

**Of Intent and Action: Implementing Personality Traits for  
Storytelling Through Concept Patterns**

by

Susan S. Song

B.S., Massachusetts Institute of Technology, 2012

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

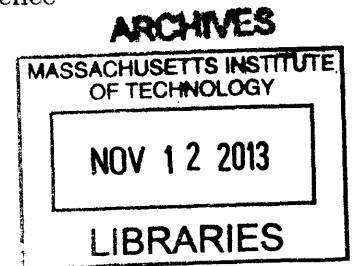
Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2012

© Susan S. Song, MMXII. All rights reserved.



The author hereby grants to MIT permission to reproduce and to distribute publicly  
paper and electronic copies of this thesis document in whole or in part in any  
medium now known or hereafter created.

Author .....  
Department of Electrical Engineering and Computer Science  
September 18, 2012

Certified by .....  
Patrick H. Winston  
Ford Professor of Engineering  
Thesis Supervisor

Accepted by .....  
Professor Dennis M. Freeman  
Chairman, Masters of Engineering Thesis Committee



# Of Intent and Action: Implementing Personality Traits for Storytelling Through Concept Patterns

by  
Susan S. Song

Submitted to the Department of Electrical Engineering and Computer Science  
on September 18, 2012, in partial fulfillment of the  
requirements for the degree of  
Master of Engineering in Electrical Engineering and Computer Science

## Abstract

Personality traits such as “kind,” “aggressive,” and “brave” are integral to storytelling because they impart succinct descriptors of character personalities. Authors apply traits to characters, readers infer characters’ traits from the narrative, and readers learn the meaning of new traits. For instance, a reader can learn the personality trait “vindictive” from Alexandre Dumas’s novel *The Count of Monte Cristo* and then use this trait to predict or explain a character’s behavior. The reader can also infer that a character from this novel, such as Edmond Dantès, is “vindictive” without needing Dumas to explicitly describe the character with this trait.

With the goal of enabling computational storytelling systems to perform the abilities stated above, I present in this thesis a concept pattern-based approach to representing intentional personality traits. I articulate the processes of trait learning, application, and inference and provide steps and insights to how these processes can be computationally implemented. I also give examples of ten personality traits represented using concept patterns inside the Genesis system and show how these traits are discovered inside well-known historical narratives and works of fiction.

Thesis Supervisor: Patrick H. Winston  
Title: Ford Professor of Engineering



## **Acknowledgments**

I would like to thank my friends at MIT for a five-year adventure exceeding my wildest dreams. Though our paths have diverged for now, I look forward to when they meet again.

I am eternally obliged to my thesis advisor, Professor Patrick Winston for his mentorship and insightful discussions. Thank you for always keeping your doors open.

Finally, I would like to thank my family. My journey through MIT and beyond would not have been possible without your undying love.



# Contents

- 1 Vision** **13**
  
- 2 Genesis** **15**
  - 2.1 Overview . . . . . 15
  - 2.2 Representation . . . . . 15
  - 2.3 Events and Elaboration Graphs . . . . . 15
  - 2.4 Rule-Based Story Analysis . . . . . 17
    - 2.4.1 Commonsense Rules . . . . . 17
    - 2.4.2 Concept Patterns . . . . . 19
  
- 3 Personality Traits** **23**
  - 3.1 Definition . . . . . 23
  - 3.2 Intentional Traits . . . . . 24
  
- 4 Trait Representation** **25**
  - 4.1 Overview . . . . . 25
  - 4.2 Example Traits . . . . . 26
    - 4.2.1 Vindictive . . . . . 26
    - 4.2.2 Aggressive . . . . . 29
    - 4.2.3 Brave . . . . . 31
    - 4.2.4 Kind . . . . . 38
    - 4.2.5 Generous . . . . . 41
    - 4.2.6 Shy . . . . . 45
    - 4.2.7 Mean . . . . . 51
  
- 5 Trait Learning** **61**
  - 5.1 Overview . . . . . 61
  - 5.2 Previous Work . . . . . 61
    - 5.2.1 Structural Alignment . . . . . 61
    - 5.2.2 Similarity Module . . . . . 62
    - 5.2.3 Near-Miss Learning . . . . . 63
  - 5.3 How to Learn Traits . . . . . 64

<b>6</b>	<b>Trait Application</b>	<b>69</b>
6.1	Overview . . . . .	69
6.2	How to Predict Behaviors . . . . .	69
6.2.1	Predicting From Trait Patterns . . . . .	70
6.2.2	Predicting From Trait and Trait-Action Patterns . . . . .	71
6.3	How to Explain Behaviors . . . . .	72
<b>7</b>	<b>Trait Inference</b>	<b>75</b>
7.1	Overview . . . . .	75
7.2	Inference From Trait Pattern and Trait-Action Pattern . . . . .	76
<b>8</b>	<b>Contributions</b>	<b>79</b>
<b>A</b>	<b>Concept Patterns</b>	<b>81</b>
A.1	Vindictive . . . . .	81
A.2	Aggressive . . . . .	82
A.3	Brave . . . . .	83
A.4	Kind . . . . .	84
A.5	Generous . . . . .	85
A.6	Shy . . . . .	86
A.7	Mean . . . . .	87
<b>B</b>	<b>Stories</b>	<b>89</b>
B.1	Vindictive . . . . .	89
B.2	Aggressive . . . . .	91
B.3	Brave . . . . .	92
B.4	Kind . . . . .	93
B.5	Generous . . . . .	94
B.6	Mean . . . . .	95



# List of Figures

2-1	Example of an elaboration graph from <i>The Count of Monte Cristo</i> . . . . .	16
2-2	Elaboration graph for “John gives Mary flowers because John likes Mary” .	16
2-3	Elaboration graph for story “Et tu Brute” . . . . .	17
2-4	Elaboration graph for story “Et tu Brute 2” . . . . .	18
2-5	Elaboration graph for story “Et tu Brute 3.” “Caesar becomes dead” did not instantiate. . . . .	18
2-6	Elaboration graph for story “Antony avenges Caesar” with the “avenge” pattern activated . . . . .	20
2-7	Elaboration graph for expanded story “Antony avenges Caesar 2” with the “avenge” pattern activated . . . . .	21
4-1	Elaboration graph for “Caesar and Ptolemy” . . . . .	28
4-2	“vindictive” pattern found in “Caesar and Ptolemy” . . . . .	29
4-3	“revenge - attempt” pattern found in “Caesar and Ptolemy” . . . . .	30
4-4	Elaboration graph for “Hitler’s Preemptive Strike” . . . . .	31
4-5	First “aggressive” pattern found in “Hitler’s Preemptive Strike” . . . . .	32
4-6	Second “aggressive” pattern found in “Hitler’s Preemptive Strike” . . . . .	32
4-7	First “intimidation” pattern found in “Hitler’s Preemptive Strike” . . . . .	33
4-8	Second “intimidation” pattern found in “Hitler’s Preemptive Strike” . . . . .	33
4-9	Third “intimidation” pattern found in “Hitler’s Preemptive Strike” . . . . .	34
4-10	Fourth “intimidation” pattern found in “Hitler’s Preemptive Strike” . . . . .	34
4-11	Elaboration graph for “Hector of Troy” . . . . .	37
4-12	“outward bravery” pattern found in “Hector of Troy” . . . . .	38
4-13	First “physical sacrifice” pattern found in “Hector of Troy” . . . . .	39
4-14	Second “physical sacrifice” pattern found in “Hector of Troy” . . . . .	39
4-15	Elaboration graph for “Little Princess” . . . . .	41
4-16	“kind” pattern found in “Little Princess” . . . . .	42
4-17	First “kind act” pattern found in “Little Princess” . . . . .	42
4-18	Second “kind act” pattern found in “Little Princess” . . . . .	43
4-19	“sharing” pattern found in “Little Princess” . . . . .	43
4-20	Elaboration graph for “Gift of the Magi” . . . . .	46

4-21	First “generous” pattern found in “Gift of the Magi” . . . . .	47
4-22	Second “generous” pattern found in “Gift of the Magi” . . . . .	48
4-23	First “material sacrifice” pattern found in “Gift of the Magi” . . . . .	49
4-24	Second “material sacrifice” pattern found in “Gift of the Magi” . . . . .	50
4-25	Callously mean / selfish . . . . .	52
4-26	Maliciously mean . . . . .	52
4-27	Comparison of “callously mean / selfish act” and “malicious mean act” patterns	53
4-28	Elaboration graph for “Great Expectations . . . . .	55
4-29	“callously mean / selfish” pattern found in “Great Expectations” . . . . .	55
4-30	“maliciously mean” pattern found in “Great Expectations” . . . . .	56
4-31	“callously mean / selfish act” pattern found in “Great Expectations” for Miss Havisham and Estella . . . . .	56
4-32	First “maliciously mean act” pattern found in “Great Expectations” for Miss Havisham and Pip . . . . .	57
4-33	Second “maliciously mean act” pattern found in “Great Expectations” for Miss Havisham and Pip . . . . .	57
4-34	First “callously mean / selfish act” pattern found in “Great Expectations” for Miss Havisham and Pip . . . . .	58
4-35	Second “callously mean / selfish act” pattern found in “Great Expectations” for Miss Havisham and Pip . . . . .	58
4-36	“callously mean / selfish act” pattern encompasses “maliciously mean act” pattern . . . . .	58
4-37	Modified “callously mean / selfish act” pattern to dissociate it from “maliciously mean act” pattern . . . . .	59
5-1	Alignment of story elements . . . . .	62
5-2	Alignment of concept patterns . . . . .	62
5-3	Arch . . . . .	63
5-4	Not an arch . . . . .	64
5-5	Elaboration graph for “Vindictive - Ferdinand harms Edmond” . . . . .	65
5-6	Learning “vindictive”: direct causation . . . . .	65
5-7	Elaboration graph for “Vindictive - Danglars harms Edmond” . . . . .	65
5-8	Structural alignment of “Vindictive - Ferdinand harms Edmond” and “Vindictive - Danglars harms Edmond” . . . . .	66
5-9	Learning “vindictive”: leads-to relationship . . . . .	66
5-10	Elaboration graph for “Edmond un-vindictively harms Albert” . . . . .	66
5-11	Inferring the underlying intent from the “revenge” pattern . . . . .	67
6-1	Partitioning of “kindness” trait concept pattern into antecedents and consequents . . . . .	70

6-2	Partitioning of “sharing” trait-action concept pattern into antecedents and consequents . . . . .	72
6-3	Explaining unexpected actions and unexpected results . . . . .	73



# Chapter 1

## Vision

Personality traits are integral to storytelling. For instance, when an author describes a character as “vindictive,” we the reader immediately form a mental model of the character’s disposition: this character is likely to harm someone if that person harms him. In other words, **applying** personality traits helps the reader predict and explain a character’s behavior, thus bridging conceptual gaps and enabling the reader to understand the character’s motivations for his actions. To use the term often articulated in high school English classes, the reader “characterizes” the subject via personality traits.

At the same time, the definition of personality traits must first **learned** by the reader in order for the reader to understand the use of such traits within stories. For instance, what is the exact scope of a “vindictive” trait? If a character is harmed by some other character, does this harmed character need to harm his injurer in return, or is the mere intent to harm enough to qualify him as being vindictive?

Finally, the author does not always label a character with a trait; instead, the reader can **infer** the character’s trait from his actions. To continue the “vindictive” example, if a character engages in repeated instances of revenge, the reader might infer that this character is vindictive. By inferring characters as possessing particular traits based on the narrative, the reader has thereby compiled compact set of keywords to describe the character’s personality (again, the process of “characterization”); in the words of Professor Patrick Winston, “labeling ideas allows us to have power over them.”

Personality traits are consequently key components in storytelling, with storytelling itself being foundational to human intelligence:

1. If we are to understand human intelligence, we must be able to capture the uniquely human ability to tell stories to ourselves and to others. Storytelling enables us to combine ideas in an infinite range of possibilities and then to share these ideas with other people.
2. If we are to tell stories to one another involving characters, we need to be able to describe these characters via succinct keywords in the form of personality traits.

I consequently articulate the following questions as a framework for pursuing this vision:

- **Representation:** how are personality traits represented?
- **Learning:** how are the definition and name of a trait learned in stories?
- **Application:** how does a reader make use of an author's labeling of a character with a trait?
- **Inference:** how can a reader identify a character as having a trait without being explicitly told this by the author?

To answer these questions, I present in this thesis a **concept pattern-based representation** for personality traits. I begin with a short overview of the Genesis system and of how it performs story analysis. I then explain the concept patterns of ten sample personality traits and show their identification in stories I have adapted from well-known novels and historical accounts. I follow these examples with a high-level discussion on how personality traits may be **learned**, **applied**, and **inferred** in stories, accompanied by brief examples of these procedures in action. Finally, the conclusion will articulate the contributions arising from this thesis.

## Chapter 2

# Genesis

### 2.1 Overview

The Genesis system (henceforth shortened to “Genesis”), developed by Professor Patrick Winston and generations of students from his Genesis group, is a comprehensive story analysis system. The underlying philosophy motivating the project is driven by the belief that computational mechanisms for processing stories should mirror how the human mind conceptually processes stories. In other words, human-like reasoning and not statistical methods must be the foundation for story understanding.

At the high level, Genesis takes a story written in plaintext English, translates the story to Genesis’s internal representation, and then performs commonsense and conceptual analysis. The processed story can then be displayed through visual representations such as a connected graph of plot elements.

### 2.2 Representation

Genesis possesses two layers of representation: the atomic layer and the event, relation, and characterization layer. First, Genesis uses START to produce a syntactic parse from the plaintext story, after which Genesis directly performs a semantic analysis of the parsed story to extract the various events, relations and characterizations. The representation for these events, relations and characterizations are built upon a foundation of atomic things, derivatives, relations and sequences found in the basic atomic layer. Furthermore, each thing, derivative, relation, and sequence belongs to a class hierarchy expressed as a list of superclasses called a “thread,” after Vaina and Greenblatt [15].

### 2.3 Events and Elaboration Graphs

Once plaintext stories are transformed into the representation described above, Genesis displays the story in an **elaboration graph** as seen in Figure 2-1. Each box in the graph

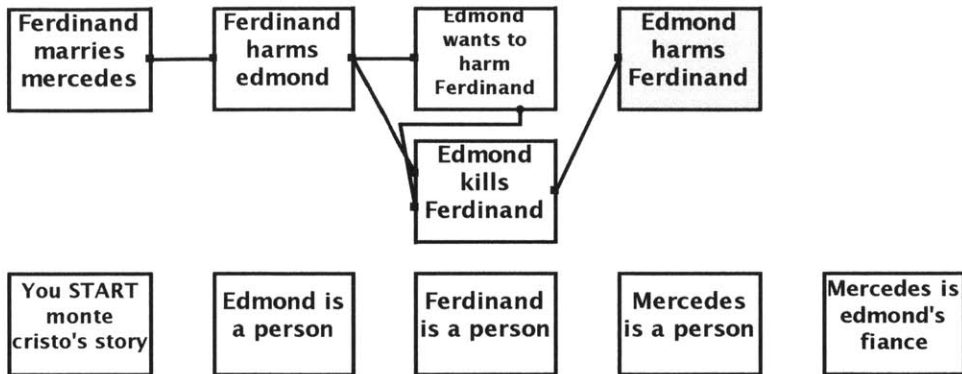


Figure 2-1: Example of an elaboration graph from *The Count of Monte Cristo*

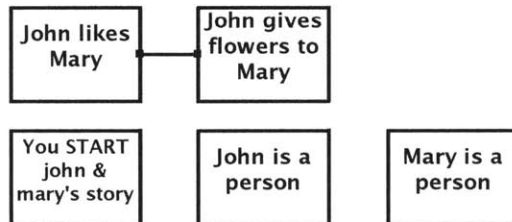


Figure 2-2: Elaboration graph for “John gives Mary flowers because John likes Mary”

represents a story element, where the elements encapsulate an event in the story. Events are connected to consequent events using wires with the story temporally progressing from left to right.

Events can be explicitly connected in elaboration graphs using the conjunction “because.” For instance, in the short story “John gives Mary flowers because John likes Mary,” the **antecedent** clause “John likes Mary” connects to the **consequent** clause “John gives Mary flowers” in the elaboration graph shown in Figure 2-2.

Nevertheless, the consequent often does not need to be explicitly expressed. For instance, if I state, “Brutus killed Caesar,” then the reader understands that “Caesar becomes dead” follows; explicitly stating “Caesar becomes dead because Brutus killed Caesar” becomes redundant. Similarly, if an author states “Caesar becomes dead” later in the story, the reader should be able to mentally link this consequent to its cause (“Brutus killed Caesar”), even if this antecedent happened earlier in the plot. In other words, when human read stories, we are able to draw inferences and conclusions from described events without needed them to be explicitly stated by the author. Genesis implements this inference behavior in the form of **rules** and **concept patterns**.



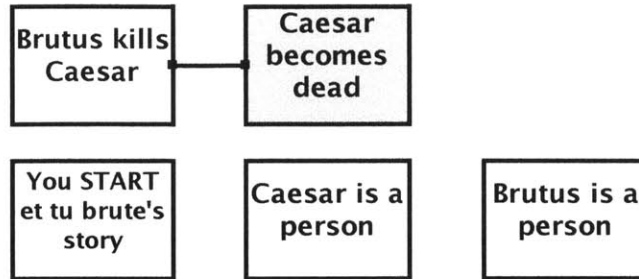


Figure 2-3: Elaboration graph for story “Et tu Brute”

## 2.4 Rule-Based Story Analysis

Humans use an extensive accumulated body of commonsense knowledge to make basic inferences while performing conceptual analysis to discover high-level patterns. Genesis performs commonsense analysis through **commonsense rules** and conceptual analysis through **concept patterns**.

### 2.4.1 Commonsense Rules

Commonsense rules introduce causal links between connected events. Genesis differentiates between two types of rules: **prediction rules** and **explanation rules**.

**Prediction rules** have the structure “if..., then...” where the “if...” is the antecedent and the “then...” is the consequent. Once these rules are articulated for a story, if at anytime during the story analysis process Genesis discovers a story element matching the antecedent, the consequent automatically triggers and add itself to the story representation. For instance, if I tell Genesis the following plaintext story (note the syntax for declaring commonsense rules and for beginning the story):

```

Start commonsense knowledge.
XX is a person.
YY is a person.
If XX kills YY, then YY becomes dead.
  
```

```

Start story titled ‘‘Et tu Brute’’.
Caesar is a person.
Brutus is a person.
Brutus kills Caesar.
The end.
  
```

Genesis then produces the elaboration graph shown in Figure 2-3. The elaboration graph contains the plot element “Caesar becomes dead” and connects the element’s box to its antecedent “Brutus kills Caesar.”

Of course, if we explicitly insert “Caesar becomes dead” into the story, then the prediction rule merely links the two plot elements as Genesis does not need to separately

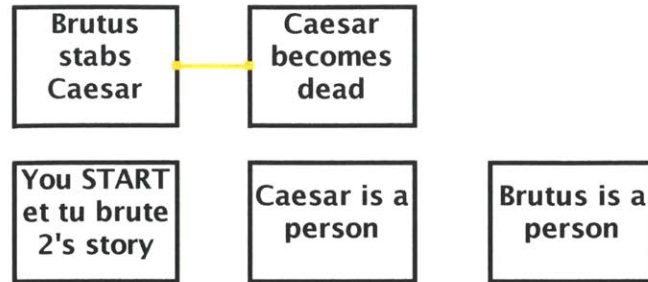


Figure 2-4: Elaboration graph for story “Et tu Brute 2”



Figure 2-5: Elaboration graph for story “Et tu Brute 3.” “Caesar becomes dead” did not instantiate.

instantiate the “Caesar becomes dead” element.

As previously stated, prediction rules always fire when the antecedent has been matched; these rules therefore *predict* consequent events in the story.

**Explanation rules** on the other hand *explain* events and infer connections that are otherwise unstated in the story. Explanation rules take the structure of “if . . . , then XX may . . .” The “may” keyword is crucial: “may” implies that when the antecedent “if . . .” appears in the story, the consequent “then . . .” could *potentially* follow. Thus if the consequent itself appears in the story, its appearance is then *explained* by the antecedent.

To illustrate the difference between explanation rules and prediction rules, if tell the modified story,

```
Start commonsense knowledge.
XX is a person.
YY is a person.
If XX stabs YY, then YY may become dead.
```

```
Start story titled ‘‘Et tu Brute 2’’.
Caesar is a person.
Brutus is a person.
Brutus stabs Caesar.
Caesar becomes dead.
The end.
```

the plot elements “Brutus stabs Caesar” and “Caesar becomes dead” are now linked through the explanation rule as seen in Figure 2-4, despite the lack of a conjunctive “because.”

