use but that it is possible to capture a large fraction of value by identifying major pillars and building from them. This table gives insight into the breadth of what humans are capable of and identifies ways to get concrete value from such a huge space of possibilities.

Chapter 5

A Computational Approach: RASHI Recognizes and Systematically Highlights Intentions

I developed the computational system RASHI (**R**ecognize **A**nd **S**ystemtically **H**ighlight **I**ntentions) to both demonstrate the viability and to refine the details of my cognitive theory. Armed with story understanding, my system reads a story, gathers explicit and objective textual evidence, interprets the evidence using models of the author's subjective frame of mind, and distills the potentially contradictory results into an argument about authorial intent. I discuss each of these capabilities in turn, using the following *Regime slaughters the freedom fighter* story as an example:

> **Regime slaughters the freedom fighter.** The Regime is a government. The regime is ruthless. Max is a boy and Max fights for freedom. Max helps our people. Max is a freedom fighter. Two barbarians were executed by the freedom fighter and the regime alleges a terrorist classification for the boy. The boy had no choice. The regime has killed Max. The regime alleges the legitimacy of the boy's death. The regime prevents freedom. We wait for the boy's body. We want

to honor him with a funeral. Tonight we will host a memorial service. This was a barbaric attack. We increase the resistance. Join us and fight with us for freedom.

Figure 5-1 depicts the flow of data within RASHI, summarizing this chapter.

5.1 Genesis Reads the Story

RASHI is built as a module on top of the Genesis Story Understanding System [Winston, 2014]. Building RASHI on a story understanding substrate ensures that not only the theory but also the components of the implementation realistically model what humans can actually do—they meet a computational imperative [Winston and Holmes, 2018]. Accordingly, RASHI's analysis consists of story-enabled operations such as pattern matching, rule-based inferences, and building complex nested representations.

I note that compositionality is a key feature of RASHI's story-enabled approach: RASHI does not merely haphazardly identify keywords or snippets of evidence but instead composes explanations in terms of the meaning of the constituents of sentences and of the story overall. This is particularly important as it is often the interaction between authorial choices that have a profound impact on gaining insight into an author's overall intentions. For example, an author may believe that harm is not always wrong, but justified in instances of self-defense. RASHI is capable of handling such concepts by, for instance, not just looking for helpful or harmful action keywords but for the relationship between subject, action, and object.

Additionally, composition allows for explainability, as RASHI can use constituent components of a given sentence or of a story as pieces of evidence supporting a final conclusion. An example of using constituent components in rhetorical analysis would be distilling an author's intention at a story level: although an author might blame an agent in a specific sentence, a coherent global assessment might determine that an author mitigates this blame by demonstrating that the victim was deserving or that the agent performed many other good deeds, too. I give an example of one approach

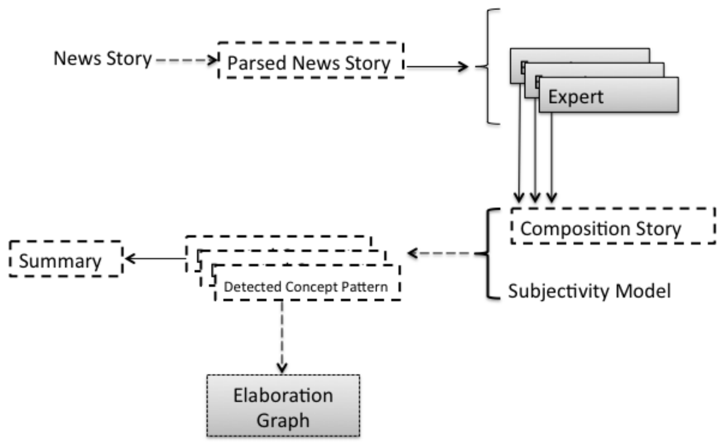

Figure 5-1: A depiction of the flow of data in RASHI. Plain text in the figure, such as News Story, indicate plain text input. Dotted arrows indicate that the processing is handled by the Genesis Story Understanding System while a solid arrow indicates that the processing is handled by RASHI. Dotted boxed text, such as Parsed News Story, indicates that data that is represented internally in the Genesis System. All such data can be externalized by Genesis into simple English. Grey boxes indicate a module of logic, where a solid border indicates that it is handled by RASHI and a dotted border indicates that it is handled by Genesis. In this figure, a News Story in simple, plain text is parsed into a Parsed News Story by Genesis. Then, RASHI applies its Experts, compiling the output of each into the composition story, which is represented internally in Genesis. Genesis then uses a provided subjectivity model, which is comprised of plain text descriptions of concept patterns, to filter the composition story, outputting a set of detected concept patterns. This set of concept patterns can be visualized as an elaboration graph by Genesis, or can be summarized by RASHI.

to such a process in Section 5.4: *RASHI Distills Evidence into a Coherent Argument*. Story understanding, therefore, provides a robust foundation for modeling rhetorical analysis.

To begin the analysis of authorial intent, RASHI takes as input a textual story consisting of about 20 simple English sentences. Throughout this work, the process for composing input stories was as follows: I would find example articles in the news or compose a sample news story from multiple perspectives. Then, I would feed it through the parser, adapting the level of language to allow for parsing. Importantly, I gathered and composed these stories independently of implementing RASHI. Instead, I gathered my corpus of stories in conjunction with delineating the potential rhetorical intentions and example mechanisms outlined in Table 4.1. The corpus of generated stories included but was not limited to depictions of police shootings, terrorist attacks, children playing together, students winning awards, and teachers helping students. Then, independently of the stories and based upon Table 4.1, I designed and implemented RASHI. Finally, I used the stories to test the success of my system, in particular in order to better *generalize* its capabilities and to handle edge cases. As a result, RASHI can handle any story that START can parse, producing more interesting results from stories that present elements of the rhetorical devices explored in Table 5.1 and discussed below.

5.2 RASHI's Experts Gather Explicit Evidence

After reading a story, RASHI looks for textual evidence of the author's rhetorical choices to begin the process of determining the author's intentions. Table 4.1, shown previously, lists a number of *possible* intentions and their corresponding mechanisms. Of course, humans are more sophisticated than this table indicates, possessing many more possible motivations and means of achieving them. This table, however, gives insight into the breadth of what humans are capable of and identifies ways to get concrete value from such a huge space of possibilities. The present computational system can detect a subset of these mechanisms. Using a system of Experts, where each Expert is a module responsible for identifying one specific rhetorical indicator, RASHI finds textual evidence that could indicate authorial intent. Implementing these Experts required making specific implementation choices; I summarize my Experts and implementation choices in Table 5.1.

As shown in the rightmost column of Table 5.1, each Expert detects rhetorical indicators using one of a few related techniques. The most common technique is to search for word meaning. *Property finding*, as used by the Evasiveness and Enemy Experts, consists of finding properties in a predefined set. These properties can either be directly stated by the author in a particular sentence, such as in the case of "alleged terrorist", or inferred from a previous characterization. For example, if Max is previously referred to as a child, then RASHI will consider this property in future instances such as when Max is referred to as a victim. Alternatively, *Synonym matching*, as used by the Youth and Abuse Experts, uses WordNet [Fellbaum, 1998] to determine which words are close in meaning. For example, RASHI's Youth Expert will match any type of juvenile, including minor, child, and kid.