Nevertheless, if I remove the “Caesar becomes dead” line from the story, the result-

ing elaboration graph can be seen in 2-5. “Caesar becomes dead” remains uninstantiated because Genesis does not know for certain whether “Caesar becomes dead” follows from “Brutus stabs Caesar” (perhaps Caesar died because a second character stabbed him).

Prediction rules follow the behavior of explanation rules in the case when both the antecedents and consequents are present in the story. If the consequent is missing, then prediction rule automatically inserts the consequent element. However, explanation rules are necessary when we do not want to automatically trigger the consequent, but instead link the antecedent and consequent when both appear.

## 2.4.2 Concept Patterns

Commonsense rules, as the name implies, are a sort of “knee-jerk” analysis of story elements: the rules predict and explain basic elements within stories. However, human-level understanding of stories obviously go much further. When we read stories, we are able to *reflect* on the events. For instance, if I share the story of Brutus killing Caesar and Marc Antony then killing Brutus because Caesar was Marc Antony’s friend, one might conclude that the second killing is no ordinary killing. Instead, Marc Antony has *avenged* Caesar’s death by killing Brutus.

Genesis enables us to identify such higher-level patterns through the triggering of **concept patterns**. For the above story, I define the following concept pattern for “avenge” to accompany the story text:

```
Start reflective knowledge.
```

```
Start description of ‘‘avenge’’.
```

```
XX is a person.
```

```
YY is a person.
```

```
ZZ is a person.
```

```
ZZ is YY’s friend.
```

```
XX’s killing YY leads to ZZ’s killing XX.
```

```
The end.
```

```
Start story titled ‘‘Antony avenges Caesar’’.
```

```
Caesar is a person.
```

```
Brutus is a person.
```

```
Antony is a person.
```

```
Antony is Caesar’s friend.
```

```
Brutus kills Caesar.
```

```
Antony kills Brutus because Brutus killed Caesar.
```

```
The end.
```

The elaboration graph then reveals that Genesis has found the pattern for “avenge” in the body of the story as seen in Figure 2-6.

Concept patterns give us additional power beyond commonsense rules in story analysis. First, a concept pattern can contain a collection of antecedents, such as in the form of “ZZ

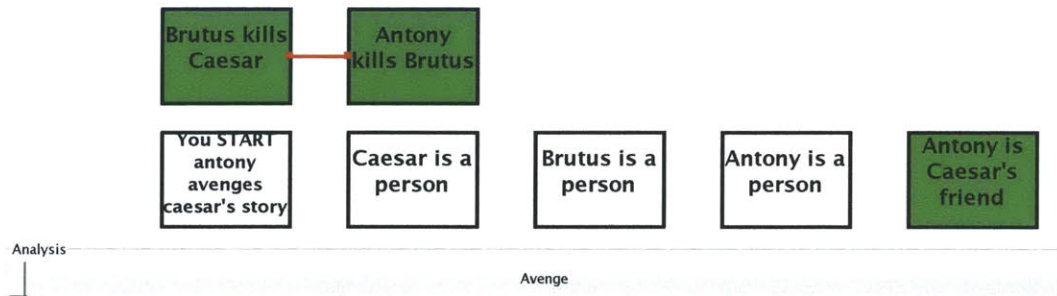


Figure 2-6: Elaboration graph for story “Antony avenges Caesar” with the “avenge” pattern activated

is YY’s friend” and “XX’s killing YY.” Though the same could be done with commonsense rules, increasing the number of antecedents (i.e. “if ZZ is YY’s friend and XX kills YY then...”) greatly burdens the rule matcher and reduces readability.

More importantly, concept patterns enables leads-to relationships through the “leads to” keyword. Whereas commonsense rules require that the antecedent and consequent are directly linked, especially for the triggering of explanation rules, “leads to” enables the concept pattern matcher to recognize the concept even if additional elements are present between the antecedent and consequent. As an example, consider the expanded “Antony avenges Caesar 2” story:

```
Start story titled ‘‘Antony avenges Caesar 2’’.
Caesar is a person.
Brutus is a person.
Antony is a person.
```

```
Antony is Caesar’s friend.
Brutus kills Caesar.
Antony fights Brutus because Brutus killed Caesar.
Antony kills Brutus because Antony fought Brutus.
The end.
```

If we keep the “avenge” concept pattern unchanged (reproduced below),

```
Start description of ‘‘avenge’’.
XX is a person.
YY is a person.
ZZ is a person.
ZZ is YY’s friend.
XX’s killing YY leads to ZZ’s killing XX.
The end.
```

we end up with the graph and activation of “avenge” in Figure 2-7.

The activation of “Antony fights Brutus,” the story element in between “Brutus kills Caesar” and “Antony kills Brutus,” indicates that Genesis also considers this element to be part of the “avenge” pattern. Thus as the “leads to” keyword implies, any action that follows the antecedent and eventually leads to the desired consequent is included in the

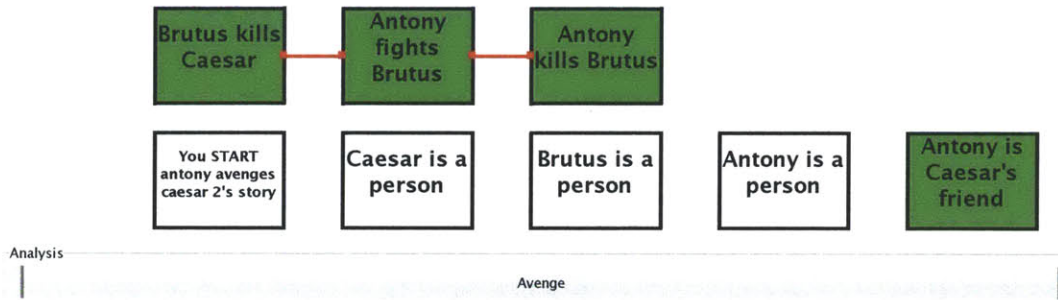


Figure 2-7: Elaboration graph for expanded story “Antony avenges Caesar 2” with the “avenge” pattern activated

concept pattern’s discovery. Concept patterns, with its flexibility and leads-to relations will become powerful mechanisms for expressing personality traits.



## Chapter 3

# Personality Traits

### 3.1 Definition

I have already made extensive use of the term “personality trait,” but just what exactly is a personality trait? Trait theorists from psychology define “traits” as habitual patterns of behavior, thought, and emotion [6], and measurement of these traits form a model for human personalities. While the strength of these traits different across individuals, they are stable influencers of outward behavior for each individual.

Psychologists have constructed numerous standardized inventories to measure and evaluate personality traits, such as the Five Factor Model (“openness,” “conscientiousness,” “extraversion,” “agreeableness,” “neuroticism”)[6], but the terminology employed in these tests hardly reflects the richness of everyday trait vocabulary. In both stories and conversations, traits are used with casual abandon to describe people. Take, for instance, the following excerpt from Charles Dicken’s *Great Expectations*:

She was not a good-looking woman, my sister; and I had a general impression that she must have made Joe Gargery marry her by hand. Joe was a fair man, with curls of flaxen hair on each side of his smooth face, and with eyes of such a very undecided blue that they seemed to have somehow got mixed with their own whites. He was a mild, good-natured, sweet-tempered, easy-going, foolish, dear fellow,—a sort of Hercules in strength, and also in weakness. [1]

With words such as “sweet-tempered” and “good-natured,” Dickens has directly given the reader a summary of Joe Gargery’s personality without resorting to a psychological evaluation of the sort, “Joe Gargery scores 10 in neuroticism and 70 in agreeableness.” Clearly the set of personality traits employed by psychologists differs from that used by a person in casual settings. Due to my interest in studying the role of personality traits in human storytelling, I have chosen to focus on these informal traits used to describe people.

However, because these traits are used informally in stories, their exact definitions are difficult to pin down. Even dictionaries do not give explicit, non-circular definitions for

traits. For instance, Oxford Dictionary’s definition for “mean” includes the following entries [12]:

- unwilling to give or share things, especially money; not generous
- unkind, spiteful, or unfair

We clearly see that in this example definition of “mean,” dictionaries define traits using more traits! Additionally, many traits have a “suitcase” quality to them (“suitcase words” is a term first defined by Marvin Minsky); they encompass a wide variety of meanings that change depending on the context. The word “boring” is one such example: a “boring” person could mean that he does not say anything of interest to the audience or that the person has not had any “interesting” life experiences. Given this ambiguity in the meaning of personality traits used in a casual setting, I have narrowed down my analysis to *intentional* personality traits.

### 3.2 Intentional Traits

Though personality traits are formally defined as habitual patterns of behavior, thought, and emotion, certain traits express habitual “thoughtless” behaviors more than others. “Thoughtless” behaviors involve limited motivations underlying the action. For instance, the trait “lazy” implies that the character is disinclined to doing things, but does it also mean that the character has *intended* to be lazy? Or is laziness merely a pattern of behavior for which the character has relegated himself to with little awareness thereof? Similarly, a character described as “rude” falls into the same intentional-versus-unintentional grey zone: is the person rude because he deliberately upsets others? Or is he rude because he is not aware of the uncouthness of his actions?

These behavioral traits, while enlightening in their own right, shine minimal light into the internal motivations of the character. Instead, they only describe outward patterns of action. **Intentional traits** on the other hand are more salient because they provide insight into the mind of the character. In other words, once we know that a character possesses a personality trait, we can better infer the character’s rationale for his outward actions. Additionally, given a trait and a rationale, we can then predict likely actions that the character will undertake. Thus I have described the components of an intentional trait:

- trait name
- underlying intention
- expressed action (optional)

In the next chapter, I discuss Genesis’s representation for intentional personality traits and give examples of implemented traits.



## Chapter 4

# Trait Representation

### 4.1 Overview

I have chosen to focus on **intentional personality traits** because they describe mental states of characters: the character wants, thinks, or desires some sort of outcome and consequently behaves with the intent of achieving this end goal. I have also articulated a difference between the trait and the specific behaviors arising from this trait, which I will now define as **trait-actions**. Traits express the intentionality of the character and a general course of action that he may want to engage in. Trait-action, on the other hand, is a specific type of action that arises from the trait. For instance, an example trait is “vindictiveness,” while an action arising from this trait is “revenge.” Similarly, while the concept pattern for “generosity” already implicitly includes an action, an explicit type of action that arises from generosity is “sharing.”

Both intentional traits and their outward trait-actions can be described through Genesis’s concept patterns. In my discussion of the Genesis system earlier in this thesis, I described how concept patterns allow the use of leads-to relationships. Leads-to links are crucial to representing intentional traits because a character’s desire for a particular outcome may not immediately result in this outcome. Instead, the character likely engages in intermediary behaviors that will eventually lead him to this end goal.

To illustrate how traits may be represented within the Genesis framework, I will show examples of ten intentional personality traits in concept pattern form. Each trait will also be accompanied by an example trait-action concept pattern and by a story (with the exception of “shyness”) showing how these patterns may be triggered.

Overall, the exercise of representing these traits as concept patterns will reveal many underlying ambiguities to the meanings of these traits. Resolving such ambiguities will therefore yield various debates and insights into the subtleties of the traits, such as in the case of “selfishly mean” versus “maliciously mean.” Furthermore, these debates illustrate why people tend to have differences in opinion when assigning traits to other people—e.g. one person’s “kind” may be another person’s “generous.” Thus not only do the examples

below showcase the idea of using concept patterns to represent traits, but the implementation approach itself forces clearer thinking about the meaning of traits while exploring the different ways in which they may be interpreted by readers.

## 4.2 Example Traits

### 4.2.1 Vindictive

The Oxford Dictionary defines “vindictiveness” as “having or showing a strong or unreasoning desire for revenge” [14]. In this very definition we already see the differentiation between *intention* (“desire”) and *action* (“revenge”). Meanwhile, revenge is “the action of inflicting hurt or harm on someone for an injury or wrong suffered at their hands” [13]. Therefore one can conclude that to describe someone as “vindictive,” that person tends to strongly want to harm someone in retribution. I define the concept pattern for vindictiveness as the following:

```
Start description of ‘‘vindictive’’.
XX is a person.
YY is a person.
YY’s harming XX leads to XX’s wanting to harm YY.
The end.
```

The concept pattern for vindictiveness expresses the idea of “harming leading to wanting to harm” via the “YY’s harming XX leads to XX’s wanting to harm YY” statement.

At the same time, the concept pattern for revenge has already been defined in previous work in Genesis as the following:

```
Start description of ‘‘revenge’’.
XX is a person.
YY is a person.
XX’s harming YY leads to YY’s harming XX.
The end.
```

Note how this pattern does not explicitly define intentionality; instead, intentionality is presumed in the “YY’s harming XX” clause. Nevertheless, we may want to broaden this idea of revenge to allow for “attempting to harm”:

```
Start description of ‘‘revenge - attempt’’.
XX is a person.
YY is a person.
YY’s harming XX leads to XX’s trying to harm YY.
The end.
```

This “revenge - attempt” pattern implicitly describes the previous “revenge” pattern. After all, if someone has harmed someone in vengeance, then it goes unstated that he has also *attempted* to harm that person.

One example story containing examples of both vindictiveness and revenge is the historical case of Pompey the Great’s fleeing to Egypt following his defeat in a civil war against

Julius Caesar. Pompey had thought that Egypt provided a safe refuge, but he greatly misjudged: the young Egyptian pharaoh Ptolemy XIII (the brother of Cleopatra VII) sought to ally himself with Caesar and assassinated Pompey as he disembarked his ship. According to certain accounts, including that from Plutarch, Caesar instead wept upon hearing news of the death of his former son-in-law (Pompey had been married to Caesar's only child, Julia). Though the historical chain of events involved other instigating factors, Caesar eventually fought Ptolemy and forced Ptomley to flee Egypt. In the process, Ptolemy drowns while crossing the Nile.

I encapsulate the above story in the following abridged plaintext version that is readable by Genesis:

Start story titled "Caesar and Ptolemy".

Caesar is a person.

Ptolemy is a person.

Pompus is a person.

Pompus is Caesar's brother-in-law.

Caesar is angry at Pompus.

Ptolemy wants Caesar's favor.

Pompus flees to Egypt because Caesar is angry at Pompus.

Ptolemy kills Pompus because Pompus flees to Egypt and because Ptolemy wants Caesar's favor.

Ptolemy harms Caesar.

Caesar wants to harm Ptolemy because Ptolemy harmed Caesar.

Caesar fights Ptolemy because Caesar wants to harm Ptolemy.

Caesar wins because Caesar fights Ptolemy.

Ptolemy flees because Caesar wins.

Ptolemy drown because Ptolemy flees.

The end.

For the sake of brevity, I have omitted the display of commonsense rules in the above story and in further stories presented in this chapter. Their full text may be found in Appendix B.

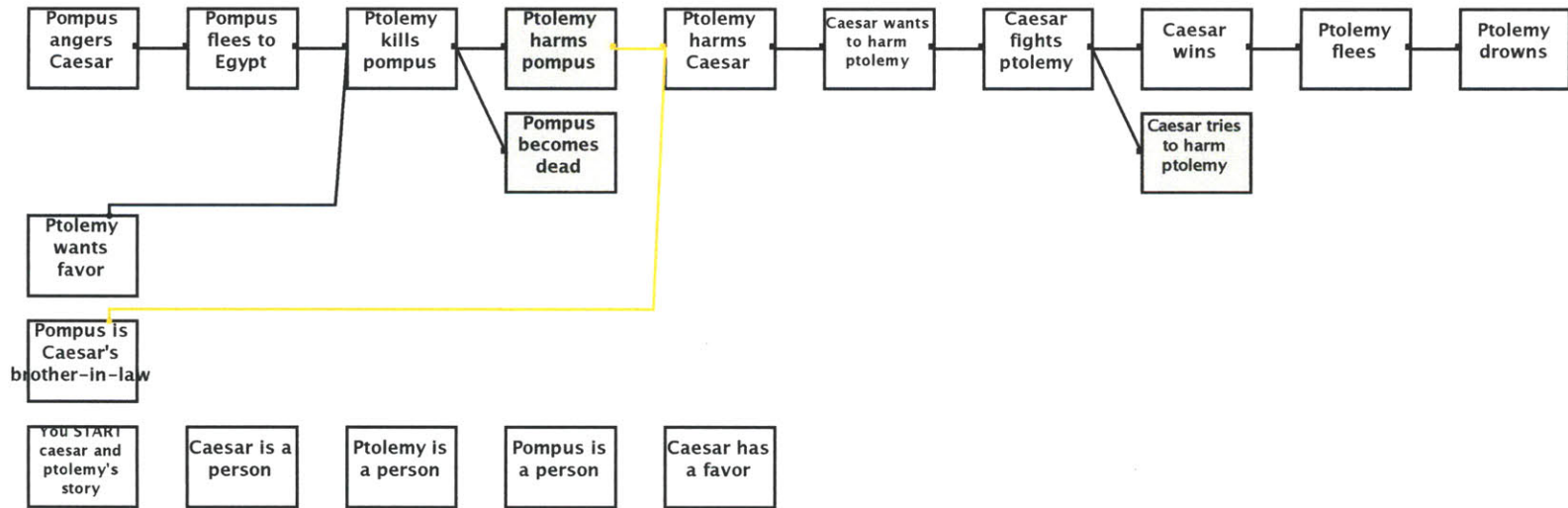


Figure 4-1: Elaboration graph for "Caesar and Ptolemy"

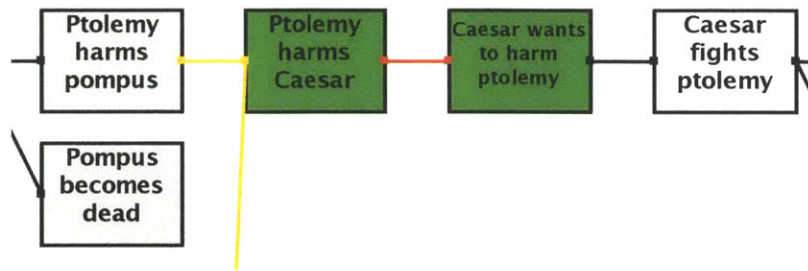


Figure 4-2: “vindictive” pattern found in “Caesar and Ptolemy”

The “Caesar and Ptolemy” story generates the elaboration graph seen in Figure 4-1 and triggers the patterns of “vindictive” (with Caesar as the vindictive character) in Figure 4-2 and “revenge - attempt” in Figure 4-3. This particular “revenge - attempt” pattern has been activated because Caesar did not directly harm Ptolemy. Instead, only during the course of Caesar’s attempt to harm Ptolemy did Ptolemy fall, but even then I argue that Caesar engages in a vengeance-motivated action of fighting Ptolemy.

#### 4.2.2 Aggressive

Returning to the Oxford Dictionary again, this dictionary defines “aggression” as “hostile or violent behavior or attitudes toward another; readiness to attack or confront” [8], with “aggressiveness” being the tendency to engage in such behavior [9]. However, I draw a distinction between behavioral aggressiveness and intentional aggressiveness. With the behavioral version, the person habitually engages in aggressive acts—he is hostile without explicitly seeking any particular ends and may also be easily provoked. In other words, one would describe this character as “hot-tempered.” Intentional aggressiveness on the other hand implies an explicit mental aim on part of the character: he purposely seeks to instill fear or caution through his aggressive actions, perhaps so that he can more easily achieve some ulterior goal.

I thus define intentional aggressiveness in the following concept pattern:

```

Start description of ‘‘aggressive’’.
XX is a person.
YY is a person.
AA is an action.
XX thinks that YY will fear XX if XX performs AA.
XX’s wanting YY to fear XX leads to XX’s performing AA.
The end.

```

I describe the pre-meditation to instill fear in the line “XX thinks that YY will fear XX if XX performs AA” and the outward aggressive behavior in “XX’s wanting YY to fear XX leads to XX’s performing AA.” Unlike vindictiveness, aggressiveness goes beyond a pure desire to harm because the aggressive character actually follows through with harming

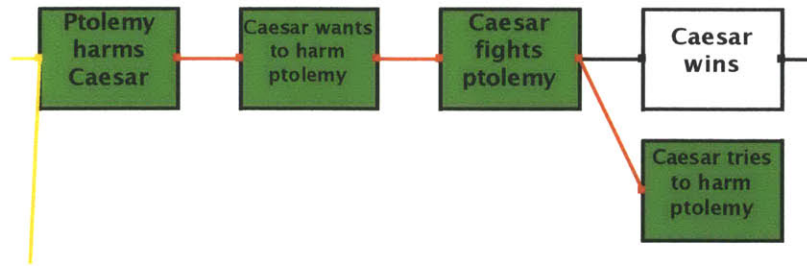


Figure 4-3: “revenge - attempt” pattern found in “Caesar and Ptolemy”

action. However, the above pattern for aggressiveness is still that of a trait and not of a trait-action because it expresses intention. Furthermore, I do not describe whether this act of aggression is a success, only that XX has performed a non-specific action that he *thinks* will lead to YY fearing him.