Two Experts employ special-purpose patterns that aren't about word meaning. The Passive Voice Expert simply checks the relevant lexical property found during parsing. The Repetition Expert, which activates after all other Experts, looks for repeated use of rhetorical strategies as found by other Experts, such as the repeated characterization of the same individual as a victim. Its search is supported by an Alias Expert that consolidates different names for the same individual to ensure accurate counting (i.e. which performs simple anaphora resolution). Finally, some Experts search not only the story itself, but the output of other Experts. The Karma Expert identifies harm dealt by and to enemies using information provided by the Victim Expert and the Enemy Expert. In this way, the Karma Expert provides a nuanced compositional understanding in terms of more basic Experts: the author may render a harm more forgivable by casting its victim as an enemy.

These matching techniques are more than just keyword matching or bag-of-word routines as they are both word and context dependent. When finding objective textual evidence, RASHI takes into account sentence structure, linking the textual evidence to

Table 5.1: RASHI gathers objective rhetorical evidence using eleven Experts. Certain Experts, like the Passive-Voice Expert and the Youth Expert, act directly on the text. Others, like the Repetition Expert or the Karma Expert, process the output of other Experts.

Expert name	Knowledge domain	Detection technique
Abuse Expert	Recognizes perpetrator of harm or violence	**Property Finding:** Action Valence (subject of negative action)
Victim Expert	Recognizes the recipient of harm or violence	Action Valence (recipient of negative action)
Benefactor Expert	Recognizes the agent of a positive action	Action Valence (subject of positive action)
Beneficiary Expert	Recognizes the recipient of a positive action	Action Valence (recipient of positive action)
Evasiveness Expert	Recognizes hedging tactics	**Property Finding:** Presence of Qualifier (e.g., alleged, believed)
Enemy Expert	Recognizes possible enemies	**Synonym Matching:** Wordnet (e.g., bad-person, wrongdoer, terrorist)
Youth Expert	Recognizes young entities	Wordnet (e.g., Juvenile, minor, etc.)
Karma Expert	Recognizes when help and harm enemies (i.e., deserving / undeserving figures).	**Misc.** Victim + Beneficiary + Enemy Expert
Passive-Voice Expert	Recognizes use of passive verb and lack of agent.	Presence of parser tag
Alias Expert	Consolidates different names that refer to the same individual (anaphora resolution)	Anaphora Resolution (e.g., via classifications: Mark is a terrorist. The terrorist...).
Emphasis Expert	Detects emphasis	Repetition Detection

the original source as well as noting any information regarding position in the sentence or parse. As an example, consider a step-by-step view into RASHI processing the following sample sentence: "Police kill attacker." First, RASHI parses this sentence using Genesis and the START parser. Then, each Expert analyzes the parsed version of this sentence or analyzes the output of other Experts. Running this sentence through RASHI's system of Experts, the Abuse Expert recognizes that kill is a verb that induces harm and uses the parsed form of the sentence to identify the *subject*, namely the Police, as the perpetrator. The Victim Expert also searches for instances of harm but instead makes note of the *object* of the sentence in question, here noting that the attacker is the recipient of harm. Already, RASHI's Experts rely upon the structure of a sentence rather than simply recognizing specific words.

Going a step further, the Karma Expert turns on after the Victim Expert and uses the Enemy Expert to determine if the author casts the victim in question as a potential enemy, here using WordNet to determine that an attacker is a form of wrong-doer. The Karma Expert, therefore, relies upon the context of the action, namely that harm was perpetrated against a potential enemy.

Importantly, RASHI's Experts use global context established from looking at the article as a whole, rather than only analyzing within a particular sentence. For instance, the Alias Expert consolidates different names that refer to the same individual. Therefore, if the above sentence "Police kill attacker" had instead been "...Bob is an attacker....Police kill Bob," RASHI's Karma Expert would incorporate the context learned on a global level—that Bob is an attacker—and note that the harm was perpetrated against a potential enemy. Similarly, the Emphasis Expert detects emphasis by looking at the story globally for instances of repetition.

Note, no conclusions have yet to be reached about what these rhetorical choices indicate about the author's intentions! In the next stage, when determining the implications of an author's rhetorical choices, RASHI will similarly use context determined within a specific sentence and on a global level.

Importantly, RASHI's system of Experts is both impartial (in that it strictly identifies textual evidence of rhetorical mechanisms without drawing conclusions) as well as modular (in that it is flexible to the addition of Experts and to using a multiplicity of methods). We can see this in the previous example, as no conclusions are reached by the Karma expert about the implications of the author's choices and the Karma Expert uses the output of both the Victim Expert and the Enemy Expert.

5.2.1 RASHI's Experts are Impartial

In this evidence-gathering stage, RASHI's Experts exhaustively search for all instances of a rhetorical device without drawing any conclusions regarding relevance or coherence. Intentionally separating evidence gathering and analysis allows RASHI to analyze authorial intent on a global level, taking into account both the author and the set of observations in its entirety. When processing the story as a composite, certain observations may turn out to be inconsequential, while others may be contradictory. For example, consider the sentence *"Child is killed by alleged terrorist."* Here, the available evidence simultaneously suggests that the terrorist should be both blamed, because they hurt a child, as well as exonerated because their role as a terrorist is alleged and the killing is described in passive voice. The implications of such contradictions are not yet noted, handled instead at a later stage. Additionally, the implications of rhetorical decisions will depend on context, such as using passive voice in a biology article versus in an article depicting a crime, and so they cannot be immediately determined. At this point, all possible matches are added, allowing RASHI to remain impartial in the evidence gathering stage.

5.2.2 RASHI's Experts are Modular

The cognitive theory itself remains uncommitted to particular implementation details. To reflect this, RASHI is modular; future Experts can easily be added and existing Experts can be changed to use various underlying implementations. For example, a different implementation of the Victim Expert could use an alternative knowledge base such as ConceptNet [Liu and Singh, 2004] rather than WordNet synonyms to classify an action as causing harm or benefit. The modular implementation allows for the simultaneous use of a multiplicity of humanly plausible methods.

5.2.3 RASHI's Experts Objectively Depict the Author's Choices

The textual evidence collected in this stage describe the author's rhetorical decisions. Consisting of a sequence of sentences in Genesis's inner language, the collection of evidence comprises a *composition story.* For example, the following is the composition story translated to English by Genesis for the *Regime slaughters the freedom fighter* example:

> **Composition Story.** The author refers to our people as a beneficiary. The author says that Max carries out a kindness. The author uses passivity for the regime. The author labels Max as passivity's object. The author refers to Max as a victim. The author says that the regime commits an inequity. The author uses passivity for Max. The author labels two barbarians as passivity's object. The author refers to two barbarians as a victim. The author says that Max commits an inequity against a malefactor. The author refers to Max as a victim. The author says that the regime commits an inequity. The author says that the regime alleges something. The author says classification to be alleged. The author says that the regime alleges something. The author says death's legitimacy is alleged. The author refers to Max as a youth. The author refers to Max as a victim repeatedly. The author says the regime commits an inequity repeatedly. The author says the regime alleges something repeatedly.

The above composition story is an *externalization* of Genesis's inner language representation into English which is why the phrasing is somewhat stilted.

RASHI's implementation demonstrates that designing within a specific domain can have a large impact: with just a few mechanisms, RASHI displays a powerful ability to identify a wide variety of potential modulating factors.

5.3 RASHI Interprets Textual Evidence and Focuses Intent Using Subjectivity Models

After RASHI gathers the composition story, it must evaluate and interpret the evidence as indicative of the author's intent to modulate meaning. The significance of a given piece of evidence depends on the author's own values or priorities and on their cultural background, including nationality, age, profession or the forum for which they are writing. For example, some cultures may view harm against a child as especially egregious; for such cultures, describing a victim as a child is a significant rhetorical move. As another example, passive voice may be a communal expectation.

Consider scientific writing in the biology community—community standards dictate the use of passive voice, especially in a methodology section. In that case, the use of passive voice does not carry the same connotations that it might in a newspaper article describing a crime. If we are to not only describe but interpret rhetorical choices, we must be able to model the author's world view and goals.

To this end, I developed *subjectivity models*, modular representations of an author's values, cultural connotations, and priorities. Subjectivity models consist of Genesis rules and concept patterns that connect the explicit textual information gathered in the composition story to a set of highly personal connotations and implications. Subjectivity models let RASHI interpret textual choices as rhetorically meaningful or meaningless. They let RASHI interpret textual choices as delivering a particular subjective effect and therefore serving a particular rhetorical goal.

Subjectivity models can encode not only cultural values but also areas of importance upon which to focus. For example, an individual author might prioritize modulating agency in a story. What it means to modulate agency will differ between authors, and so a subjectivity model can serve to focus the analysis on a specific topic while also clearly delineating an author's values. Figure 5-2 shows a sympathy-focused subjectivity model for an author who comes from a culture in which children are highly protected. Figure 5-3 shows an agency-focused subjectivity model, including concept patterns of `Passivity`, `Blame`, `Retribution`, `Misguided Kindness`, and `Kindness` and two rules. RASHI uses the Genesis Story Understanding System to detect instances of these concept patterns within the composition story, which depicts the full set of the author's rhetorical choices. Currently, subjectivity models are hand-crafted to represent a cohesive set of priorities. In the future, I hope to provide RASHI with the capability of hypothesizing over a universe of subjectivity models about which most likely represents an author's priorities. Section 2.3.3: *Genesis Uses Common Sense in the Form of Rules and Concept Patterns* discusses the benefits of using concept patterns instead of rules, even when there is no functional difference between the two.

In this way, subjectivity models implement knowledge of *authorial purpose* as

```
If xx refers to yy as a victim, then xx elicits sympathy for yy.
If xx evokes yy’s childhood and xx elicits sympathy for yy, then xx
increases sympathy for yy.
If xx refers to yy as a victim and xx labels yy as passivity’s object,
then xx increases sympathy for yy.

Start description of "Childhood".
xx and yy are entities.
xx refers to yy as a youth.
Consequently, xx evokes yy’s childhood.
The end.

Start description of "Victimhood".
xx is an entity and yy is a person.
xx refers to yy as a victim.
Consequently, xx evokes sympathy for yy.
The end.
```

Figure 5-2: A subjectivity model for a sympathy-focused author. This model contains concept patterns for an author who comes from a culture in which children are more highly protected.

described in the table of author intentions and example mechanisms (Table 4.1). By describing an author’s focus or cultural perspective, subjectivity models provide tools to systematically uncover an author’s intentions from rhetorical choices. Effectively, concept patterns in a subjectivity model provide the mapping backwards from rhetorical technique to underlying authorial intention. When RASHI applies the agency-focused subjectivity model to the example article’s composition story, RASHI generates the elaboration graph in Figure 5-4 detailing its conclusions.

A simplifying assumption I make with these subjectivity models is that an author’s own values will dictate how the author will attempt to modulate the story to affect an audience. This assumes that the author’s values align perfectly with the audience’s, eliminating some theory-of-mind difficulties. An interesting future direction would be to enable RASHI to reason about an author’s mental model of the audience.

```
If xx decreases yy’s agency, then xx mitigates yy’s blame.
If xx says that yy commits an inequity, then xx blames yy.

Start description of "Passivity".
xx uses passivity for yy.
Consequently, xx decreases yy’s agency.
The end.

Start description of "Blame".
xx and yy are entities.
xx says that yy commits an inequity.
Consequently, xx blames yy.
The end.

Start description of "Retribution".
xx and yy are entities.
xx says that yy commits an inequity against a malefactor.
Consequently, xx mitigates blame against yy.The end.

Start description of "Misguided Kindness".
xx and yy are entities.
xx says that yy carries out a kindness against a malefactor.
Consequently, xx casts doubt on yy’s intentions.
The end.

Start description of "Kindness".
xx and yy are entities.
xx says that yy carries out a kindness.
Consequently, xx casts admiration on yy’s intentions.
The end.
```

Figure 5-3: An example subjectivity model for an agency-focused author. Five concept patterns and two rules model how agency changes when the passive voice is used and how agency translates into blame, retribution, kindness, and misguided kindness and both filter and highlight relevant pieces of a composition story.

5.4 RASHI Distills Evidence into a Coherent Argument

In the previous stage, RASHI used a subjectivity model to determine what textual evidence is relevant to the given author and to generate conclusions about the impli-

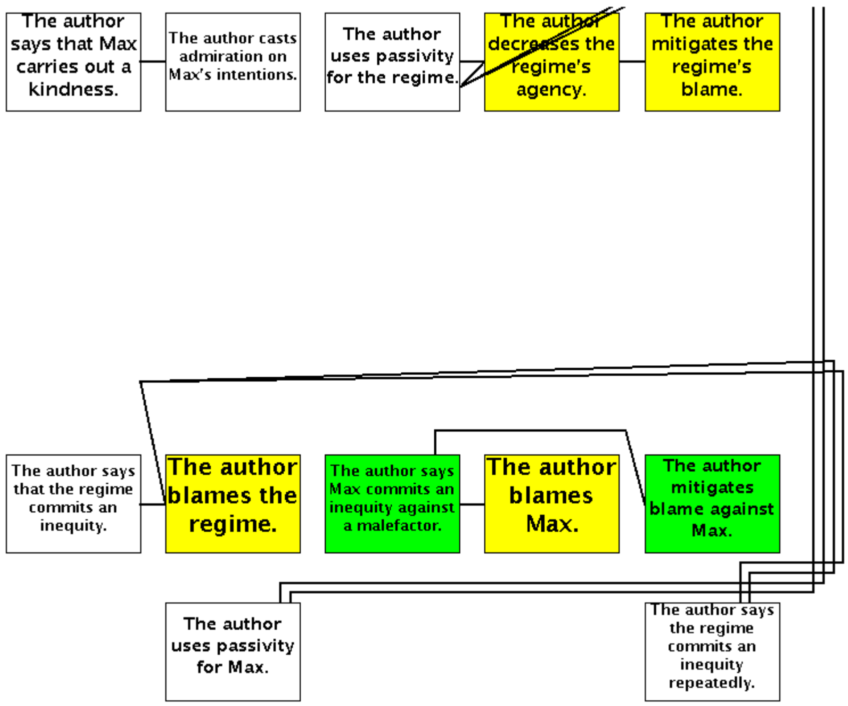

Figure 5-4: This elaboration graph depicts the textual evidence that RASHI found relevant to the author's intention to modulate the text regarding agency. RASHI uses the agency-focused subjectivity model (Figure 5-3) to determine which evidence is relevant in the composition story and to draw immediate conclusions. Here, RASHI finds evidence of Agency and Passivity, Blame, Retribution, and Kindness but does not find evidence of Misguided Kindness.

cations of the author's rhetorical choices. This procedure is exhaustive, identifying many, potentially conflicting, pieces of evidence. In the next stage, therefore, RASHI tries to reconcile the available evidence, consolidating it into a coherent picture of the author's intentions.