An instantiated action form of aggressiveness is “intimidation,” where someone’s desire to instill this fear leads to him performing an action that accomplishes this goal:

```
Start description of ‘‘intimidation’’.
XX is a person.
YY is a person.
AA is an action.
XX’s performing AA leads to YY’s fearing XX.
The end.
```

In the line “XX’s performing AA leads to YY’s fearing XX,” I implicitly force the action “AA” to be some sort of act that instills fear as opposed to merely an act that XX thinks will instill fear (i.e. in the case of “aggressive”).

These concept patterns may be seen in the following fictional (but loosely based on history) story of Adolf Hitler, Philippe Pétain (head of the Vichy government during World War II) and Winston Churchill:

```
Start story titled ‘‘Hitler’s Preemptive Strike’’.
Hitler is a person.
Churchill is a person.
Petain is a person.
```

```
Poland is an entity.
```

```
Hitler wants Churchill to fear Hitler because Hitler wants to win the
war.
```

```
Hitler thinks that Petain will fear Hitler if Hitler invades Poland.
Hitler thinks that Churchill will fear Hitler if Hitler invades Poland.
```

```
Hitler invades Poland because Hitler wants Petain to fear Hitler.
Hitler invades Poland because Hitler wants Churchill to fear Hitler.
```

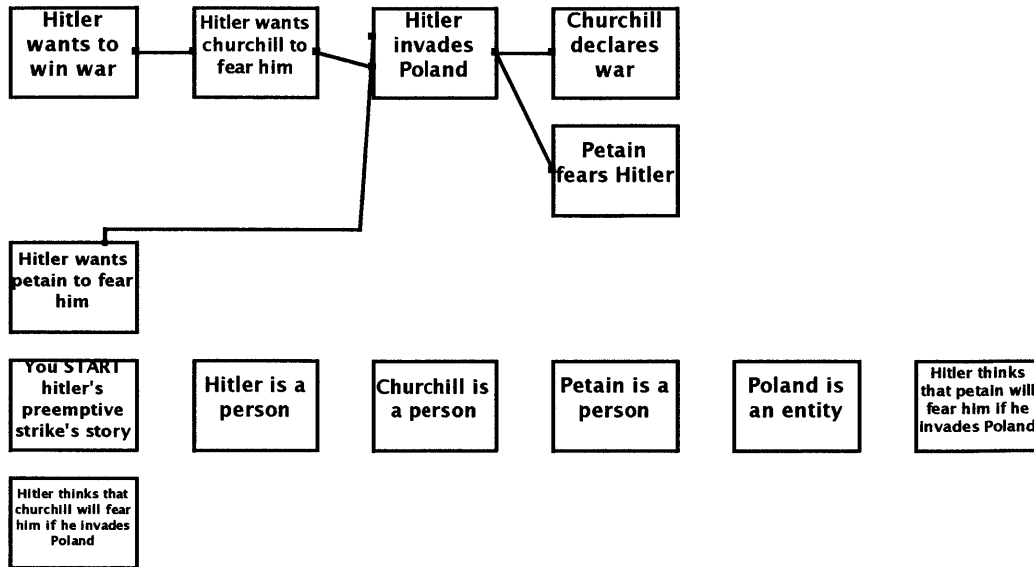


Figure 4-4: Elaboration graph for “Hitler’s Preemptive Strike”

Petain fears Hitler because Hitler invaded Poland.  
 Churchill declares war because Hitler invaded Poland.  
 The end.

Genesis produces the elaboration graph shown in Figure 4-4 and the concept patterns for “intimidate” and “aggressive” shown in Figures 4-5 to 4-10. Interestingly, Genesis finds multiple “intimidation” patterns because of the leads-to relationship. While the first discovery (Figure 4-7) is the most meaningful, Genesis discovers the other three because *any* action AA that results in “Petain’s fearing Hitler” fits the specifications for the intimidation pattern. One possible way to filter out these extraneous patterns would be to set additional specifications for AA. Perhaps AA could be limited to action verbs as opposed to non-action abstract verbs such as “wanting” or “thinking.”

### 4.2.3 Brave

The third set of concept patterns I have defined involves the trait of “bravery.” Both the Oxford Dictionary and Wikipedia discuss bravery in the context of courage and fear: bravery is the willingness to commit some action in spite of the threat or state of pain and fear [10][16]. Wikipedia goes further to differentiate between “physical courage” and “mental courage” (with courage being synonymous with bravery), where the physical variety involves bodily harm while the mental variety involves social harm, such as being ostracized [16]. I have chosen to focus on physical bravery due to Genesis implementation limitations in describing socially-related behavior avoidance. Nevertheless, I further distinguish physical bravery into “outward bravery” and “inward bravery.” “Outward bravery” describes a

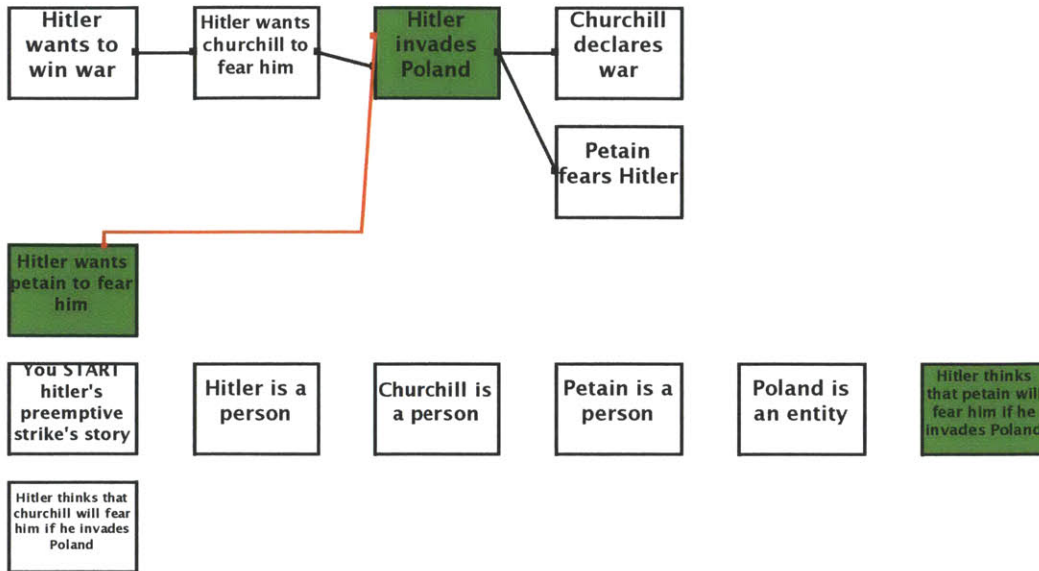


Figure 4-5: First “aggressive” pattern found in “Hitler’s Preemptive Strike”

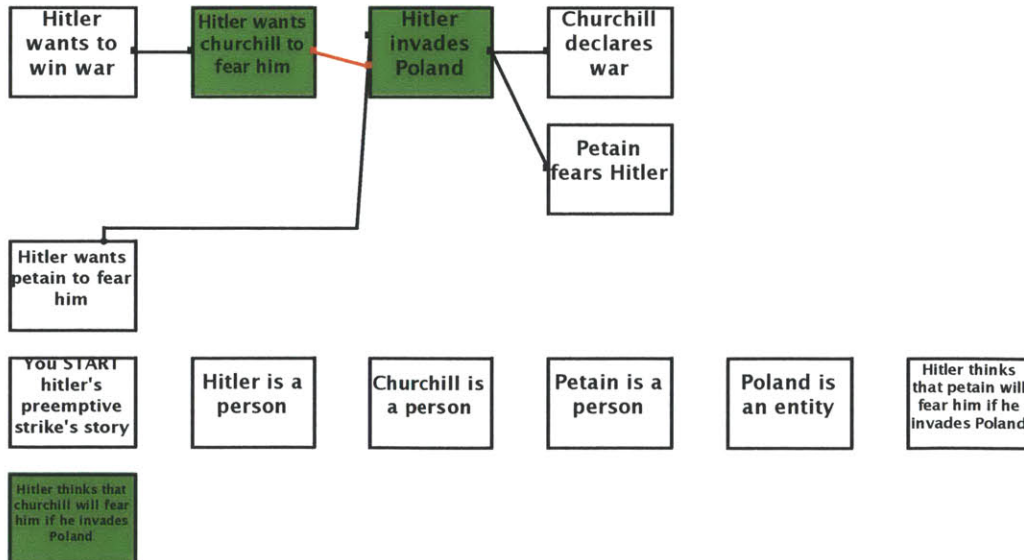


Figure 4-6: Second “aggressive” pattern found in “Hitler’s Preemptive Strike”



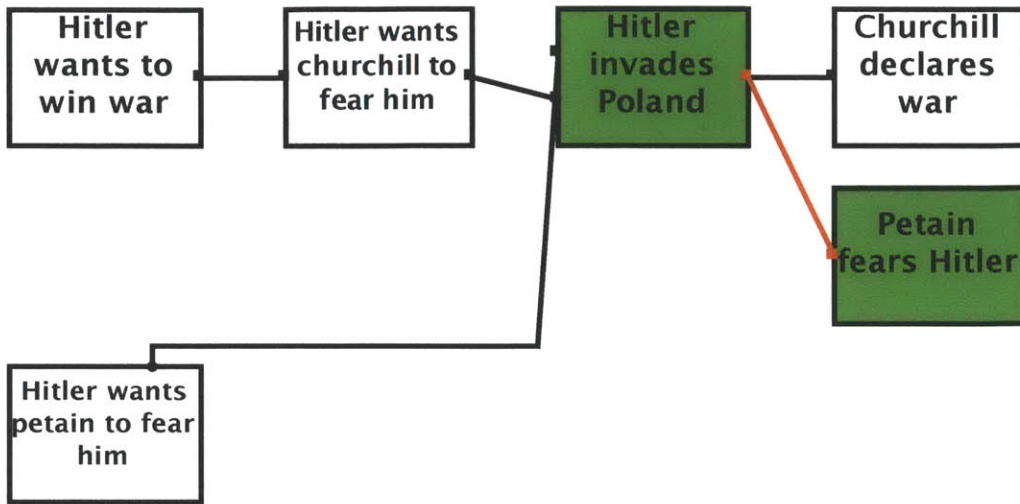


Figure 4-7: First “intimidation” pattern found in “Hitler’s Preemptive Strike”

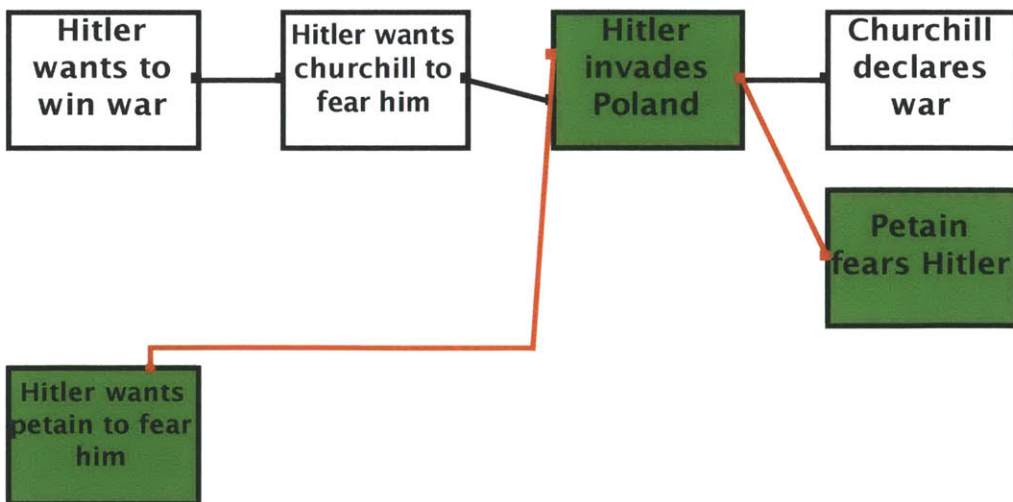


Figure 4-8: Second “intimidation” pattern found in “Hitler’s Preemptive Strike”

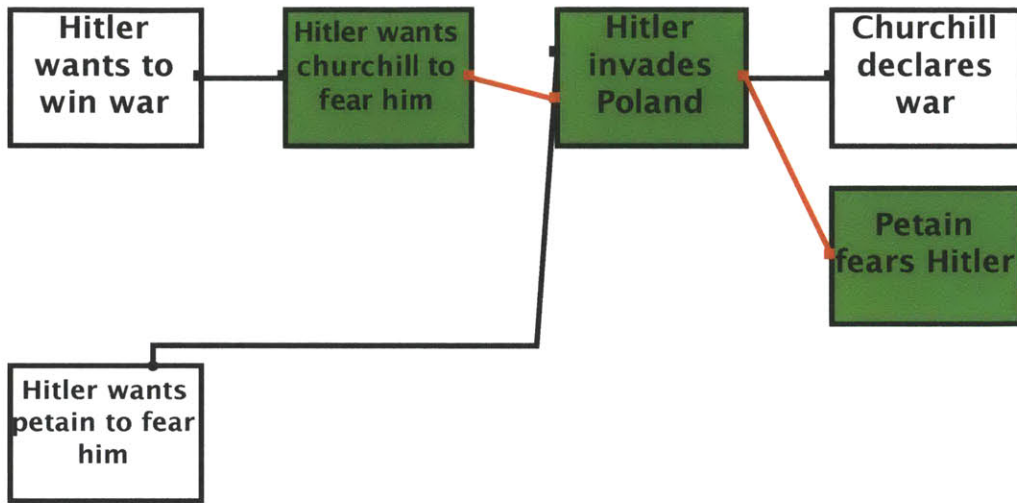


Figure 4-9: Third “intimidation” pattern found in “Hitler’s Preemptive Strike”

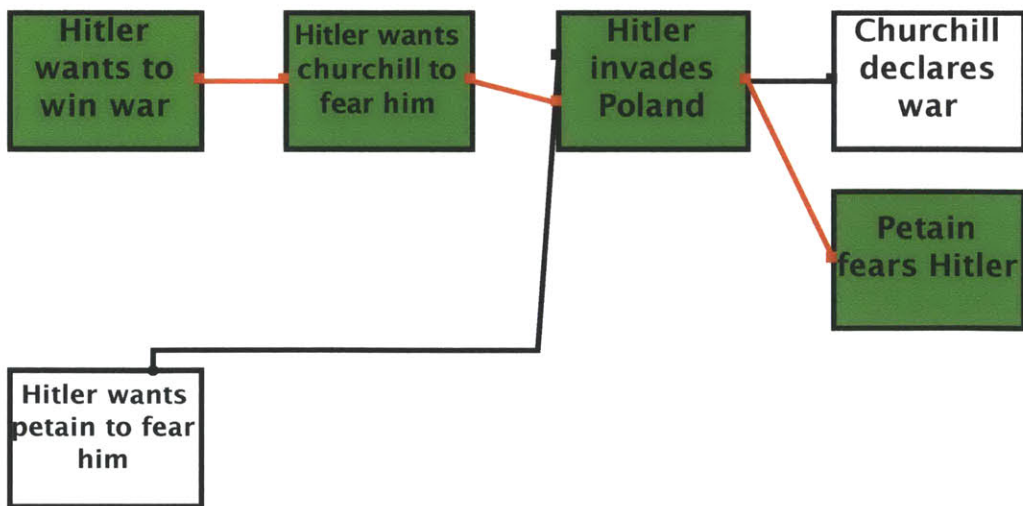


Figure 4-10: Fourth “intimidation” pattern found in “Hitler’s Preemptive Strike”

personality trait where one desires to benefit others through a potentially dangerous action, while an “inwardly” brave character seeks to benefit oneself. Real-life examples of outward bravery would be when firemen save people from burning buildings, for example, while inward bravery could be that of a disabled person conquering physical pain to learn how to walk again.

I describe these types of physical bravery through the following patterns<sup>1</sup>:

Start description of “outward bravery”.

XX is a person.

YY is a person.

AA is an action.

XX thinks that XX will become injured if XX performs AA.

XX thinks that YY will become happy if XX performs AA.

XX’s wanting YY to become happy leads to XX’s performing AA.

The end.

Start description of “inward bravery”.

XX is a person.

YY is a person.

XX thinks that XX will become injured if XX performs AA.

XX thinks that XX will become happy if XX performs AA.

XX’s wanting to become happy leads to XX’s performing AA.

The end.

Similar to “aggressiveness,” these two physical bravery patterns involve a general action accompanying the intention, with the intention described in the two “XX thinks…” lines in each pattern. A particular trait-action then arising from “bravery” is “sacrifice.” In a sacrificial act, the character performs an action that leads to someone else becoming happy while the character injures himself in the process:

Start description of “sacrifice”.

XX is a person.

YY is a person.

AA is an action.

XX’s performing AA leads to XX’s becoming injured.

XX’s performing AA leads to YY’s becoming happy.

The end.

However, is an action sacrificial if XX does not know ahead of time the potential harm arising from AA? I argue that the answer depends on interpretation: to one reader, sacrifice does entail prior knowledge, while to another, only XX’s intention to benefit YY with his action and XX’s unintentionally becoming injured in the process are enough to classify AA as a sacrificial act.

Because of the direct parallels between “outward bravery” and “inward bravery,” I have omitted the inclusion of a story showing the inward version. Nevertheless, to show the “outward bravery” and “sacrifice” concept patterns in action, I adapted Homer’s *The Iliad*

---

<sup>1</sup>Due to a Genesis parsing bug, I have substituted “harmed” for “injured” in the bravery concept patterns and all further patterns where “harmed” is used as an adjective.

poem for Genesis. *The Iliad* tells the story of Hector, Priam, and Achilles, where Hector sacrifices himself by fighting Achilles in an attempt to liberate Troy and please his father Priam (the king of Troy) [5]:

Start story titled ‘‘Hector of Troy’’.

Achilles is a person.

Hector is a person.

Priam is a person.

Troy is an entity.

Troy is under siege.

Priam is unhappy because Troy is under siege.

Hector wants Priam to become happy because Priam is unhappy and because Hector is Priam’s son.

Hector thinks that Hector will become injured if Hector fights Achilles.

Hector thinks that Priam will become happy if Hector fights Achilles.

Hector fights Achilles because Hector wants Priam to become happy.

Hector becomes injured because Hector fights Achilles.

Hector dies because Hector becomes injured.

Priam becomes happy because Hector fights Achilles.

Priam becomes unhappy because Hector dies.

The end.