For the example article and the agency-focused subjectivity model (Figure 5-3), RASHI generates the following summary:

Modulation of agency. Overall, the author directs the majority of blame at the regime. Not only did the author blame the regime most frequently, but also the author casts favor on other agents. The author did so by excusing other agents' actions by casting their inequities as retribution, while not doing so for the regime. Similarly, the author casts favor on other agents as compared to the regime by referring to other agents as having done good deeds, while not doing so for the regime. Although the author uses passive voice for the regime, they do so only once and so it can not be concluded that the author mitigates this blame by describing the regime's action as passive.

RASHI uses a series of hand-coded heuristics that comprise a decision tree to summarize the author's intentions (see similar discourse production techniques, e.g. [Davey and Longuet-Higgins, 1978]). The summary process forms an argument regarding a specific goal stemming from a central concept pattern. In the example summary, the central concept pattern is "Blame" from the agency-focused subjectivity model found in Figure 5-3. The process is the same regardless of the topic of focus and is described below.

The summarizer first tallies the different participants in the "Blame" concept patterns occurring throughout the story. The idea is that repetition can help resolve conflicting or ambiguous authorial cues. For example, repetition can serve as a heuristic to identify an author's primary target, as a consistently blamed character, presumably, indicates that the author deliberately intends to place blame.

Next, having found a character who stands out as a primary target, the summarizer searches for secondary effects that may mitigate or exacerbate the blame placed on the target. The summarizer checks to see whether blame is exclusively placed on the target or whether the blame is spread across others. Using hand-coded knowledge about the relationship between concept patterns (such as blame, retribution, and kindness), the summarizer also checks which characters are portrayed as having done kindnesses and whether any blameworthy actions constitute acts of retribution. When the target character is excused in this way, the summarizer describes the blame as comparatively less; when other characters are excused in this way, the summarizer describes the target's blame as comparatively greater (an approach taken by, for example, [Sayan, 2014]).

To generate natural-sounding discourse, the summarizer effectively walks down a tree, determining at each stage if the latest evidence strengthens or contradicts previous observations, and filling in the appropriate built-in text template accordingly. For example, RASHI first "branches" according to if the most blamed character is the *only* blamed character (choosing between "Overall, the author directs the majority of the blame at `xx`" or "Overall, the author unilaterally blames `xx`"). Then, if the author exacerbated the character's guilt in any way, RASHI adds "Not only did the author blame `xx` most frequently, but also...". If RASHI finds multiple exacerbating factors, such as casting others' harm as retribution and describing others as having done good deeds, RASHI bridges the explanations with "Similarly,...". The opposite effect occurs when there is evidence of the author mitigating the blame rather than exacerbating it.

Currently, these summarizations techniques are hand-coded and are meant to be a systematic means of capturing the domain of knowledge currently embodied in RASHI. The summarizer is build on top of an extendable framework and exciting next steps include exploring additional domains and building the decision tree directly from concept patterns.

At present, RASHI uses Genesis for parsing stories, analyzing Genesis's internal, hierarchical representation of sentences rather than the text directly. Because of this structure it is possible to identify roles in a sentence like subject and object and parts of speech, used for identifying the action of a sentence. The START parser also exposes markers such as when a sentence is passive (to enable matching, Genesis stores a passive sentence in innerese in its active form, maintaining the passive marker exposed by START). Finally, RASHI uses Genesis to recognize the concept patterns that comprise an author's subjectivity model, as discussed in Section 5.3: *RASHI Interprets Textual Evidence and Focuses Intent Using Subjectivity Models*.

5.5 RASHI Recognizes Many Different Author Motivations

To show the breadth of RASHI's capabilities, I demonstrate RASHI's results on two example stories and three subjectivity models. The first example story is *Regime slaughters freedom fighter*, as described above. The second story is an alternative telling of the first story from a different perspective:

> **A mother and child were murdered by a terrorist.** The killing is a terror attack. Terror attacks hurt our city. The mother and child were going to the doctor. The mother and child waited at the bus stop. The terrorist killed the mother and child with kitchen knives. The mother was killed by the terrorist while protecting the child. The terrorist stabbed the mother and child in the back. The funerals are tomorrow and the city encourages the community to come. The city will increase security because of the terrorist attack. People need to remain alert.

In Table 5.2, I juxtapose RASHI's conclusions regarding agency (Figure 5-3), sympathy (Figure 5-2), and doubt (Figure 5-5) in these two stories, where each concepts is defined according to the subjectivity model referenced by figure.

```
If xx alleges yy, then xx casts doubt on yy.

Start description of "Object of Doubt".
The author says that yy is alleged.
Consequently, the author casts doubt on yy's label.
The end.
Start description of "Decrease Credibility".
xx and yy are entities.  xx says that yy alleges something.
Consequently, xx decreases yy's credibility.
The end.
```

Figure 5-5: The subjectivity model modelling an author whose focus is on casting doubt or on making statements lose credibility include a rule linking allegations to doubts, and two concepts regarding being the object of doubt and decreased credibility.

Here, subjectivity models act like filters, serving to focus RASHI on a specific area of relevance to the author. It is possible that an author might care about all of these things and nothing prohibits RASHI from using a combined subjectivity model. To demonstrate the wide variety of conclusions reached, I include RASHI's summarizer's output for each combination of area of focus and example story. As demonstrated through this side by side comparison, RASHI could ideally be used to create a political spreadsheet, comparing various news sources across different dimensions.

While contradictions are reconciled via the summarizer, it is still interesting and important to see the apparent contradictions in an author's rhetorical choices. Beyond striving for a coherent argument about the author's intentions, RASHI brings to light the many nuances that exist within the same piece and can potentially be used to help authors ensure that their rhetorical choices match their implicit intentions during the editing process. Additionally, RASHI allows readers to note possible concessions made by the author, which can be helpful during political negotiation in terms of finding a potential avenue for compromise.

Table 5.2: A comparison of RASHI analyzing two different stories describing the same event along three dimensions: sympathy, agency, and doubt.