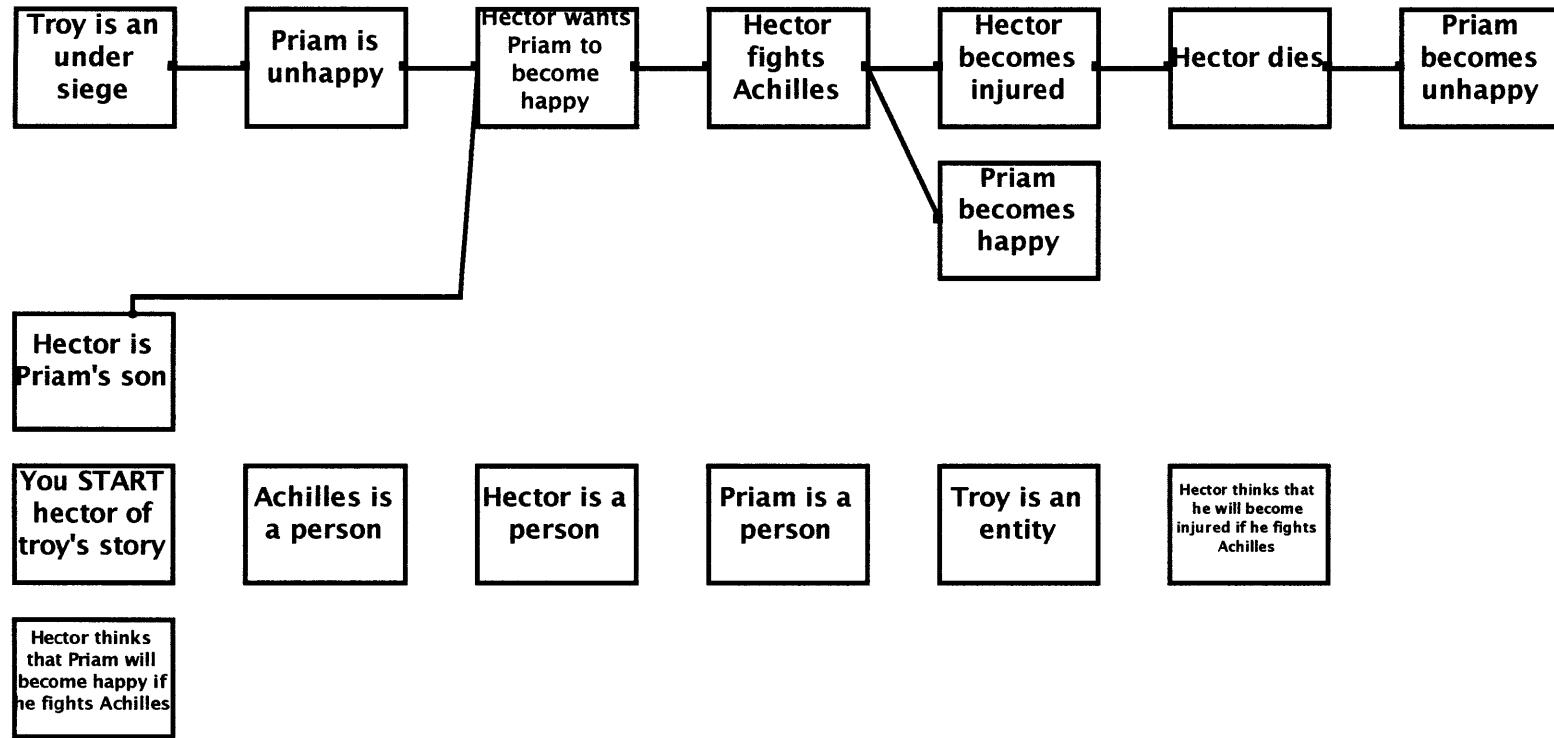


Figure 4-11: Elaboration graph for "Hector of Troy"

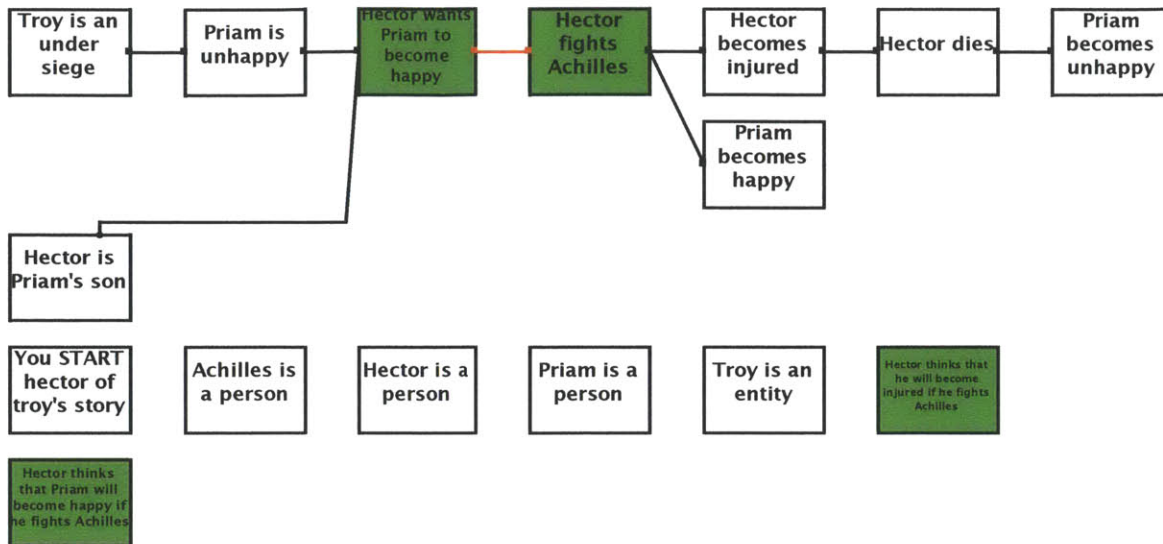


Figure 4-12: “outward bravery” pattern found in “Hector of Troy”

The elaboration graph for the story can be seen in Figure 4-11. Genesis finds the the “outward bravery” (Figure 4-12) and “sacrifice” concept patterns (Figures 4-13, 4-14), as expected. However, similar to the case with “intimidation,” Genesis discovers two “physical sacrifice” patterns due to the mechanics of the leads-to relationship.

#### 4.2.4 Kind

“Kind” is a difficult trait to explain both computationally via concept patterns and verbally in colloquial English. What is exactly is kindness? The Oxford Dictionary describes it as “the quality of being friendly, generous, and considerate” [11]. The definition of kindness, which includes synonyms such as “friendly” and “generous,” implies that kindness has a “suitcase” property, where the word aggregates an entire collection of meanings and takes on different definitions depending on context. In other words, “kindness” could mean “generous” or “friendly” in specific scenarios. Due to this ambiguity, I have settled on the one clear commonality underlying the various definitions of kindness: kindness is the quality of wanting someone to become happy without expecting a reciprocating act.

However, the definition for “kindness” should not discourage cases where the kind person wants himself to be happy while performing a kind act for someone else. Hedonistically-speaking, one can argue that the person performs the kind act not just to make the beneficiary happy, but also to derive a personal happiness from making this beneficiary happy.

I therefore describe “kind” in two separate concept patterns:

```
Start description of ‘‘kind’’.
XX is a person.
YY is a person.
```

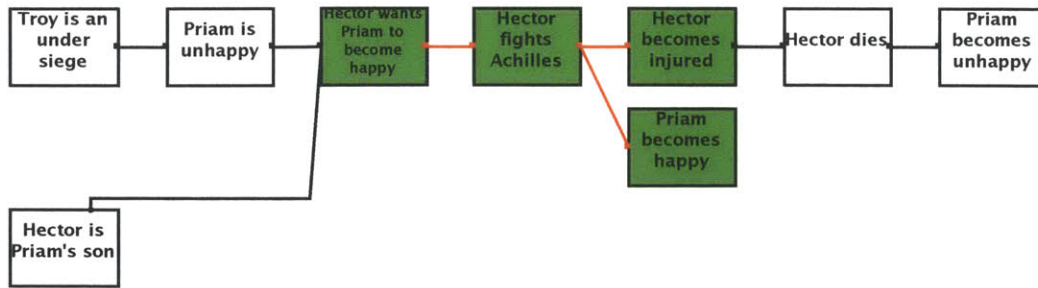


Figure 4-13: First “physical sacrifice” pattern found in “Hector of Troy”

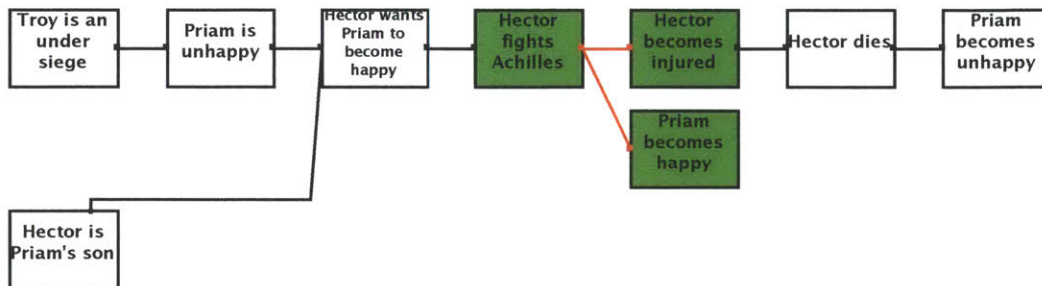


Figure 4-14: Second “physical sacrifice” pattern found in “Hector of Troy”

AA is an action.  
 XX thinks that XX’s performing AA leads to YY’s becoming happy.  
 XX’s wanting YY to become happy leads to XX’s performing AA.  
 The end.

Start description of ‘‘hedonistic kind’’.  
 XX is a person.  
 YY is a person.  
 AA is an action.  
 XX thinks that XX’s performing AA leads to YY’s becoming happy.  
 XX thinks that YY’s becoming happy leads to XX’s becoming happy.  
 XX’s wanting to become happy leads to XX’s performing AA.  
 The end.

This first concept pattern, “kind,” directly expresses the first definition I articulated for kindness: it is the quality of wanting someone else to become happy and then performing an action that the kind person thinks will lead to the beneficiary’s happiness. The second pattern, “hedonistic kind” has a selfish quality to it: it expresses the idea that wanting someone to become happy derives from wanting oneself to become happy. In other words, the motivation for any person’s actions, even if they derive from altruistic intentions, ultimately finds its source in wanting oneself to become happy. This second pattern also allows for a sort of false and manipulative variety of “kind,” where the supposedly kind character performs an act that appears to be for the sake of another character’s kindness, but is

actually for oneself due to the “XX’s wanting to become happy leads to XX’s performing AA” line.

A potential way to disambiguate these different flavors of kindness is to emphasize how much reciprocation or self-directed happiness the character expects: does the character perform AA with the ultimate intention of becoming happy himself? Or does he perform AA with the primary goal of benefitting the other person, with becoming happy himself as an unintended consequence? Such thoughts are unfortunately difficult to express within the Genesis framework but provide insight into the suitcase property of “kind.”

In the meantime, I define the trait-action, “kind act” for “kind” as the following:

```
Start description of ‘‘kind act’’.
XX is a person.
YY is a person.
AA is an action.
XX is YY’s friend.
XX’s performing AA leads to YY’s becoming happy.
The end.
```

The trait-action pattern “kind act” explicitly looks for YY’s becoming happy as a result of XX’s action. However, because Genesis allows XX and YY to be bound to the same entity in the line “XX’s performing AA leads to YY’s becoming happy” in absence of other constraints, I found it necessary to add the line “XX is YY’s friend” to differentiate XX from YY. I introduce this line purely as a workaround, and it can be removed once Genesis has the ability to distinguish XX and YY as separate characters even if the only constraint is “XX’s performing AA leads to YY’s becoming happy.”

Finally, I define a more specific trait-action for “kind” in the form of “sharing”:

```
Start description of ‘‘sharing’’.
XX is a person.
YY is a person.
ZZ is an entity.
XX owns ZZ.
YY’s wanting to use ZZ leads to XX’s allowing YY to use ZZ.
The end.
```

Note that XX’s allowing YY to use ZZ is premised on two antecedents. First, YY must desire the object ZZ. Secondly, XX needs to own the entity ZZ, thereby enabling XX to share it with YY.

I have adapted Frances Hodges Burnett’s *A Little Princess* to show these two patterns and “kind act” trait-action in Genesis [11]. Due to matcher limitations, I express Sara and Becky’s interaction with the doll (the entity named “Emily”) as “used” instead of “play” to enable matching with the desired concept patterns. Also, as previously noted, I found it necessary to explicitly state “Sara is Becky’s friend” as an additional constraint.

```
Start story titled ‘‘Little Princess’’.
Sara is a person.
Becky is a person.
```



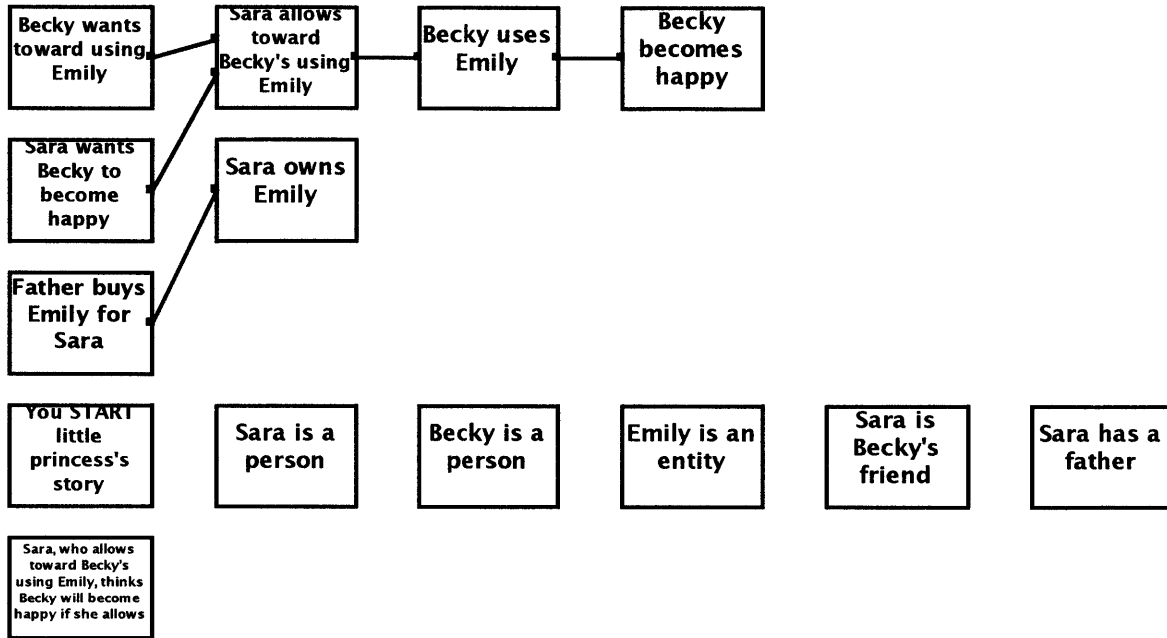


Figure 4-15: Elaboration graph for "Little Princess"

Emily is an entity.

Sara is Becky's friend.

Sara owns Emily because Sara's father bought Emily for Sara.  
Becky wants to use Emily.

Sara thinks that Becky will become happy if Sara allows Becky to use Emily.  
Sara allows Becky to use Emily because Becky wants to use Emily.  
Sara allows Becky to use Emily because Sara wants Becky to become happy.

Becky uses Emily because Sara allows Becky to use Emily.  
Becky becomes happy because Becky uses Emily.  
The end.

Genesis produces the elaboration graph for "Little Princess" as seen in Figure 4-15, with the trait "kind" being discovered in 4-16. Furthermore, Genesis finds two instances of "kind act" (Figures 4-17, 4-18) and one instance of "sharing" (Figure 4-19).

#### 4.2.5 Generous

The suitcase-like definitions of kindness and sharing led me to explore "generosity" in further detail. Generosity is a trait used to describe someone who willingly shares his time, material goods, affections—in other words, something one is capable of giving—to people who either

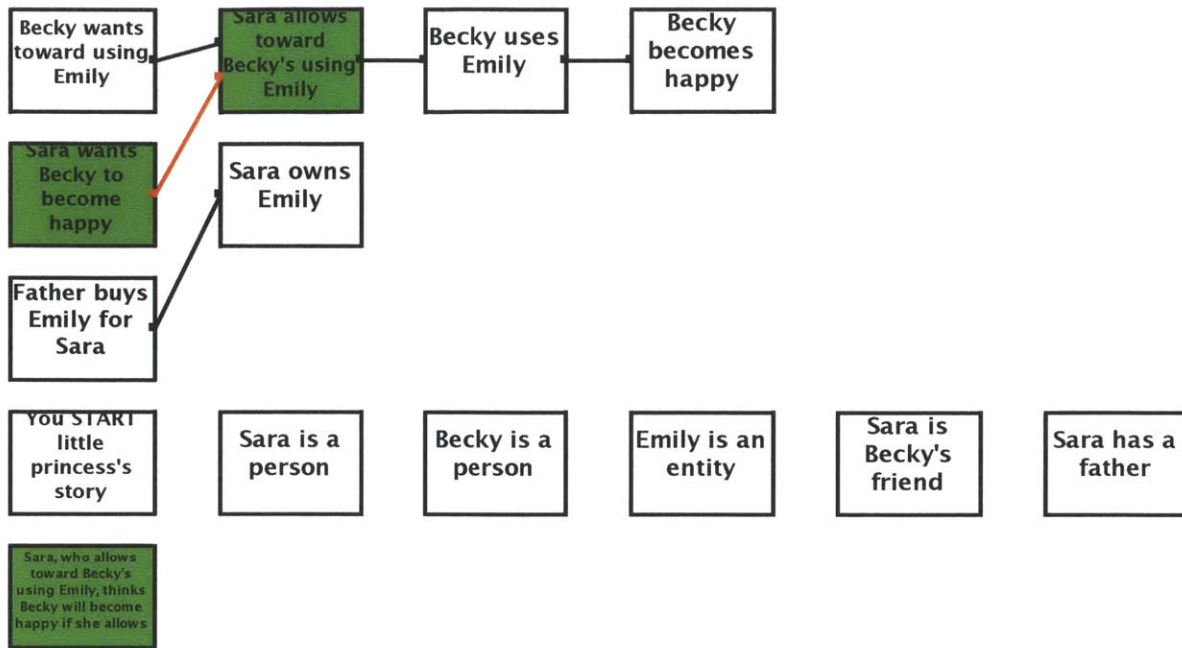


Figure 4-16: "kind" pattern found in "Little Princess"

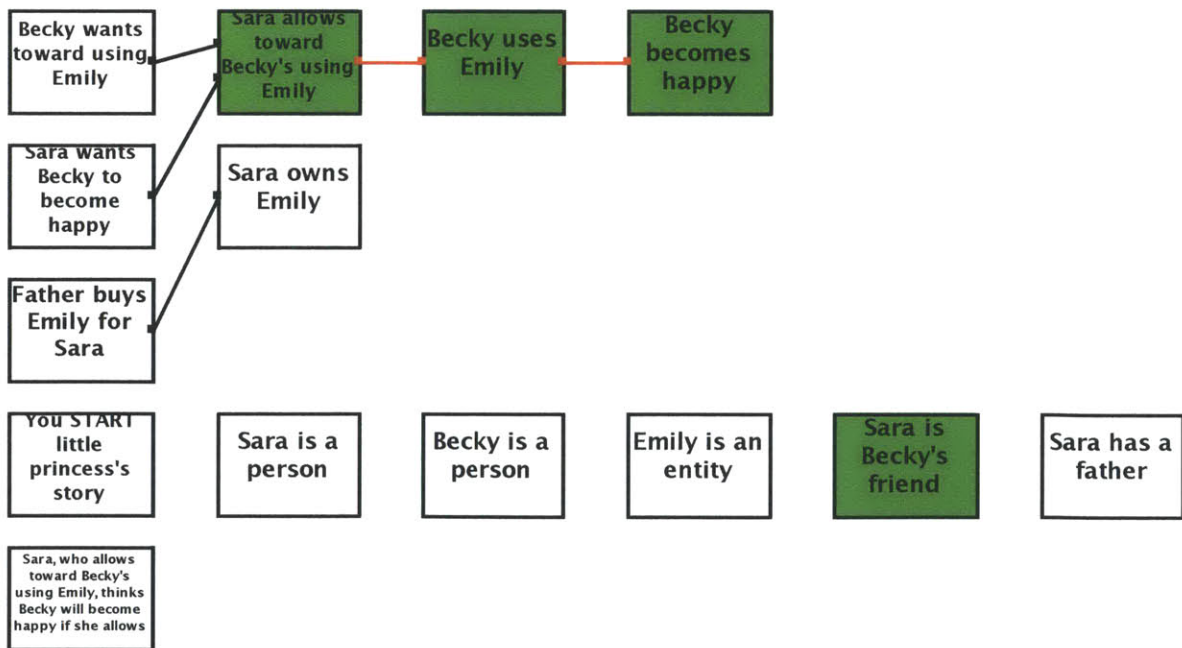


Figure 4-17: First "kind act" pattern found in "Little Princess"

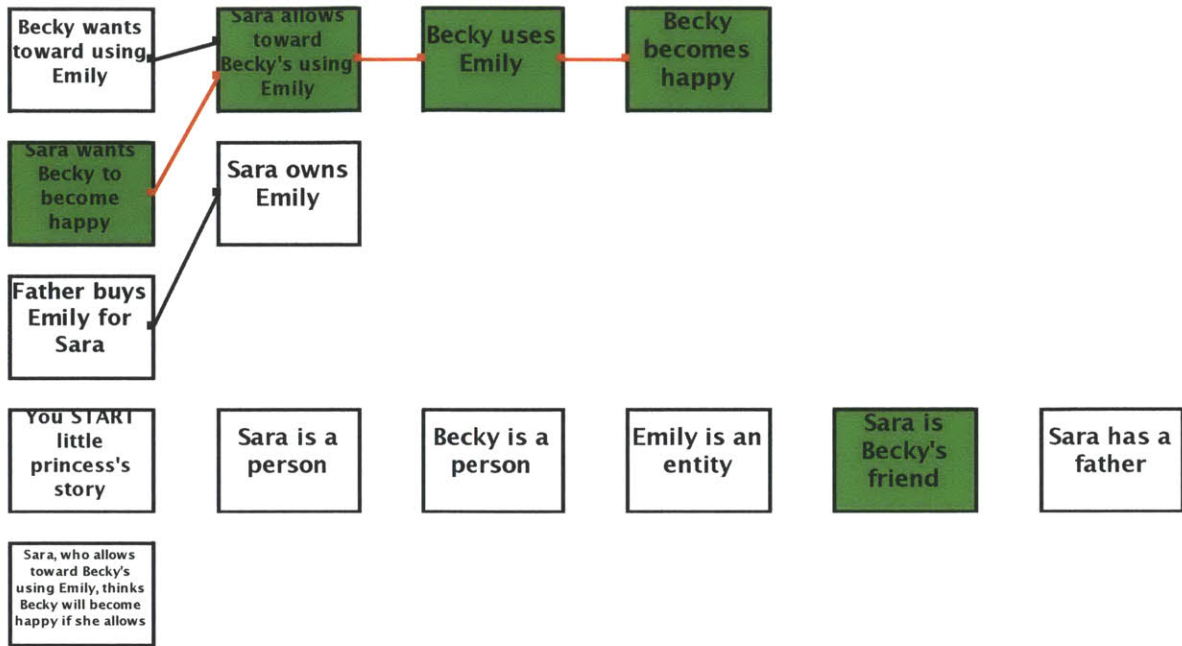


Figure 4-18: Second “kind act” pattern found in “Little Princess”

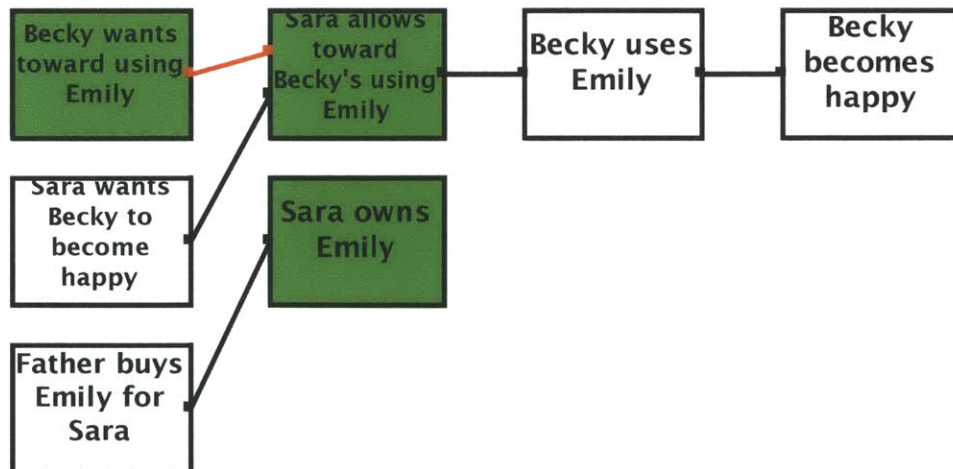


Figure 4-19: “sharing” pattern found in “Little Princess”

desire such things or to people this generous person *thinks* will want it. Furthermore, the gift-giving lacks an expectation of reciprocity from the recipient because the giver ultimately wants the recipient to become happy as opposed to feeling obliged to give something in return. In other words, an intent of causing unconditional kindness underlies both the generosity and kindness traits and explains why kindness encompasses generosity.