Regime slaughters freedom fighter.	*Mother and child murdered by terrorist.*
Agency. Overall, the author directs the majority of blame at the regime. Not only did the author blame the regime most frequently, but also the author casts favor on other agents. The author did so by excusing other agents' actions by casting their inequities as retribution, while not doing so for the regime. Similarly, the author casts favor on other agents as compared to the regime by referring to other agents as having done good deeds, while not doing so for the regime. Although the author uses passive voice for the regime, they do so only once and so it cannot be concluded that the author mitigates this blame by describing the regime's action as passive.	**Agency**. Overall, the author unilaterally blames the terrorist. The author's use of passive voice for the terrorist does not substantially mitigate the placement of blame because the author unilaterally and repeatedly blames the terrorist.
Sympathy. Overall, the author evokes the most sympathy for Max. Not only did the author refer to Max as a victim most frequently, but also the author emphasizes this sympathy by referring to Max as a child. Similarly, but to a lower degree, the author emphasizes sympathy for Max by using the passive voice when describing Max as receiving harm. By doing so, the author increases attention on Max's victimhood.	**Sympathy**. Overall, the author evokes the most sympathy for the mother. Not only did the author refer to the mother as a victim most frequently, but also the author emphasizes this sympathy by using the passive voice when describing the mother as receiving harm. By doing so, the author increases attention on the mother's victimhood.
Doubt. Overall, the author solely questions the credibility of the regime. The author does so by stating that the regime made qualified claims, such as by 'alleging', 'suspecting', or 'believing' rather than directly claiming or stating. At the same time, because the regime qualified its remarks, the author casts doubt on the regime's claims as well. More specifically, the author casts doubt on the terrorist classification and on the boy's death's legitimacy.	[No modulation of doubt.]

Chapter 6

Taking RASHI to the Next Level: Making Stories Relevant

In this chapter, I discuss how to take RASHI to the next level by enabling it to bridge cultural divides by making stories more culturally appropriate and by telling stories more persuasively to targeted audiences. I present exploratory work that is very much in its early stages: I demonstrate the theoretical approach with a first-step towards a prototypical expansion of RASHI that transforms a story to be more sympathetic to an audience with a different value system for sympathy.

6.1 Align Goals by Aligning Inner Stories

Consider an excerpt from the usual example, *Regime Slaughters Freedom Fighter*:

> ...Max is a boy...The regime has killed max.

In some cultures, it would seem that the author of this story is trying to intensify sympathy for Max (and intensify blame for the regime) by mentioning Max's role as a child. In RASHI, you can represent the concept of sympathy and intensified sympathy due to the victim being a child in the subjectivity model. The subjectivity model

```
Start description of "Victimhood".
xx is an entity and yy is a person.
xx refers to yy as a victim.
Consequently, xx elicits sympathy for yy. The end.

Start description of "Childhood".
xx and yy are entities.
xx refers to yy as a youth.
Consequently, xx invokes yy's childhood.
The end.

Start description of "Intensified Sympathy For a Child".
xx and yy are entities.
xx invokes yy's childhood.
xx elicits sympathy for yy.
Consequently, xx intensifies sympathy for yy.
The end.
```

Figure 6-1: An author might elicit sympathy for a particular entity by referring to them as a victim. At the same time, an author might emphasize someone's role in society, such as by referring to them as a child. Such intentions can have compounding effects: An author whose intentions are to intensify sympathy for a particular entity might do so through the rhetorical decision of referring to the person that they elicited sympathy for as a child. RASHI captures these rhetorical mechanisms in three concept patterns: `sympathy`, `childhood`, and `intensified sympathy for a child`.

includes three concepts `Sympathy`, `Childhood` and `Intensified Sympathy for a Child`, represented explicitly in a subjectivity model in Figure 6-1 and visually in Figure 6-2.

It is not necessarily the case, however, that every culture views "taking candy from a baby" as especially wrong, or in other words that hurting a child carries special connotations. Consider a particular nationalistic society, in which the worst type of person that you can hurt is a patriot. If I asked you to tweak the characteristic example for an audience of people from this second, nationalistic society, what might you do?

Clearly, if the goal is to make the article as compelling as possible to the second audience, then it needs to include elements relevant to that particular culture.

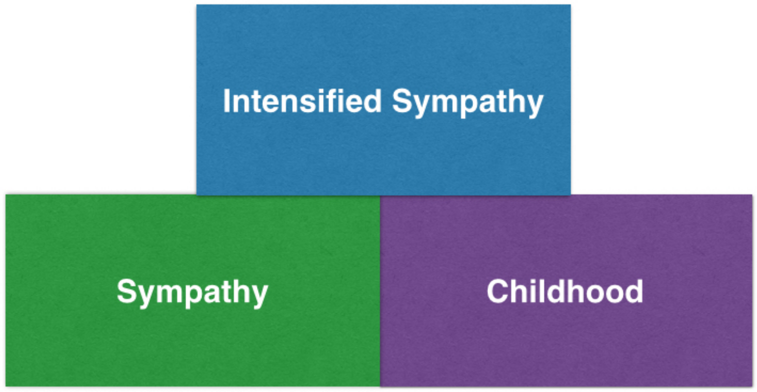

Figure 6-2: In one culture, perhaps in the culture of the United States, an author's intention may be to intensify sympathy for a particular recipient of harm by referring to the victim as a child. We can visualize the concepts combining as a pyramid, with the concept of `Intensified Sympathy` built on top of the concepts of `Sympathy` and `Childhood`

More specifically, the inclusion of relevant elements can be though of as a transformation of the pyramid from Figure 6-2 such that `intensified sympathy` is based upon `sympathy` and `patriotism` rather than `sympathy` and `childhood`. This transformation can be visualized in Figure 6-3 and expressed explicitly in the subjectivity model found in Figure 6-4. To effectuate this change, you might want to ask: is Max a patriot? Can we add this particular description to make the victim more sympathetic or can we emphasize this aspect in another way?

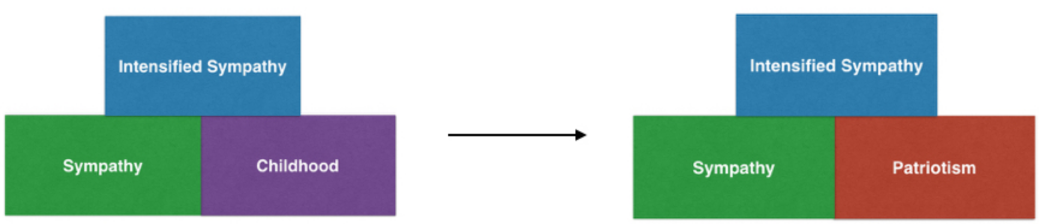

Figure 6-3: In order to transform an article such that it still achieves intensified sympathy in a second culture, its necessary to determine which conceptual building blocks for `intensified sympathy` are unique to the culture, which elements are missing, and which elements already are produced by the first culture.

For our simple example above you might want to describe the murder as:

```
Start description of "Victimhood".
xx is an entity and yy is a person.
xx refers to yy as a victim.
Consequently, xx elicits sympathy for yy. The end.

Start description of "Patriotism".
xx and yy are entities.
xx believes yy is a patriot.
Consequently, xx invokes yy's patriotism.
The end.

Start description of "Intensify Sympathy for Patriot".
xx and yy are entities.
xx invokes yy's patriotism.
xx elicits sympathy for yy.
Consequently, xx increases sympathy for yy.
The end.
```

Figure 6-4: An author might intend to increase sympathy for a victim by mentioning that the victim is a *patriot*, rather than as a child, as depicited in this subjectivity model.

> ...Max is a patriot...The Regime has killed Max...

or more subtly as:

> ...Max helps our country...The Regime has killed Max...

In both examples, the intentions are the same: generate sympathy for Max, and as much of it as possible. So in many ways, the inner-story of the author writing each article can be entirely aligned across a particular goal. But how the author chooses to manifest that inner-story into an outer-story will vary according to their own, or their audience's, cultural values and priorities. Therefore, composing or transforming an article to bridge a cultural gap comes down to aligning the inner-stories on a high level, and then determining how to achieve the high level goal with the constraints imposed by a particular culture or value system. In RASHI, such a cultural or value system is made explicit in the subjectivity model.

So, in order for RASHI to take a story written for one culture and transform it

	STEP	APPROACH
1.	Find overlapping intentions between Cultures A and B	*Consequently clause*
2.	Detect missing authorial choices	*Back-Chain*
3.	Determine how to manifest authorial choices in story	*Reverse Expert*
4.	Update original story	*Inject (delete)*
5.	Re-analyze	*Rerun*

Figure 6-5: In order to transform an article written for culture A into a form that is relevant for culture B, RASHI must find the overlapping intentions in the two culture models, detect which authorial choices are missing in order to achieve the motivation of culture B, determine how to manifest these choices in the original article, update the original article, and reanalyze the resulting piece.

for a second, it would need to: (1) recognize what an author's values or intentions are for a particular piece, (2) detect if this intention is already achieved with respect to the second culture and if not determine if there's any way to achieve this intention, and finally, (3) determine a method for transforming the original article into one that would resonate with the second culture, reanalyzing the article accordingly. These steps, and the beginnings of the approach taken to implement them in RASHI as will be described in the following sections, are outlined in Figure 6-5.

6.2 A Story Modulation Approach to Cultural Adaptation

In this section, I describe a theoretical approach for aligning authorial intentions across cultures. Throughout, I outline how RASHI can computationally model this theoretical framework and demonstrate a first-step, simple prototype built with RASHI that can transform the example story to elicit intensified sympathy due to Max's patriotism rather than childhood. Both the theoretical approach and the computational model prioritize introducing the fewest number of changes required so that a transformed article remains as faithful to the original text as possible.

6.2.1 Infer Authorial Values or Intentions

An author's world view might contain many intentions representing different values and priorities, but the author may not invoke every one in a given article. The values and priorities that an author does invoke in a particular piece serve as high level goals that an author would hope to achieve when communicating with their audience or with a new audience from a different culture. So the first step in a theoretical framework for transforming an article into a new cultural perspective is to determine which intentions an author invokes within a particular piece.

RASHI Infers Authorial Values and Intentions

Determining which values are relevant to an author in a particular article is simple – its one of the primary functions of RASHI. Because it is built on top of the Genesis Story Understanding System, RASHI can simply access which concepts were instantiated from the author's subjectivity model. The instantiated concepts will be exactly those concepts that are relevant to the author in the original article, as determined by their rhetorical choices. Once RASHI has determined which concepts are relevant, it then becomes a matter of determining which of these concepts are also relevant to the second culture.

6.2.2 Find Overlapping Intentions Between Cultures

Although an author may invoke multiple values within a particular piece, not every value will be modeled or relevant and therefore achievable in the second culture. The second step, therefore, is to find overlapping intentions between the two cultures.

RASHI Finds Overlapping Intentions via Consequently Clause

One aspect of concept patterns, discussed in Chapter 2: *Story Understanding and Genesis*, is that they may have a *consequently* clause. When a concept is matched, that consequently clause is reinserted into the story. RASHI relies on consequently clauses in concept patterns to build the elaboration graph depicting its hypothesis regarding the author's intentions and their implications. The author's intentions therefore are contained in consequently clauses. For example, consider the concept of `Intensified Sympathy for a Child` defined in Figure 6-1. By both invoking an entity's childhood and eliciting sympathy for that entity, an author consequently intensifies sympathy for that entity.

Therefore, finding overlapping intentions between two cultures can be modeled as the process of finding matching consequently clauses. More precisely, RASHI needs to determine which instantiated and therefore relevant concepts from culture A match with general, uninstantiated concepts that model culture B's values. This process is represented in Figure 6-6. The matching process, as discussed in Chapter 2: *Story Understanding and Genesis*, is handled by the Genesis Story Understanding System and produces a binding that maps entities from the general concept pattern variables to the instantiated concept pattern.

6.2.3 Determine What's Needed: Detect Missing Authorial Choices

The set of concepts from the original article that were found to overlap with intentions in the second culture form a set of requirements or goals for transforming the article for a second culture. To do so, an author would need to add to or remove elements from the original article so that they satisfy the conditions of each goal.

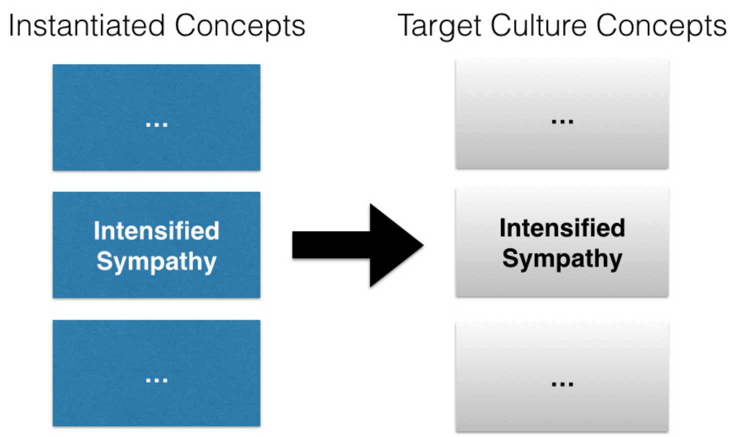

Figure 6-6: The first step in transforming an article to be more culturally relevant to a different audience is determining which of the meta-goals that are present in the original article overlap between the two cultures. To do so, RASHI searches for instantiated concept patterns that match with uninstantiated concept patterns in the target culture.

RASHI Uses Back-Chaining to Detect Missing Authorial Choices

In RASHI, this goal set consists of the concept patterns from the second culture's subjectivity model whose consequently clauses matched instantiated concept patterns from the first culture's subjectivity model. Before RASHI can transform the original article to be applicable to a second culture, RASHI must determine what rhetorical choices are missing, if any, in order to make the article match with the concept patterns that form the goal set.

RASHI looks to the antecedents of the concept patterns in the goal set to define the set of rhetorical choices missing from the original article. Visually, these are the building blocks demonstrated for `Intensified Sympathy` in Figure 6-3. This process is again simple because of the Genesis Story Understanding System as RASHI is able to directly access the antecedents for any concept pattern.

To maintain a consistent intention for specific actors between the original and the transformed versions, RASHI applies the bindings found in the instantiated concept patterns of the original article to the antecedents necessary to produce the concept pattern in the second culture. In this example, this means that in RASHI must transform the original article so that it achieves "The author elicits sympathy for Max" and "The author invokes Max's patriotism."

Before moving forward, RASHI checks to see if any of the needed antecedents have already been achieved in the original article by matching goal antecedents to elements of the composition story. In this example, RASHI determines that the author already elicits sympathy for Max but that the author does not invoke Max's patriotism. RASHI's goal therefore is to make the original article invoke Max's patriotism.

Operating on the heuristic of prioritizing simplicity, RASHI back-chains recursively to see if there are any other concept patterns that would result in a particular goal being achieved. For example, with RASHI's goal as "The author invokes Max's patriotism," RASHI searches for other concept patterns in the second culture's subjectivity model that would achieve this result. Here, RASHI finds that by referring to Max as a patriot, the author can invoke Max's patriotism. RASHI therefore replaces its goal of "the author invokes Max's patriotism" with "the author believes Max is a patriot."