To capture the definition for generosity, I define its concept pattern:

Start description of ‘‘generous’’.

XX is a person.

YY is a person.

EE is an entity.

XX thinks YY will become happy if YY receives EE.

XX owns EE.

XX’s wanting YY to become happy leads to XX’s providing EE to YY.

The end.

As one can see, the pre-meditation for giving is expressed in “XX thinks YY will become happy if YY receives EE” and in “XX’s wanting YY’s becoming happy leads to XX’s providing EE to YY.” First, XX infers that the entity EE will make YY happy. Then, because XX owns EE (EE is not limited to material goods and could be extended to time, affections, or other resources) and XX wants to make YY happy, XX gives this EE entity to YY. Similar to the previously seen kindness patterns, the result of YY becoming happy is not explicit because only the intent and resulting generalized action define a trait concept pattern.

Now, to express a corresponding trait-action pattern, I define a pattern for something I call “material sacrifice.” Paralleling “physical sacrifice,” where the character gives up his physical well being for the sake of another character’s happiness, “material sacrifice” involves giving up something the character owns while gifting the other character with an entity that will make this other character happy. The exchange is expressed in “XX’s providing EE to YY leads to XX’s losing FF” below:

Start description of ‘‘material sacrifice’’.

XX is a person.

YY is a person.

EE is an entity.

FF is an entity.

AA is an action.

XX’s performing AA leads to XX’s providing EE to YY.

XX’s performing AA leads to XX’s losing FF.

XX’s providing EE to YY leads to YY’s becoming happy.

The end.

The classic example of material sacrifice is that of O. Henry’s “The Gift of the Magi,” where a young couple give up their respective treasures in order to purchase accessories for the other’s own treasure. Jim, who cherishes his heirloom watch, pawns it to purchase combs for his wife Della’s long hair while Della sells her hair to buy a gold chain for Jim’s watch. From this ironic Christmastime exchange, the couple discover that the sacrifices ultimately

matter more than the material goods:

And here I have lamely related to you the uneventful chronicle of two foolish children in a flat who most unwisely sacrificed for each other the greatest treasures of their house. But in a last word to the wise of these days let it be said that of all who give gifts these two were the wisest. Of all who give and receive gifts, such as they are wisest. Everywhere they are wisest. They are the Magi.

[4]

To capture the acts of material sacrifice described in the above passage from “The Gift of the Magi,” I sketch the following version of the short story:

Start story titled ‘‘Gift of the Magi’’.

Della is a person.

Jim is a person.

Gold Chain is an entity.

Combs is an entity.

Hair is an entity.

Watch is an entity.

Della thinks Jim will become happy if Jim receives the Gold Chain.

Della owns the Gold Chain because Della sells the Hair.

Della loses the Hair because Della sells the Hair.

Della provides the Gold Chain to Jim because Della owns the Gold Chain.

Della provides the Gold Chain to Jim because Della wants Jim to become happy.

Jim thinks Della will become happy if Della receives the Combs.

Jim owns the Combs because Jim sells the Watch.

Jim loses the Watch because Jim sells the Watch.

Jim provides the Combs to Della because Jim owns the Combs.

Jim provides the Combs to Della because Jim wants Della to become happy.

Della becomes happy because Jim provides the Combs to Della.

Jim becomes happy because Della provides the Gold Chain to Jim.

The end.

Genesis’s elaboration graph is shown in Figure 4-20. Both “generous” (Figures 4-21, 4-22) and “material sacrifice” (Figures 4-23, 4-24) concept patterns are found with two instances of each discovered due to the symmetry of the plot.

#### 4.2.6 Shy

Though the emotion of shyness has been widely studied by psychologists, resulting in a rich body of literature on why it exists to how shyness can be overcome, shyness can be generally defined as the feeling of anxiety a person feels in social situations, especially ones involving strangers. In the context of the Genesis framework, shyness provides an interesting case study of **reification of actions** and **action avoidance**. As mentioned in the introduction to this chapter, I was unable to fully implement the trait concept pattern in Genesis at

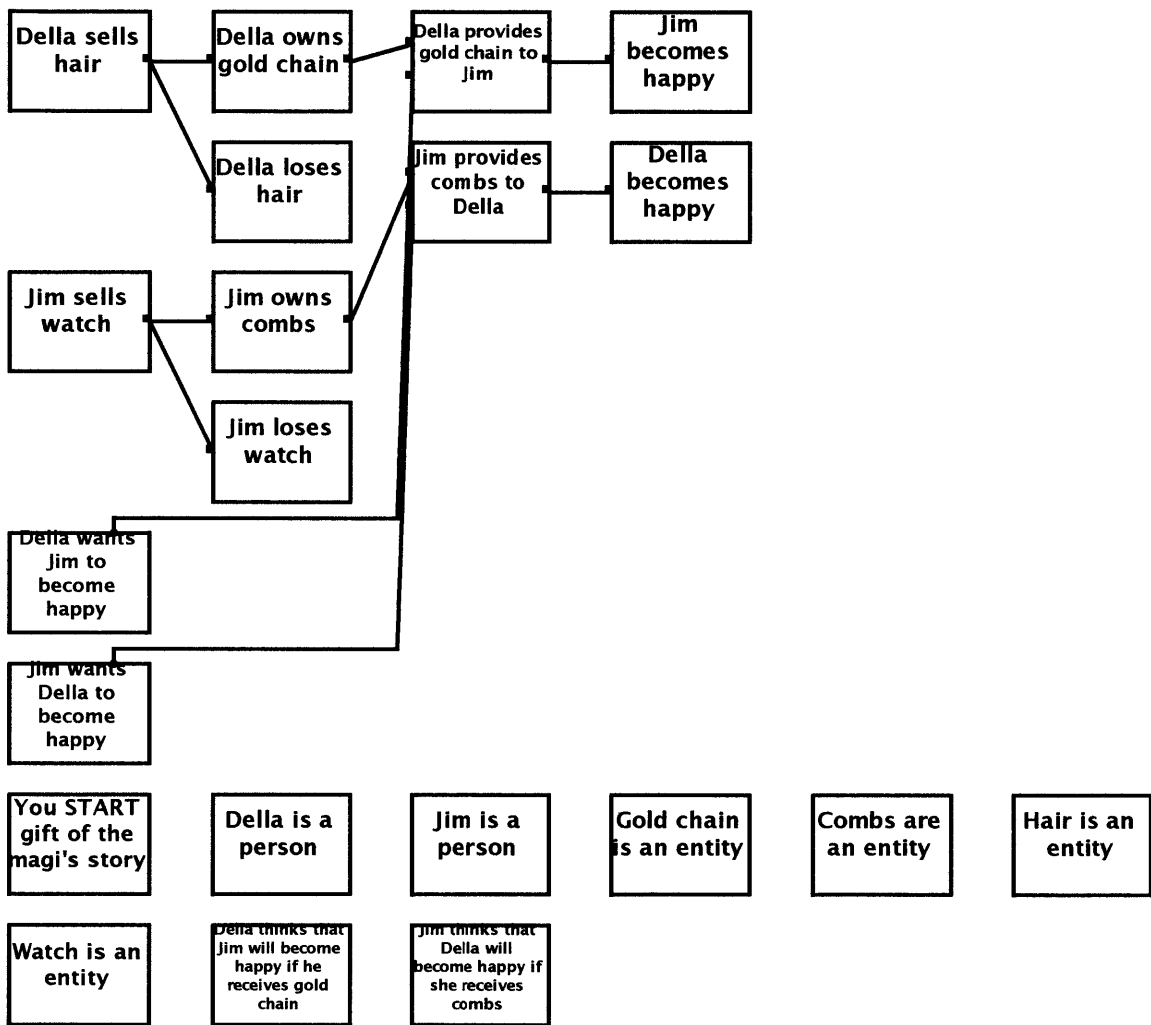


Figure 4-20: Elaboration graph for "Gift of the Magi"

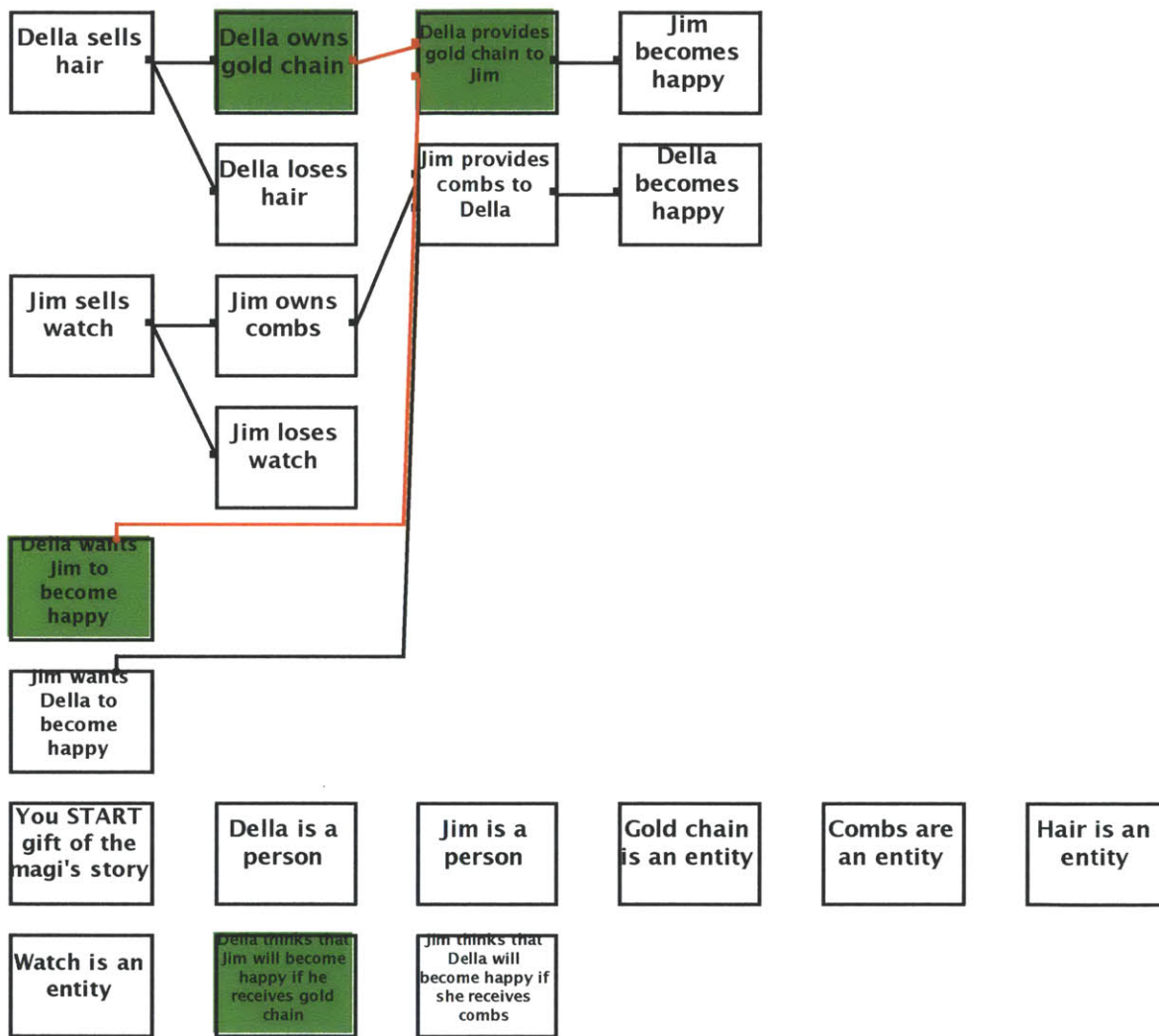


Figure 4-21: First “generous” pattern found in “Gift of the Magi”

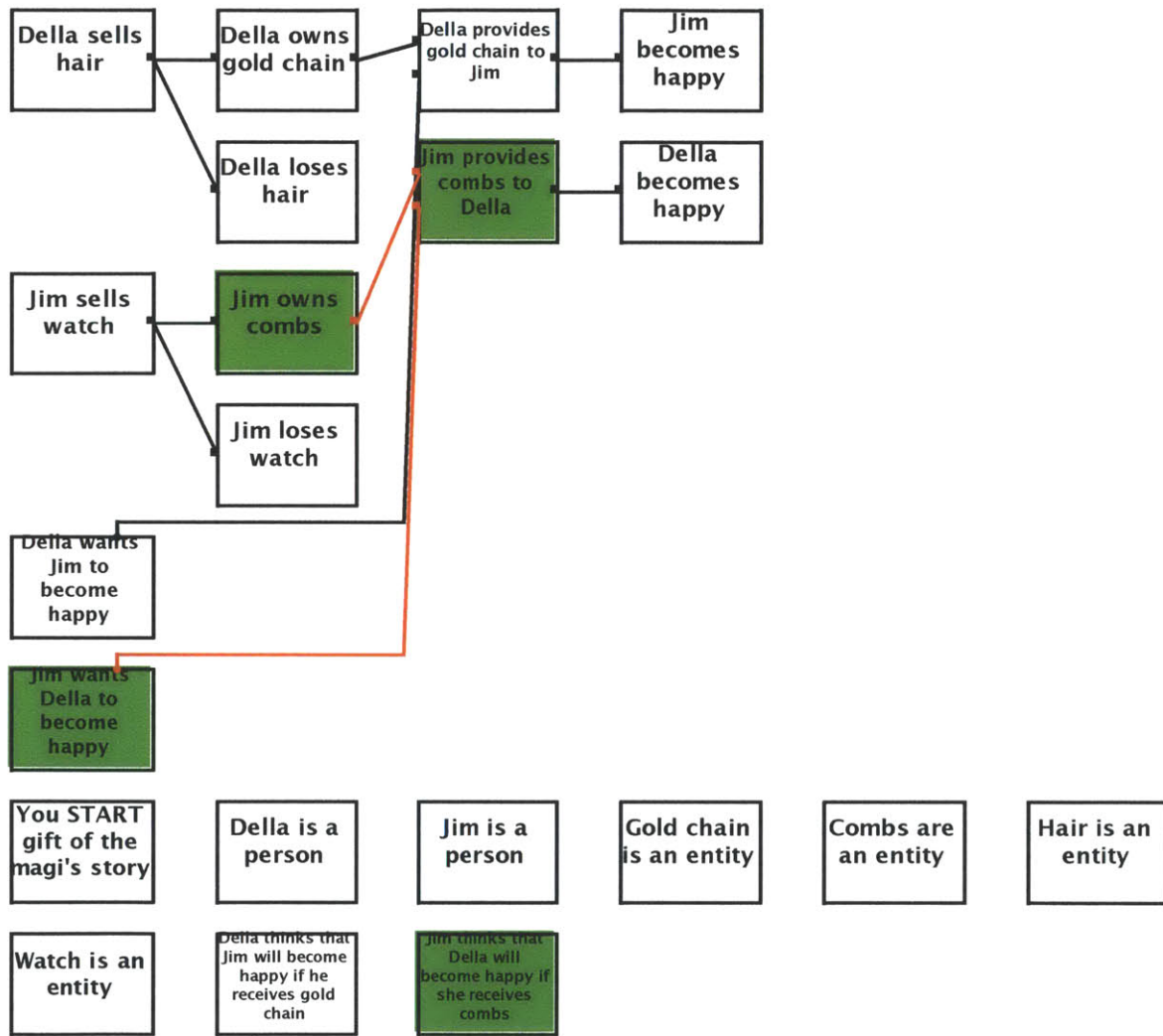


Figure 4-22: Second “generous” pattern found in “Gift of the Magi”



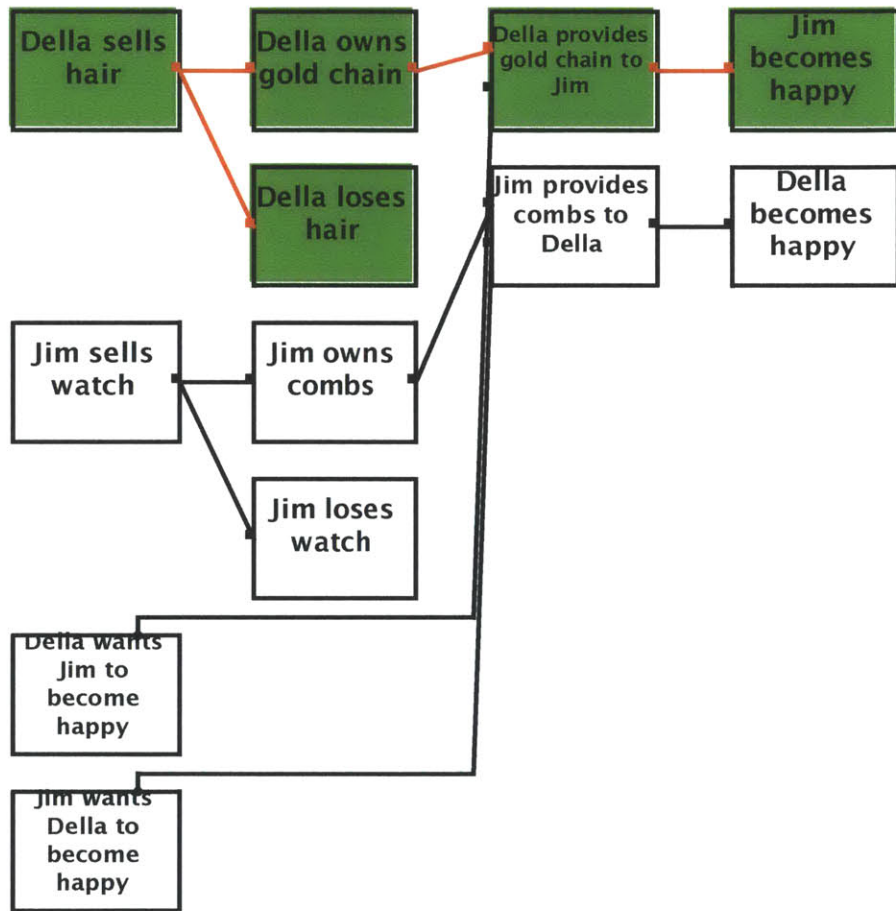


Figure 4-23: First “material sacrifice” pattern found in “Gift of the Magi”

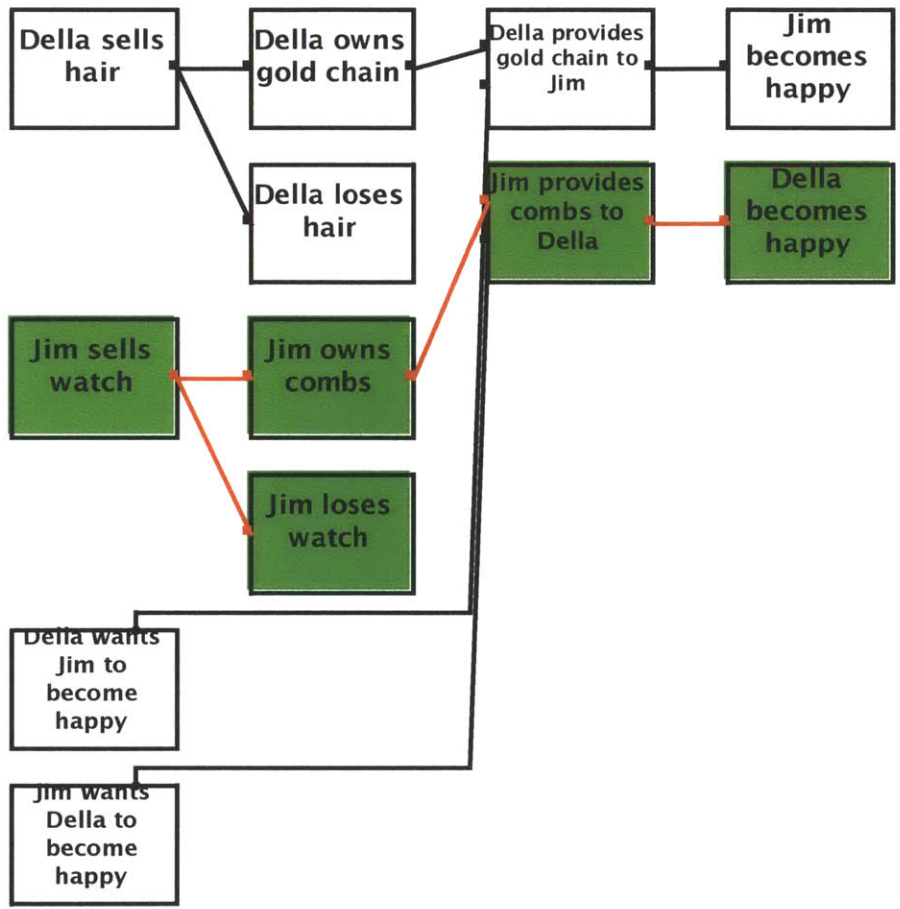


Figure 4-24: Second “material sacrifice” pattern found in “Gift of the Magi”

this time due to the lack of a mechanism for expressing the reification process, so I instead elucidate the unique features of this trait and its corresponding concept pattern, “shy - avoidance”:

Start description of ‘‘shy - avoidance’’.

AA is an action.

XX is a person.

AA is a social interaction.

XX thinks that XX will become embarrassed if XX performs AA.

XX’s not wanting to become embarrassed leads to XX’s avoiding AA.

The end.

First, shyness involves **reification**, or the concretization of an action AA in the line “AA is a social interaction.” Because shyness encompasses an entire class of actions, that of social interactions, I need to first check that AA is indeed a social interaction before defining how XX will perform (or in the case of shyness, avoid) AA. Though Genesis lacks the ability, reification of actions is an incredibly powerful device as one is then able to define intentional traits that revolve around engaging in or avoiding (to name two sample behaviors) an entire class of actions. For example, one can also define “fearful” as the avoidance of any “self-harming” actions.

**Action avoidance**, secondly, is another notable quality of the shyness pattern because avoidance can be expressed in multiple ways. One method is to describe XX’s avoidance of AA in explicit words inside a story as seen above. However, detection of action avoidance heavily depends on whether story analysis is performed in real time or performed post-reading of the entire narrative.

If story analysis is performed after the entire narrative has been read, Genesis should be able to easily infer action avoidance: the character XX performed some alternative action BB such that AA was not performed. To strengthen this “avoidance,” the story could mention that XX deliberately thought that performing BB would lead to his avoiding AA.

However, if story analysis is performed in real time, inferring action avoidance is much harder. The reader or Genesis (perhaps via explanation rules) must predict the possible paths that XX could take at that point in the story and then evaluate the taken path after the story has been read to see if the target path has been deliberately avoided.

Overall, shyness presents a fascinating case study due to its dependence on reification of an entire class of actions and avoidance of actions. Implementation of these two capabilities in Genesis will enable traits such as shyness to be defined while opening up the possibility of defining behaviors such as liking/disliking or avoidance/preference that relate to categories of actions.

#### 4.2.7 Mean

Meanness, like its antonym “kind,” is another “suitcase” personality trait that is broadly used while taking on subtle sub-definitions depending on context. Though the Oxford

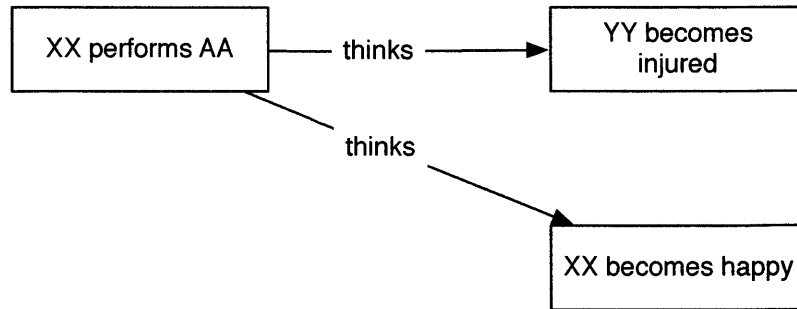


Figure 4-25: Callously mean / selfish

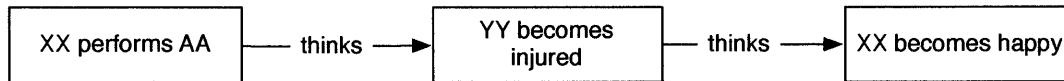


Figure 4-26: Maliciously mean

Dictionary defines it as “not generous,” “unkind,” “spiteful” and “unfair” [12], I instead focus on the general underlying meaning of “mean” as the willingness to harm someone for personal gain. Furthermore, to emphasis the intentionality of this desire to harm, a mean character needs to know ahead of time that his action will harm another character, while thinking that this action will be self-beneficial.

However, the different stages at which this mean character becomes happy highlights the subtlety of meanness. The character could think that a particular action AA will lead to his happiness but with AA causing harm to a secondary character, as seen in Figure 4-25. Such an interpretation exemplifies a “callous” or “selfish” version of meanness:

```

Start description of ‘‘callously mean / selfish’’.
XX is a person.
YY is a person.
XX thinks that XX harms YY if XX performs AA.
XX thinks that XX will become happy if XX performs AA.
XX’s performing AA leads to YY’s becoming injured.
The end.
  
```

On the other hand, the mean character could derive happiness directly from a harmful action as exemplified in Figure 4-26, i.e. a “malicious” meanness:

```

Start description of ‘‘maliciously mean’’.
XX is a person.
YY is a person.
XX thinks that XX harms YY if XX performs AA.
XX thinks that XX will become happy if XX injures YY.
XX’s performing AA leads to YY’s becoming injured.
The end.
  
```

The differences become clearer in the actualized trait-action patterns of both types of meanness, where XX goes through with the action AA and achieves his desired happiness:

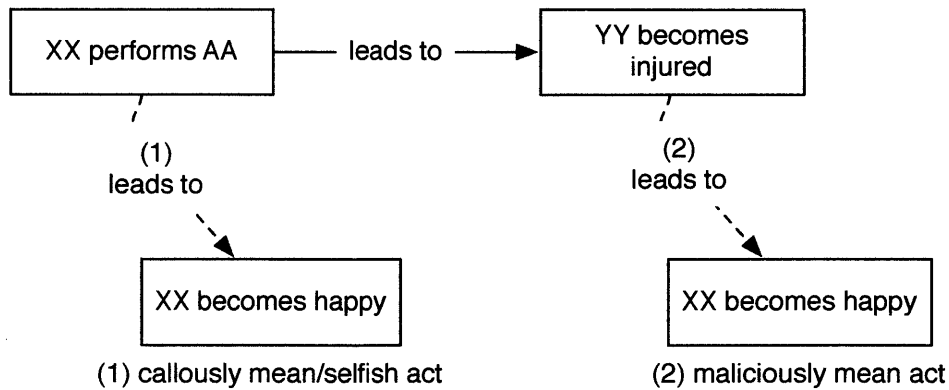


Figure 4-27: Comparison of “callously mean / selfish act” and “malicious mean act” patterns

Start description of “callously mean / selfish act”.

XX is a person.

YY is a person.

XX’s performing AA leads to YY’s becoming injured.

XX’s performing AA leads to XX’s becoming happy.

The end.

Start description of “maliciously mean act”.

XX is a person.

YY is a person.

XX’s performing AA leads to YY’s becoming injured.

XX’s harming YY leads to XX’s becoming happy.

The end.

The event traces of these two trait-action patterns are contrasted in Figure 4-27.

To test these trait and trait-action patterns, I adapted Charles Dickens’s *Great Expectations* for Genesis. *Great Expectations* tells the tale of the orphan Pip who comes into great wealth thanks to a mysterious benefactor. Along the way, he meets the spinster Miss Havisham, who as a result of being left at the altar by her fiancé many years ago has been living in her decaying mansion wearing her old wedding gown ever since. Wanting to inflict the same pain she felt from being deserted, Miss Havisham adopts a girl named Estella and raises her to share a distrust of men. As part of this vicarious revenge scheme, Miss Havisham introduces Pip to Estella and encourages Pip to fall in love with the girl, knowing that Estella will eventually break Pip’s heart. [1]

Miss Havisham’s manipulation of Pip and Estella demonstrates the two types of meanness that I have described. Her raising of Estella to be an unfeeling woman exemplifies callous meanness and selfishness because Miss Havisham cares little for how her own desire for revenge, vicariously enabled through her manipulation of Estella, ultimately harms her adopted daughter. Meanwhile, Miss Havisham intentionally seeks to harm Pip through her manipulation of him because she thinks this will lead to her becoming happy.

I express the plot of *Great Expectations* in the following story. The name “Miss Hav-

isham” has been substituted with “Eliza” due to Genesis parsing troubles with unusual names; “Eliza” is the first name of the real-life spinster upon whom Dickens modeled Miss Havisham.

Start story titled “Great Expectations”.

Estella is a person.

Eliza is a person.

Pip is a person.

Eliza thinks that Eliza will injure Pip if Eliza manipulates Pip.

Eliza thinks that Eliza will become happy if Eliza injures Pip.

Eliza manipulates Pip.

Pip falls in love with Estella because Eliza manipulates Pip.

Pip becomes heart-broken because Pip falls in love with Estella and because Estella does not marry Pip.

Estella does not marry Pip because Estella becomes cold.

Pip becomes injured because Pip becomes heart-broken.

Eliza becomes happy because Pip becomes injured.

Eliza thinks that Eliza will injure Estella if Eliza manipulates Estella

Eliza thinks that Eliza will become happy if Eliza manipulates Estella.

Eliza manipulates Estella.

Estella becomes cold because Eliza manipulates Estella.

Estella becomes injured because Estella becomes cold.

Eliza becomes happy because Eliza manipulates Estella.

The end.

The elaboration graph for “Great Expectations” is shown in Figure 4-28, while the trait concept patterns for “callously mean / selfish” (Figure 4-29) and “maliciously mean” (Figure 4-30) are found among their respective plot elements as expected.

As for the discovery of the trait-action patterns, the results from the elaboration graph yield interesting discoveries. First, Genesis discovers “callously mean / selfish act” within Miss Havisham’s manipulation of Estella because Estella ultimately becomes injured from Miss Havisham’s selfish machinations (Figure 4-31).

Secondly, Genesis uncovers two “maliciously mean act” patterns from Miss Havisham’s injuring of Pip. For one, she maliciously harms Pip through her direct manipulation (Figure 4-32). At the same time, Miss Havisham also maliciously harms Pip through her manipulation of Estella, because Estella’s resulting cold-hearted nature ultimately breaks Pip’s own heart, thus hurting him (Figure 4-33). Therefore, Miss Havisham maliciously harms Pip in two directions: directly manipulating him and indirectly by manipulating Estella.

However, Genesis discovers two additional instances of “callously mean / selfish act” as displayed in Figures 4-34 and 4-35. Note that the plot elements highlighted are the same as that of the two “maliciously mean act” patterns that have already been shown. Genesis believes “maliciously mean act” to be a sub-pattern of “callously mean / selfish act” because

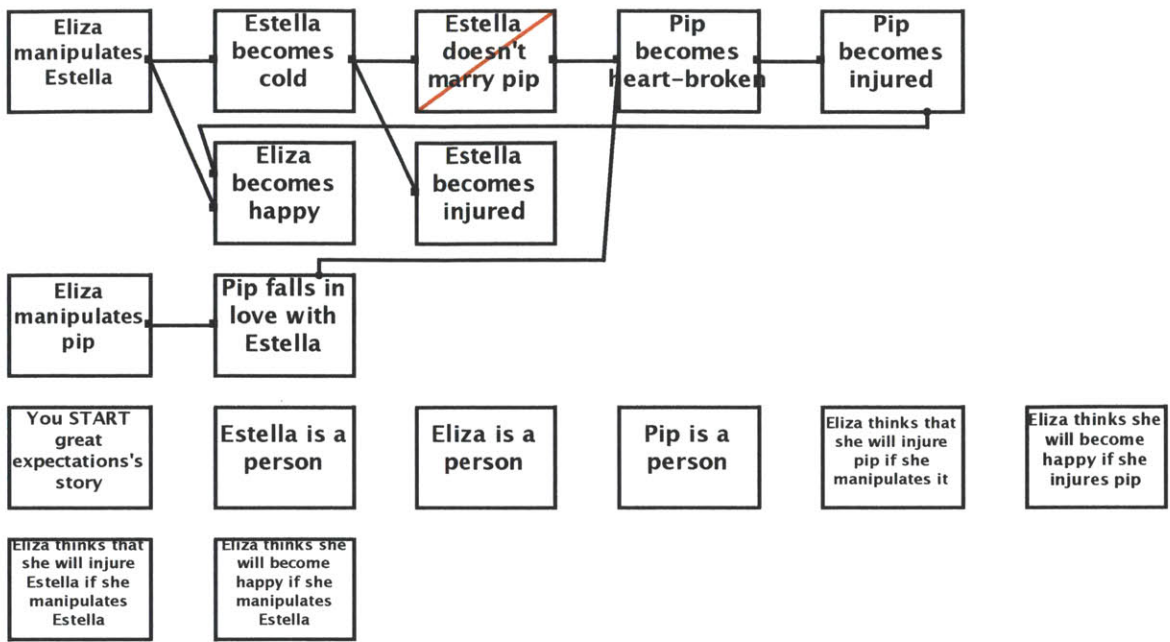


Figure 4-28: Elaboration graph for “Great Expectations”

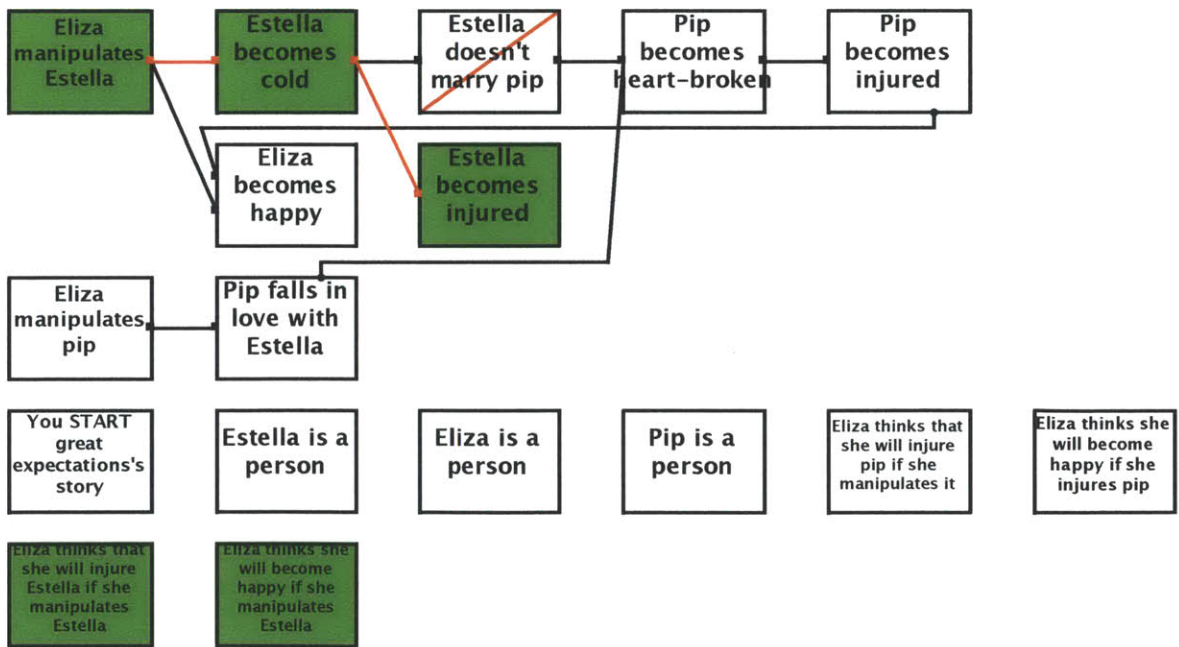


Figure 4-29: “callously mean / selfish” pattern found in “Great Expectations”

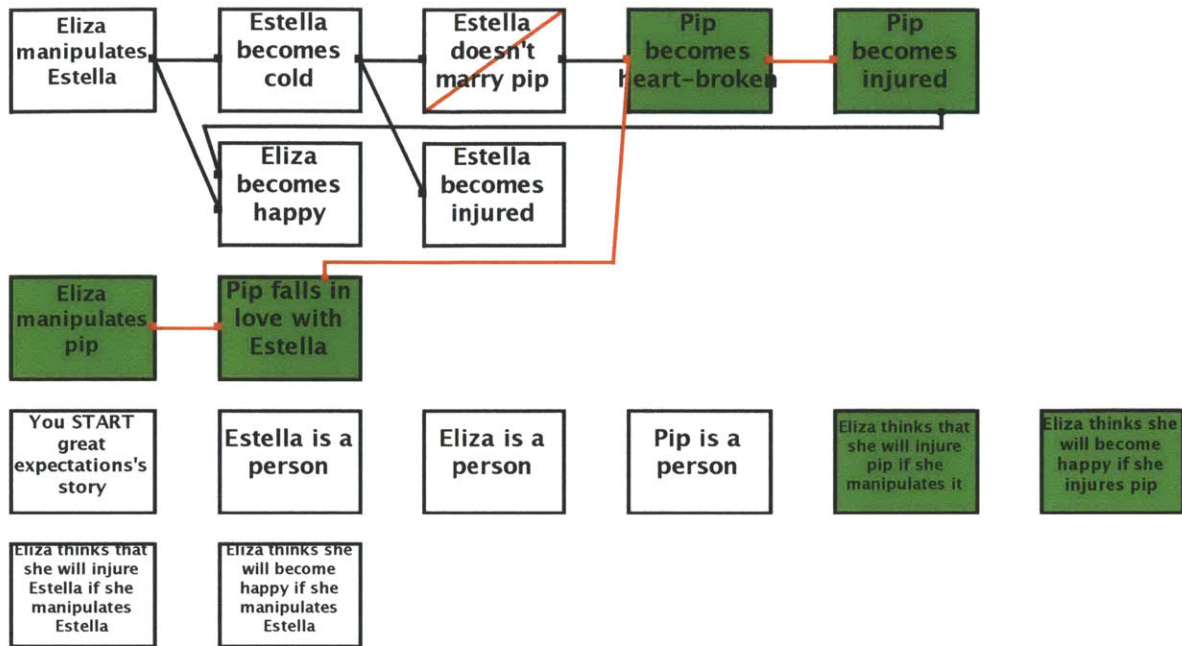


Figure 4-30: “maliciously mean” pattern found in “Great Expectations”