6.3 Transforming a News Piece for a Different Audience

Once the conditions required to satisfy the goal set are explicitly defined, the original article must be transformed to satisfy these conditions. At the most basic level, this would mean adding in any missing elements (or removing any strictly prohibitive elements). At the next level up, it would mean removing those elements that are no longer relevant. At the highest level, this process must also account for ensuring that the article flows smoothly, and for determining where is the best place to insert particular elements.

6.3.1 Determine How to Manifest Authorial Choices

So far, RASHI has determined which intentions of an author are important in a particular piece, which of those important concepts are relevant in the second culture, and which rhetorical decisions an author needs to make in order to achieve these same intentions for a second culture. But how can RASHI determine how to manifest these authorial choices?

RASHI is comprised of a system of Experts, each of which is responsible for recognizing and highlighting particular rhetorical choices. Therefore, producing a particular rhetorical choice in the original article that was not present beforehand is a matter of *reversing* RASHI's experts. Incorporating the idea of reversing RASHI's experts is very much aligned with Radul and Sussman's Propagator Programming Model [Radul and Sussman, 2009], where RASHI's Experts serve as autonomous machines. In particular, it is important that information used in the decision making process is not lost to ensure that the process is invertable.

The inverse process will look different for any given Expert, with some Experts having a simpler inverse process than others. Importantly, everything in RASHI is linked. Every sentence in the composition story maps to the original sentence or sentences from which it is sourced. Similarly, every portion of the elaboration graph maps to the sentences in the composition story from which it is derived. Therefore, the original story's source material is available for the given Expert to use when trying to invert a particular intention.

The simplest inversion that an Expert might do is to do nothing at all. Consider an antecedent of a concept pattern that says "the author believes that Max is a patriot".

To handle the explicit base case, RASHI has an Explicit ("do-nothing") Expert. The Explicit Expert simply let's through a particular sentence from the original story if its tagged as "obvious," wrapping it in the clause "the author believes that". The explicit Expert's inverse is to unwrap the composition story's sentence, stripping away the statement explicitly referencing the author's belief. Currently, only the Explicit Expert's inverse is implemented in RASHI. The Explicit Expert generates elements to be inserted into the original story and marks them with an "obviously" tag.

Similar to the Explicit Expert, the Passive Expert would have a clear inverse process—making the sentence passive or active, accordingly. The Victim Expert can be implemented on a basic level by using passive voice, such as "Max was harmed," removing the complication of determining an assailant.

A more interesting and complicated case to consider is the Abuse Expert. If a concept pattern in the second culture's subjectivity model requires that a specific entity be cast as an abuser, how can RASHI determine *whom* the entity should abuse? One approach to this is to expand the subjectivity model to include defaults such as cultural scapegoats.

Generalizing the idea of cultural scapegoats, the choice of analogies, metaphors, and references are extremely culturally and domain specific. Therefore, when targeting a new audience RASHI must be able to identify:

- ☐ **religious metaphors or allusions** (i.e, referencing the New Testament versus the Torah),
- ☐ **national symbols** (i.e., comparing or depicting historical lessons from George Washington versus David Ben-Gurion), and
- ☐ **cultural boogeymen**, i.e., the American use of boogeyman versus the Russian use of Baba Yaga.

and then substitute more relevant choices according to the values and culture of the target audience. As with the cultural scapegoat example, an approach to allowing RASHI to make such substitutions might be to include cultural defaults in the subjectivity model.

6.3.2 Update the Original Story and Analyze with RASHI to Confirm Success

The result of each Expert's inverse would be an action to perform on the original story—for example, an addition or a deletion of a particular element, or for additions, where the element should be placed in the story. At this time, RASHI does not consider deletions and insertions are done by injecting the element through Genesis, and does not specify the location within the original story.

As a last step, RASHI analyzes the transformed piece as if it were a new article. In its current implementation, RASHI analyzes the transformed piece using both the original and the second culture's subjectivity model. Because RASHI does not currently delete elements, the newly transformed article satisfies the preconditions for both the original and the second culture. In future iterations, RASHI would only analyze the transformed article using the second culture's subjectivity model to check the instantiated concept patterns to confirm that the transformation successfully achieved the desired authorial intentions.

6.4 What Comes Next?

I have explained theoretically how RASHI can be used to bridge cultures by transforming an article written for one culture to be more applicable in another, but have of course only taken a small, first-step towards enabling RASHI to do so.

The next step is to enable all of RASHI's Experts to be invertable, through the techniques outlined in this chapter and through techniques yet to be discovered, and there are big motivating factors for doing so. Demonstrating this theoretical framework computationally would open up a world of applications in:

- ☐ **politics**, helping a diplomat transform their speech to be more applicable to their audience even while maintaining their home country's original intentions,
- ☐ **knowledge distribution**, helping an article pass through government censorship such that citizens living within countries with oppressive regimes can be exposed to outside media, and
- ☐ **education**, helping children experiment with how articles are transformed to be more persuasive, so that they can be better prepared when others are trying to manipulate them.

Chapter 7

Contributions

In this thesis, I argued that story understanding is not complete without story modulation and I have defined a cognitive theory of how humans understand not just what is said, but how it is said. I posited that reasoning about rhetorical intent must be compositional, knowledge-based, and explainable—features that are well suited to a story-understanding substrate. To flesh out the theory, I articulated nine explicit motivations that categorize an author's intentions.

By way of demonstration and elaboration of the theory, I developed RASHI, a computational framework that identifies authorial intent using a subset of these motivations and corresponding mechanisms—modulating doubt, sympathy, and agency. With its story-based substrate, RASHI is able to go beyond mere keyword matching, analyzing the structure of sentences and how sentences relate in a story to form a coherent picture of an author's intentions. RASHI not only collects evidence, but distills it into a coherent argument, identifying mitigating factors and reconciling contradictions. In this way, RASHI represents a step toward a rich, cognitive and computational understanding of human communication, taking the Genesis Story Understanding System to the next level. In the process, I made the following contributions:

I proposed a cognitive theory of critical analysis. I claimed that humans conduct their analysis on the sentence level and on the story level as compared to simply relying on bag-of-words techniques. Rather than looking for particular key words or counting word frequencies, humans use global context provided by the story as a whole to understand the relationship between words and sentences and to form a coherent understanding of the story. Humans also use local context provided within a sentence to understand the meaning of words. Additionally, I claimed that humans draw on previously acquired, domain-specific knowledge about authorial intentions and the mechanisms that are available to them and that humans can explain their reasoning at any step. I argued both that a cognitive model of critical analysis must satisfy these three claims and that story understanding provides an ideal substrate for doing so.

I identified nine author intentions and corresponding mechanisms. I highlighted a conceptual subdivision between objective and subjective components of rhetorical analysis. Additionally, the formalization of intentions and associated mechanisms lays strong groundwork for future study and expansion of RASHI. By subdividing the problem of identifying author intentions into finding objective, explicit evidence and attributing motive, I separated concerns of collecting data from concerns of subjectively interpreting it, allowing for transparency and explainability. I broke down the problem of what an author means into an analysis of rhetorical choices and their implications.

I implemented Experts that find evidence of rhetorical purpose through a subset of these mechanisms. The Experts can independently analyze a story while allowing for an integrated and then summarized analysis of their conclusions. I demonstrated through the combined result the impact that a Society of Mind [Minsky, 1986] approach can have in a specific domain by allowing for a modular while compatible collaboration of subcomponents, each of which can explain the reasoning behind its conclusions. Although RASHI depends on being provided with a subjectivity model, I envision RASHI instead using the meta-text of an article, such as the date and location it was written and basic biographical information of the author, to search for the most plausible subjectivity model and resulting conclusions.

I developed a summarizer that distills evidence into a coherent picture of authorial intent. RASHI analyzes the interplay between the author's rhetorical choices. RASHI determines the amount of weight that should be given to contradictory pieces of evidence by recognizing that the author might use techniques to emphasize or mitigate contradictory factors and by using a hierarchy of rhetorical intensity to decide between effects. The summarizer is just one way in which RASHI explains itself and the result is a natural language, coherent summary of its argument describing authorial intentions.

I outlined how RASHI can create bridges between cultures. I proposed a theoretical approach and demonstrated a basic prototype for taking RASHI to the next level by enabling it to transform an article to effectively communicate to a second culture. By analyzing the author's intentions as manifested in their rhetorical choices, RASHI can determine an author's priorities for a given piece. RASHI can then determine which of these prioritized intentions are represented and therefore achievable in a second culture's subjectivity model and what rhetorical decisions are missing in order to achieve those intentions in the new culture. Finally, I discussed how to take RASHI to the next level by designing its Experts to be invertable so that the desire for a rhetorical choice can be manifested into a story element to transform the original article. This theoretical approach and prototype demonstrate how RASHI can be used to assist diplomats, disseminate information to people living under censorship, and how to educate children to be less swayed by propaganda.

I demonstrated the critical role of story understanding Story understanding techniques such as concept recognition, knowledge patterns of if-then rules for common sense, and WordNet integration allowed for RASHI to explain its decisions and be flexible across subjectivity models. Story understanding leads to a natural implementation of compiling the composition story of textual evidence and then interpreting the composition story according to a subjectivity model, removing the need to rely on human experts. Comparing a story understanding approach and my architectural design to other techniques demonstrated the crucial role that structured text, as compared to bag-of-word approaches, plays in taking systematic bias analysis to the next level.

Overall, through both my theoretical and computational contributions, I demonstrated the power of modeling rhetorical analysis as the process of interpreting an author's writing as goal directed. I showed that by working within a specific domain it is possible to define a concrete, impactful set of building blocks and the tools to use these building blocks effectively. I demonstrated the value in modeling rhetorical analysis as the process of gathering objective textual evidence and then of analyzing it with respect to an author's subjective frame of mind. Most importantly, I argued that in order to have a rich, human level understanding of communication, we must be able to reason not just about what we say, but how we say it.

Bibliography

[Berwick and Chomsky, 2015] Berwick, R. C. and Chomsky, N. (2015). *Why Only Us: Language and Evolution*. The MIT Press.

[Chong and Druckman, 2007] Chong, D. and Druckman, J. N. (2007). Framing Theory. *Annual Review of Political Science*, 10(1):103–126.

[Davey and Longuet-Higgins, 1978] Davey, A. C. and Longuet-Higgins, H. C. (1978). *A Computational Model of Discourse Production*, pages 125–136. Springer US, Boston, MA.

[Entman, 1993] Entman, R. M. (1993). Framing: Toward Clarification of a Fractured Paradigm. *Journal of Communication*, 43(4):51–58.

[Fellbaum, 1998] Fellbaum, C., editor (1998). *WordNet: An Electronic Lexical Database*. MIT Press, Cambridge, MA.

[Hallahan, 1999] Hallahan, K. (1999). Seven Models of Framing: Implications for Public Relations. *Journal of Public Relations Research*, 11(3):205–242.

[Jackendoff, 1983] Jackendoff, R. (1983). *Semantics and Cognition*. Current studies in linguistics series. MIT Press.

[Katz, 1997] Katz, B. (1997). Annotating the World Wide Web using natural language. In *Computer-Assisted Information Searching on Internet*, pages 136–155. Le Centre De Hautes Etudes Internationales D'informatique Documentaire.

[Langley, 2017] Langley, P. (2017). A Cognitive Systems Analysis of Personality and Conversational Style. In *Proceedings of the Fifth Annual Conference on Advances in Cognitive Systems*, Troy, New York. Cognitive Systems Foundation.

[Liu and Singh, 2004] Liu, H. and Singh, P. (2004). Conceptnet – A Practical Commonsense Reasoning Tool-Kit. *BT Technology Journal*, 22(4):211–226.

[Minsky, 1986] Minsky, M. (1986). *The Society of Mind*. Simon & Schuster, Inc., New York, NY, USA.

[Morris and Peng, 1994] Morris, M. and Peng, K. (1994). Culture and cause: American and chinese attributions for social and physical events. *Journal of Personality and Social Psychology - PSP*, 67:949–971.

[Pang et al., 2002] Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up? Sentiment Classification using Machine Learning Techniques. *CoRR*, cs.CL/0205070.

[Radford et al., 2017] Radford, A., Józefowicz, R., and Sutskever, I. (2017). Learning to Generate Reviews and Discovering Sentiment. *CoRR*, abs/1704.01444.

[Radul and Sussman, 2009] Radul, A. and Sussman, G. J. (2009). The art of the propagator. In *Proceedings of the 2009 international lisp conference*, pages 1–10.

[Sayan, 2014] Sayan, S. (2014). Audience Aware Computational Discourse Generation for Instruction and Persuasion. Master's thesis, Electrical Engineering and Computer Science Department, MIT, Cambridge, MA.

[Socher et al., 2013] Socher, R., Perelygin, A., Wu, J., Chuang, J., Manning, C. D., Ng, A. Y., and Potts, C. (2013). Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1631–1642, Stroudsburg, PA. Association for Computational Linguistics.

[Taylor, 2018] Taylor, A. (2018). Computational Recognition and Comprehension of Humor in the Context of a General Error Investigation System. Master's thesis.

[Touri and Koteyko, 2015] Touri, M. and Koteyko, N. (2015). Using corpus linguistic software in the extraction of news frames: Towards a dynamic process of frame analysis in journalistic texts. *International Journal of Social Research Methodology*, 18(6):601–616.

[Winston, 2011] Winston, P. H. (2011). The strong story hypothesis and the directed perception hypothesis. *AAAI*.

[Winston, 2014] Winston, P. H. (2014). The Genesis Story Understanding and Story Telling System: A Twenty-First Century Step toward Artificial Intelligence. Memo 019, Center for Brains Minds and Machines, MIT.

[Winston, 2015] Winston, P. H. (2015). Genesiss implementation substrate. `https://groups.csail.mit.edu/genesis/Documentation/frames.pdf`.

[Winston, 2019] Winston, P. H. (2019). Introductory remarks to essay 6. In Solomon, C. and Xiao, X., editors, *Inventive Minds: Marvin Minsky on Education*. MIT Press.

[Winston and Holmes, 2018] Winston, P. H. and Holmes, D. (2018). The Genesis Enterprise: Taking artificial intelligence to another level via a computational account of human story understanding. MIT DSpace, Computational Models of Human Intelligence Community, CMHI Report Number 1. URL `http://hdl.handle.net/1721.1/119651`.