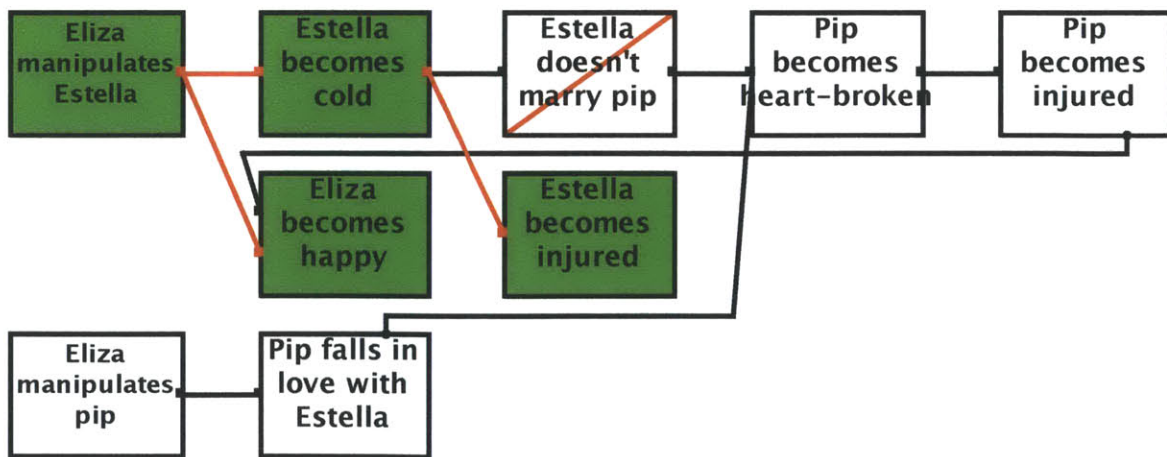


Figure 4-31: “callously mean / selfish act” pattern found in “Great Expectations” for Miss Havisham and Estella



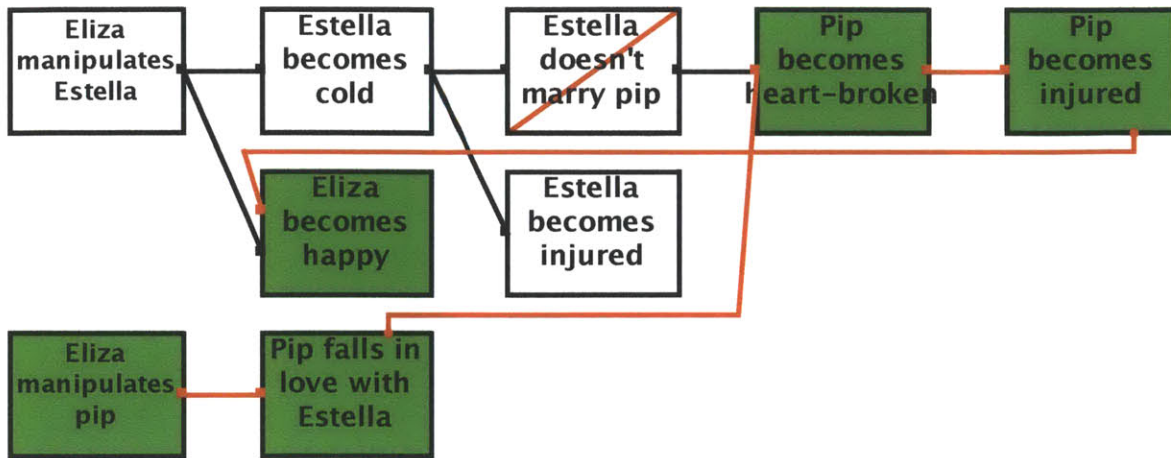


Figure 4-32: First “maliciously mean act” pattern found in “Great Expectations” for Miss Havisham and Pip

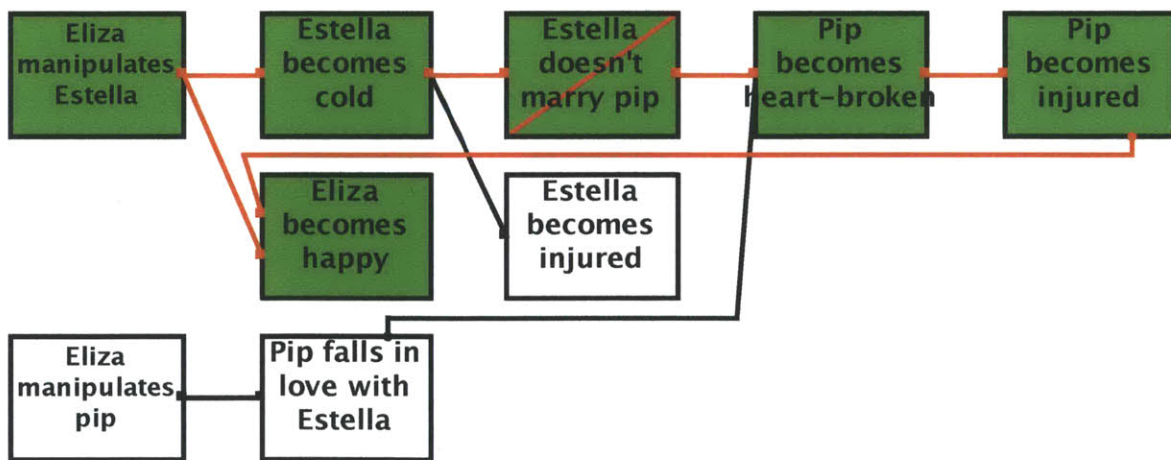


Figure 4-33: Second “maliciously mean act” pattern found in “Great Expectations” for Miss Havisham and Pip

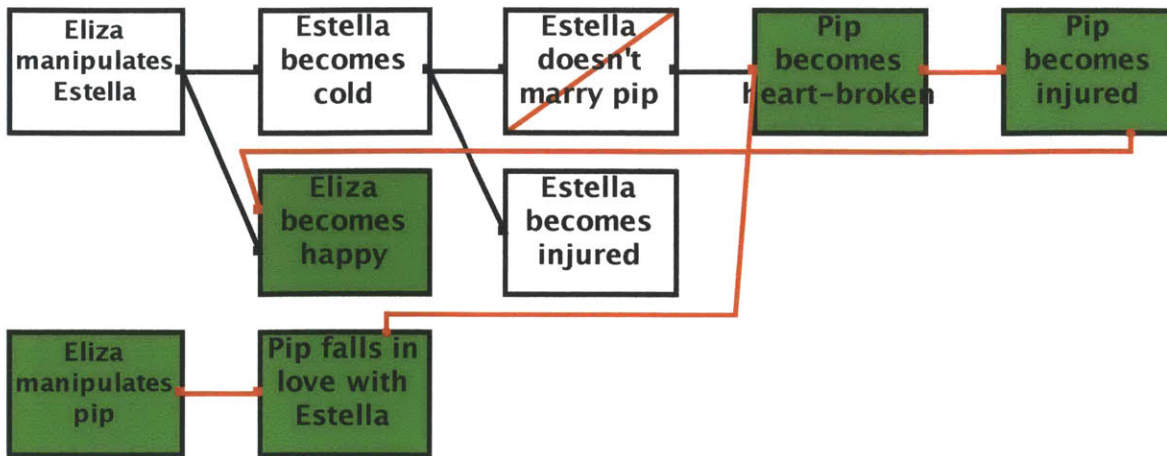


Figure 4-34: First “callously mean / selfish act” pattern found in “Great Expectations” for Miss Havisham and Pip

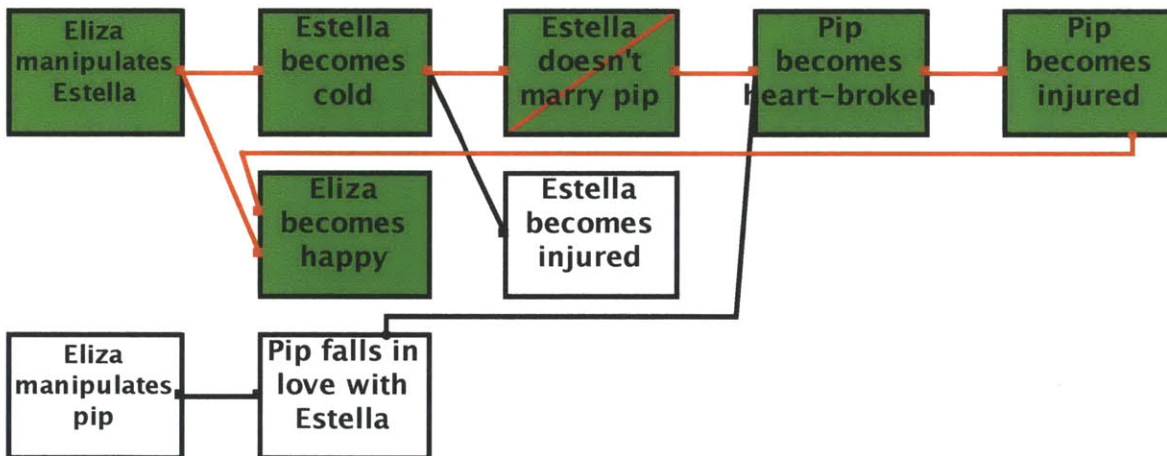


Figure 4-35: Second “callously mean / selfish act” pattern found in “Great Expectations” for Miss Havisham and Pip

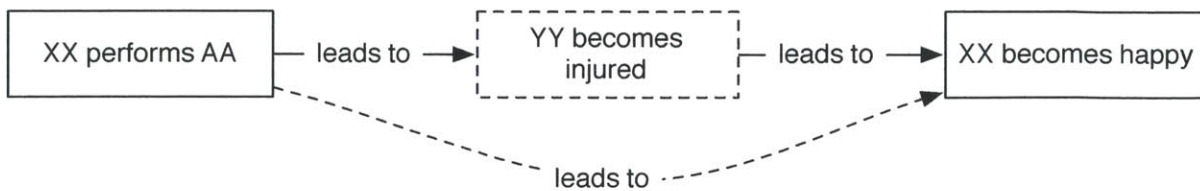


Figure 4-36: “callously mean / selfish act” pattern encompasses “maliciously mean act” pattern

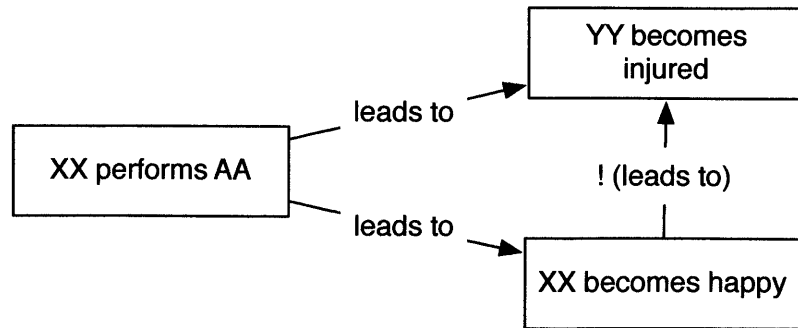


Figure 4-37: Modified “callously mean / selfish act” pattern to dissociate it from “maliciously mean act” pattern

of the way it handles the leads-to relationship. Returning to the comparison of these two trait-action patterns from Figure 4-27, the “leads to” for “XX becoming happy” from the “callous mean / selfish act” encompasses the second “leads to” for “maliciously mean act.” In other words, any plot element path that begins with “XX performs AA” and ends with “XX becomes happy” such that “XX performs AA” also leads to “YY becomes injured” is considered to be callously mean/selfish (see Figure 4-36). If Genesis finds “YY becomes injured” in the path of “XX performs AA” leading to “XX becomes happy,” then Genesis also discovers “maliciously mean act.”

If one desires Genesis to discover the two trait-actions separately, one will then need to specify that the path from “XX performs AA” to “XX becomes happy” *cannot* include “YY becomes injured,” as seen in Figure 4-37. Alas, this specific feature of “does not lead to” has yet to be implemented in Genesis, so “callously mean / selfish act” will have to include “maliciously mean act” for now.



## Chapter 5

# Trait Learning

### 5.1 Overview

The sample traits described in Section 4 exemplify *user-generated* concept patterns. On the other hand, for Genesis to learn a new personality trait, it will have to discover and generate a new concept pattern from the raw story text. Given how I have chosen to represent personality traits, learning new traits can nevertheless be considered a special case of concept pattern learning.

In my discussion of trait learning in this section, and also of trait application in Section 6 and trait inference in Section 7, I will be showing partial implementations or hand simulations of the examples. The aim of the discussion will instead be to describe potential frameworks for performing these three activities.

### 5.2 Previous Work

Previous work arising from Genesis group members, including that by Patrick Winston, Mark Finlayson, Matthew Fay, and Caryn Krakauer, have all focused on the problem of learning new concepts in general (not restricted to Genesis' concept pattern framework). I describe some of their work and how they may be applied to learning personality traits in the following subsections.

#### 5.2.1 Structural Alignment

In his thesis “Enabling Imagination Through Story Alignment,” Matthew Fay designs and implements a structural alignment engine that, among other functionalities, compares two stories by aligning their plot units (plot units being synonymous with plot elements) [3]. For instance, given the stories of Disney’s “The Lion King” and Shakespeare’s *Macbeth*, Fay’s algorithm produces the following plot element-level alignment in Figure 5-1<sup>1</sup>.

---

<sup>1</sup>Figure reproduced with permission from [3]

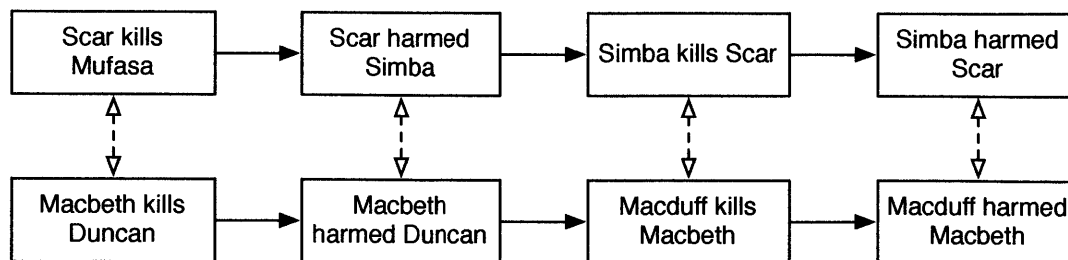


Figure 5-1: Alignment of story elements

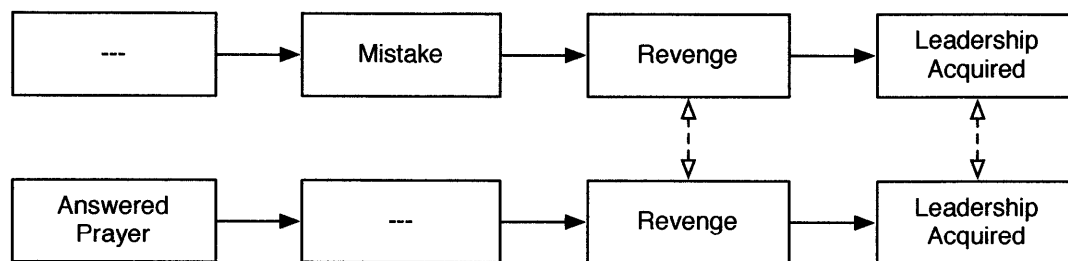


Figure 5-2: Alignment of concept patterns

More importantly, the engine can perform an alignment at the concept pattern level. Given two stories, the first containing patterns for “Mistake,” “Revenge,” and “Acquired Leadership,” and the second containing “Answered Prayer,” “Revenge,” and “Acquired Leadership,” the alignment process produces results shown in Figure 5-2<sup>2</sup>. Note that plot elements and patterns that are not matched between the two stories are ignored as seen by the empty boxes.

Fay’s work enables a side-by-side comparison of plot and concept pattern structure between two stories such that the alignment is independent of the characters themselves. In other words, while the flow of who-did-what matters in terms of parallelizing two story structures, the actual character names can be ignored for the output. Story alignment consequently becomes a valuable tool for learning the patterns for personality traits. Although the characters, names, and intermediary actions might differ across stories, character intention and outward action/result, the two factors defining any particular intentional personality trait will remain the same and thus can be extracted from these stories through story alignment.

### 5.2.2 Similarity Module

Caryn Krakauer then extends Fay’s work in her own thesis, “Story Retrieval and Comparison Using Concept Patterns” by enabling the discovery of concept patterns that are not explicitly defined by the user [7]. First, given the elaboration graph of a story (of size  $n$

<sup>2</sup>Figure reproduced with permission from [3]

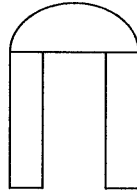


Figure 5-3: Arch

in elements) and a user input in the form of size of concept pattern, her “Concept Pattern Discovery Algorithm” extracts all the possibly concept patterns of that size, e.g.  $m$ , in time  $O(n^m)$ , resulting in  $O(m^n)$  patterns. Of course, for larger stories, the number of possible patterns of size  $m$  quickly blows up, so a filtering step follows: once the concept patterns have been extract from a collection of stories, only the patterns that appear across more than one story are filtered out.

Then, using Fay’s structural alignment work, Krakauer compares patterns generated across stories by finding and then comparing their alignment structures. This method of comparison is vital for distinguishing between patterns that share events but not structure. For instance, Krakauer gives the example of “Mary hits Sally leads to Sally hits Mary leads to Mary yells at Sally” as being fundamentally dissimilar to “Mary hits Sally and Mary yells at Sally leads to Sally hits Mary.”

Krakauer also performs word-based generalizations to extract higher-level concept pattern. For instance, the concept of “revenge” is characterized by the pattern, “A harming B leads to B harming A,” but “harming” includes an assortment of actions ranging from “kick” to “kill.” Using the semantic hierarchy of WordNet, Krakauer calculates a “thread distance” to determine if two concept patterns express the same idea despite using different vocabulary.

Krakauer’s approach to extracting and comparing concept patterns, especially when used in conjunction with Fay’s structural alignment engine, enables the possibility of learning trait concept patterns from a collection of stories.

### 5.2.3 Near-Miss Learning

One final tool relevant to trait concept learning is that of Patrick Winston’s near-miss learning. Winston’s near-miss learning articulates a solution to the problem of concept learning through the presentation of negative examples [17]. Called “near-misses,” these negative examples differ from the target concept by one or more salient differences such that these near-misses are almost but not quite the same as the target. As a result, the differences then highlight a feature that is a required component of the target concept.

For example, once the student has been taught that Figure 5-3 is an arch but that Figure 5-4 is not an arch, the student then concludes that for a pile of blocks to be an arch, the two supports cannot touch.

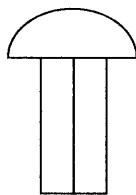


Figure 5-4: Not an arch

Near-miss learning is a powerful approach to learning by negative examples and highly applicable to trait learning: once the reader begins learning a trait, the mental representation of the concept can be refined through near-miss examples.

### 5.3 How to Learn Traits

To show how these ideas regarding concept pattern learning can be applied to trait learning, I illustrate by hand the learning process for the “vindictive” trait concept pattern. I previously defined the pattern for “vindictive” as the following:

```
Start description of ‘‘vindictive’’.
XX is a person.
YY is a person.
YY’s harming XX leads to XX’s wanting to harm YY.
The end.
```

In English, vindictiveness is the *desire* to harm someone after that person has harmed oneself. The outward action for vindictiveness is revenge:

```
Start description of ‘‘revenge’’.
XX is an entity.
YY is a entity.
XX’s harming YY leads to YY’s harming XX.
The end.
```

To begin the learning process, I (as the teacher) present the following story adopted from *The Count of Monte Cristo* regarding vindictiveness [2]:

```
Start story titled ‘‘Vindictive - Ferdinand harms Edmond’’.
Edmond is a person.
Ferdinand is a person.
Ferdinand harms Edmond.
Edmond wants to harm Ferdinand because Ferdinand harmed Edmond.
The end.
```

Genesis produces the elaboration graph seen in Figure 5-5.

At this point, Genesis should have an internal model for this trait with the names generalized to XX and YY in Figure 5-6:

I then present a second story “Vindictive - Danglars harms Edmond”,



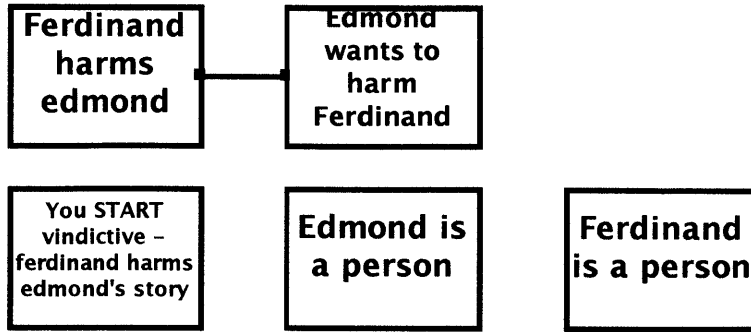


Figure 5-5: Elaboration graph for “Vindictive - Ferdinand harms Edmond”

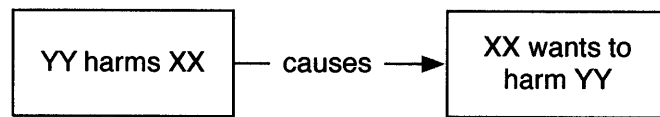


Figure 5-6: Learning “vindictive”: direct causation

Start story titled “Vindictive - Danglars harms Edmond”.

Edmond is a person.

Danglars is a person.

Danglars betrays Edmond.

Danglars harms Edmond because Danglars betrays Edmond.

Edmond becomes unhappy because Danglars harmed Edmond.

Edmond wants to harm Danglars because Edmond becomes unhappy.

The end.

Genesis then displays an expanded elaboration graph as seen in Figure 5-7.

However, if Genesis structurally aligns the two stories, it arrives at the comparison in Figure 5-8 such that Danglars’s and Ferdinand’s harming of Edmond along with Edmond’s subsequent desires to harm his respective injurers line up despite the intermediary plot elements. The “because” relation that is learned from the first story generalizes to a “leads

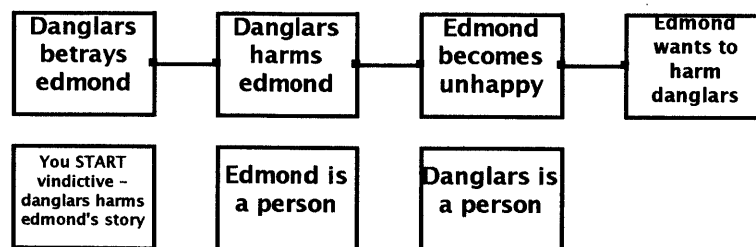


Figure 5-7: Elaboration graph for “Vindictive - Danglars harms Edmond”

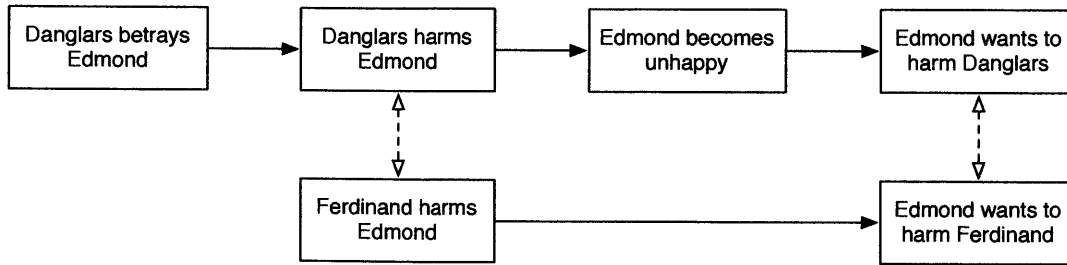


Figure 5-8: Structural alignment of “Vindictive - Ferdinand harms Edmond” and “Vindictive - Danglars harms Edmond”



Figure 5-9: Learning “vindictive”: leads-to relationship

to” relationship in Figure 5-9.

Though not necessary in this particular example, I may want to emphasize the leads-to relationship via a negative example of “not-vindictive” by giving a story where this relationship is not present.

Start story titled “Not vindictive - Albert harms Edmond”.  
 Edmond is a person.  
 Albert a person.  
 Albert harms Edmond.  
 Edmond wants to harm Albert.  
 The end.

The elaboration graph for “Edmond un-vindictively harms Albert” is produced in Figure 5-10 with the abstract model being the same.

In contrasting the model of not-vindictive against the one generated from the positive vindictive stories, Genesis should realize that the absence of the “leads to” link in the near-miss example illustrates a salient difference between the two. Therefore Genesis should identify the link between “YY harms XX” and “XX wanting to harm YY” as a requirement for the vindictive concept pattern.

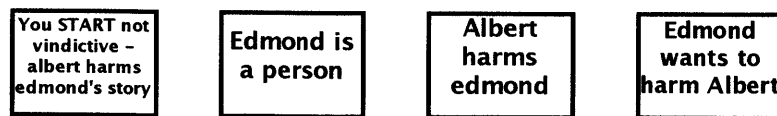


Figure 5-10: Elaboration graph for “Edmond un-vindictively harms Albert”

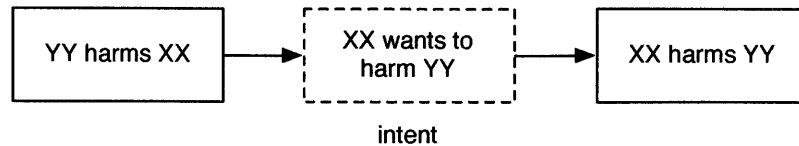


Figure 5-11: Inferring the underlying intent from the “revenge” pattern

Nevertheless, what if the storyteller only narrates the outward actions of the trait without privileging the reader with the character’s underlying intention? In other words, how can one infer the underlying intention driving the expressed action? For instance, the author could directly describe a character XX’s engaging in revenge without expressing the underlying vindictiveness. Genesis then needs to figure out that XX’s harming YY is a deliberate act that follows from YY’s harming XX, and that XX *wanted* to harm YY (Figure 5-11) from the revenge pattern “YY’s harming XX leads to XX’s harming YY.” Granted, the leads-to relationship implicitly implies that XX’s harming YY follows from YY’s harming XX; thus the retaliatory harming was no mere random act of violence. However, the intention can be strengthened through additional elements from the story.

First, assuming that the Genesis is not privy to the character’s complete mental intentions, if the story nevertheless gives evidence of premeditation, e.g. if a character plans to kill someone after that person has harmed him by preparing weapons or luring the victim to the desired location, then Genesis should *infer* that the character intended the kill. To make the parallel to the real-life legal system clearer, accidental manslaughter should not be considered evidence of a “vindictive” personality whereas murder should be, even if the same provocation and harm-leads-to-harm structure occur in both scenarios. Secondly, the post-action reaction matters too. If Genesis discovers that the character becomes happy or profits in some way after taking part in a “revenge” pattern, more than likely this character intended for the revenge to happen and acted deliberately to set it up.

To generalize these ideas, auxiliary physical and mental evidence can provide additional insight into the character’s internal thoughts. Such interpolated explanations for the character’s behaviors should then be inserted into the overall plot so that later on the knowledge about intention can be used to infer the trait of the character (see “Trait Inference,” Section 7). More importantly, when Genesis makes comparisons between multiple stories, Genesis should then discover similar intentionality elements and include them as part of the trait concept pattern.



## Chapter 6

# Trait Application

### 6.1 Overview

Authors may not always define their characters with specific traits early on in a story and instead let the reader discover the traits using the narrative itself. However, when the author does **apply** a trait to a character, this application then enables the reader to **predict** and **explain** the character's behavior:

- **Prediction:** what is he likely to do?
- **Explanation:** does his behavior make sense or does it violate expectations?

Note the temporal asymmetry between the two types of analyses: prediction happens during the reading process while explanation happens post-reading.

Trait application obviously assumes that the reader (i.e. Genesis) has already learned the meaning of the trait, either in trait concept pattern form or in both trait and trait-action concept pattern forms. These patterns should have also been named because Genesis needs to reference the corresponding concept pattern once the storyteller directly identifies the character with a particular trait. For instance, once Edmond Dantès has been identified as “vindictive” inside *The Count of Monte Cristo*, Genesis should be retrieving and referencing the vindictive concept pattern. One also assumes that the reader's or Genesis's representation for a trait aligns with that of the storyteller, though I will briefly discuss later in this chapter how the trait concept patterns may be modified while reading a narrative.

Similar to Section 5 on trait learning, the examples shown here have been partially implemented or hand simulated.

### 6.2 How to Predict Behaviors

If Frances Hodgson Burnett begins *A Little Princess* with the sentence, “Sara Crewe is kind,” what can the reader expect from this Sara character? Will she want to perform actions that make people happy if they are unhappy? Will she engage in acts of kindness

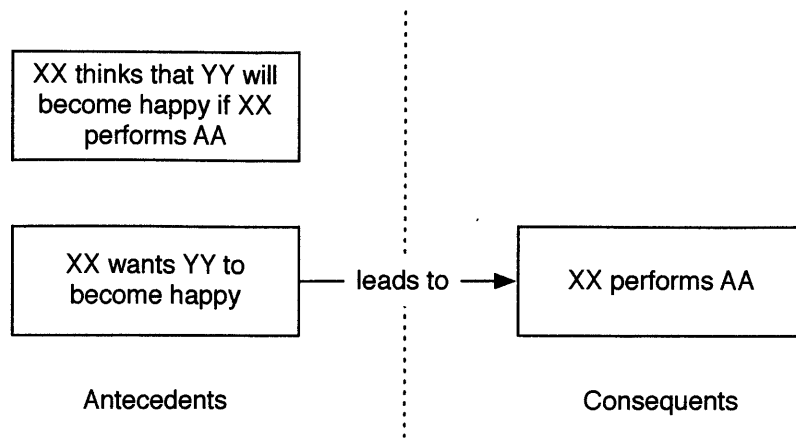


Figure 6-1: Partitioning of “kindness” trait concept pattern into antecedents and consequents

such as sharing and comforting? Expectations differ depending on prior reader knowledge of this “kindness” trait, and this knowledge can be summarized into two cases: trait concept patterns, and both trait and trait-action concept patterns.

In the first case of trait concept pattern, the reader only knows that “kindness” is represented by the following pattern:

```

XX is a person.
YY is a person.
AA is an action.
XX thinks that YY will become happy if XX performs AA.
XX’s wanting YY to become happy leads to XX’s performing AA.
  
```

In other words, if YY is unhappy and XX wants YY to become happy, then XX will perform some action AA that XX thinks will make YY happy. The reader then has to use this concept pattern to make predictions of Sara’s behavior. One possibility is to partition the trait concept pattern into antecedents and consequents by modeling the flow of the various statements and leads-to relationship as seen in Figure 6-1 where the antecedents are the boxes on the left and the consequents are all that follows to the right. Making a prediction then becomes the act of checking to see if the antecedents are ever fulfilled. If they are, then the reader may reasonably expect the consequents to follow.

### 6.2.1 Predicting From Trait Patterns

In the case of *A Little Princess*, once the reader knows that Sara is kind, the reader then could expect that when Sara thinks an action will make someone (e.g. Sara’s friend Becky) happy, and she wants that person to become happy, then Sara *may* perform the action that grants Becky happiness.

Implementation-wise in Genesis, the system will need to convert the trait concept pattern into some sort of predictive mechanism. One possibility is to convert the concept

pattern into rules, where the rule for kindness is:

```
XX is a person.  
YY is a person.  
AA is an action.  
If XX thinks that YY will become happy if XX performs AA and XX wants  
YY to become happy, then XX may perform AA.
```

The concept pattern has been framed as an explanation rule because we want to emphasize the predictive nature of knowing a character's trait.

However, even in the relatively simple case of "kindness," converting from the concept pattern results in a complicated expression because the antecedents must all be enumerated in the rule statement.

Additional knowledge about actions linked to a particular trait is crucial; knowing just the intention is not enough to predict actions. Consider this example of vindictiveness:

```
XX is a person.  
YY is a person.  
YY's harming XX leads to XX's wanting to harm YY.
```

The antecedent is YY's harming XX. Once we know that some character, say Edmond Dantès from *The Count of Monte Cristo* is vindictive and once we encounter the antecedent for vindictiveness, we expect that Dantès will want to harm anyone who injures him. However, what sort of behavior can we expect from this "wanting to harm?" We clearly need outside knowledge pertaining to what "harming" entails. Perhaps the reader has already learned that actions such as "kill," "destroy," and "fight" constitute "harming." Thus Dantès, because he is vindictive, may participate in these actions once he has been harmed. Still, Genesis requires knowledge about outward actions corresponding to a trait to make predictions because just knowing the intention via trait concept patterns does not suffice.

### 6.2.2 Predicting From Trait and Trait-Action Patterns

Prediction clearly becomes substantially more powerful once Genesis has learned both the trait concept pattern and corresponding actions in the form of trait-action patterns. For instance, the "kind" trait could potentially be linked to not only the "kind act" pattern but also to "sharing" and "helping." Once the character has been endowed with a particular trait, the Genesis should predict behavior by answering the following questions:

1. Does the situation fulfill the trait concept pattern's antecedents?
2. If so, does the situation also fulfill the any trait-action concept pattern's antecedents?

For instance, a sample trait-action pattern for kindness is "sharing":

```
Start description of "sharing".  
XX is a person.  
YY is a person.  
ZZ is an entity.
```

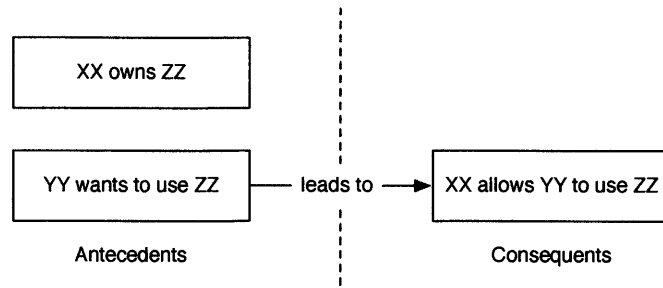


Figure 6-2: Partitioning of “sharing” trait-action concept pattern into antecedents and consequents

XX owns ZZ.  
 YY’s wanting to use ZZ leads to XX’s allowing YY to use ZZ.  
 The end.

The antecedents are “XX owns ZZ” and “YY’s wanting to use ZZ” (Figure 6-2). Combined with the two antecedents from the “kind” trait pattern, we can now make a prediction that XX may let YY use ZZ when all four of these antecedents have been fulfilled.

Therefore, to make specific predictions given knowledge of a character’s trait, Genesis should use the concept patterns both for the trait and for its corresponding trait-actions.

### 6.3 How to Explain Behaviors

The complement to predicting behavior is explaining actions after they have been performed. Knowing a character’s personality traits (as told by the author) allows the reader or Genesis to verify whether the character’s actions are justified given his predispositions or whether his actions are unexpected.

The first case of justified behavior does not provide extraordinarily insightful explanations. After all, assuming Genesis has learned both trait pattern and trait-action patterns, Genesis should then be expecting the consequents for each pattern to be carried out once the antecedents match; it should be no surprise when the consequents do appear.

On the other hand, what if the character defies expectations and performs actions contrary to these consequents? This defiance can occur at two main stages as exhibited in Figure 6-3. The first case is when the action that the character has undertaken is a negation of the existing trait-action patterns linked to the trait. An example would be if Sara Crewe, despite being described as “kind,” kicks Becky after expressing a desire of wanting Becky to become happy. Sara has now seemingly harmed Becky.

The second defiance is when the final result does not happen as predicted. Perhaps Sara’s wanting Becky to become happy leads to Sara sharing her doll with Becky, but then Becky becomes unhappy in the final consequent from the leads-to relationships. What do these two types of anomalous behavior imply about Sara?



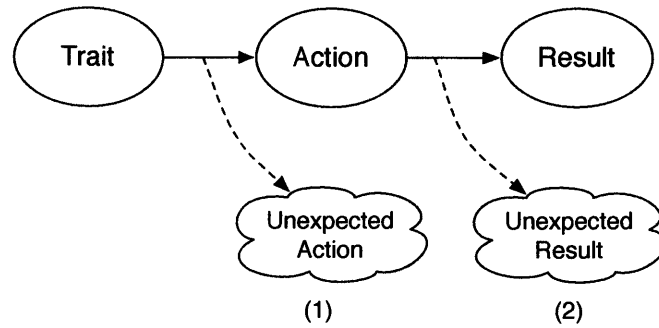


Figure 6-3: Explaining unexpected actions and unexpected results

Genesis should attempt to explain the two unexpected plot twists in a few ways. In the case of unexpected action only (i.e. Sara kicks Becky), Genesis or the reader should question Sara’s true intentions. Did she really want Becky to become happy? Or is Sara actually *unkind* given her outward actions? As a result, the reader or Genesis may **modify** their own perception of Sara despite what the author initially described. Characters change throughout a story, and perhaps Sara’s personality evolved during the plot.

Genesis can also **expand** its collection of trait-action patterns linked to the trait. Assuming that Sara is actually kind, then the kicking anomaly reinforces the idea that “wanting someone to become happy” is the crux intent of kindness; actions that follow, even if they involve harming, may be allowed under this trait. To state this another way, intentions ultimately matter more than the action itself. Perhaps Sara genuinely thought at that time that kicking Becky would result in Becky’s becoming happy.

Given that the reader and Genesis are privy to the end result of Becky’s own state through the narrative, if Becky becomes unhappy, it is more likely that Sara is unkind. On the other hand, if Becky does become happy from being harmed, then one has learned that harming could lead to happiness.

What if only the final result is unexpected? For instance, despite Sara’s best efforts in sharing the doll, Becky becomes unhappy. The reader might then conclude that Sara’s attempt at kindness was simply poorly executed or some other incident happened along the way that derailed her efforts.

Explaining negative examples using knowledge of a character’s trait is a powerful inference tool for story readers. The reader essentially sets up expectations for the characters during the course of the story. If these expectations are unfulfilled or defied, they garner additional notice from the reader who then may modify his perception of the character.



# Chapter 7

## Trait Inference

### 7.1 Overview

When storytellers avoid directly telling the reader what the characters' personalities are ahead of time, the reader must then *characterize* the actors as the plot progresses. Therefore, to infer whether a character possesses a trait without being explicitly told such by the author, Genesis needs to match learned patterns against the character's thoughts and behaviors. I describe in this section potential ways for this inference process to be performed, with implementation of the examples and suggestions to be accomplished in future work.

To begin the inference process, assume that Genesis has learned both the trait and trait-action concept patterns. For instance, it knows that the "aggressive" trait is linked to acts of intimidation:

```
Start description of 'aggressive'.
XX is a person.
YY is a person.
AA is an action.
XX thinks that YY will fear XX if XX performs AA.
XX's wanting YY to fear XX leads to XX's performing AA.
The end.
```

```
Start description of 'intimidate'.
XX is a person.
YY is a person.
AA is an action.
XX's performing AA leads to YY's fearing XX.
The end.
```

Genesis should use this learned knowledge to infer a character's personality trait in two ways: inference through trait concept pattern and inference through the trait-action pattern. Both approaches share similar steps. For instance, as Genesis progresses through a story, it might encounter plot elements that trigger the discovery of either a trait or trait-action pattern in the same manner as how Genesis discovered the traits from Section 4 on trait representation.

Nevertheless, just discovering one instance of the trait or trait-action pattern does not suffice for trait assignment because the character could have been engaging in a one-time behavior. For a character to possess a trait, Genesis must see that it is a *habitual* thought or action by definition of personality traits. Consequently a desirable approach is to collect either a group of stories sharing the same characters and/or to identify the pattern multiple times in the same story before trait assignment. Note that for Genesis to assign a trait, the assignment threshold involves a certain level of ambiguity. In reality, one human reader's tolerance for vindictiveness could merely be two revenge acts, while for another reader vindictiveness needs to be confirmed through multiple revenges.

One can determine this threshold through real-life test studies involving human readers being given a corpus of stories and then inferring traits from the characters. For instance, readers could be presented with differing versions of *The Count of Monte Cristo* and then asked, "Would you consider Edmond Dantès 'aggressive' " or "Is Monte Cristo vindictive?" Then one should set Genesis's computational thresholds based on these experimental results.

Another possibility is to allow the threshold to fluctuate depending on the reader's **mental model**, because the threshold does not need to be computationally static. After all, Eastern and Western human readers likely have differing culturally-based thresholds for inferring personality traits in stories. Genesis could mimic this by having two different thresholds, one for a "Western" reader and another for a "Eastern" reader (with the readers being different internal Genesis models), and then comparing the story analyses after both readers have processed the same corpus of stories.

Overall, once this threshold has been achieved for some character and trait, the reader should conclude that the character possesses this trait. Such knowledge should then be stored for later retrieval, especially for trait application purposes.

## 7.2 Inference From Trait Pattern and Trait-Action Pattern

The main difference between inferring from trait concept patterns and trait-action patterns arises from how much information to which the reader is privy. More specifically, the reader sometimes does not have access to direct knowledge of character intentions.

This access to knowledge heavily depends on story perspective. If a story is presented in first person for instance, then it is likely that the reader not only hears the narrator's thoughts and intentions, but also hears this narrator's speculation on other people's intentions. Note that the narrating character will likely express biases about other characters, so this bias must be removed. On the other hand, an objective third person perspective will lead to little direct information about mental states and intentions but a subjective third person perspective will indeed yield such insights.

When the mental intentions are hidden, the reader will have to either assign traits solely based on trait-action patterns or infer the underlying mental intentions. For example, if someone frequently engages in acts of intimidation, either in the same story or across

multiple stories, the reader could conclude that the linked trait, “aggressive,” is present in the character. On the other hand, the reader could use heuristics discussed in the “Trait Learning” (Section 5), to infer the underlying mental states, especially through auxiliary evidence such as premeditation that accompanies the trait-actions.



## Chapter 8

# Contributions

In the course of this thesis, I have made the following contribution towards the vision of integrating personality traits into computational story understanding:

- Defined **intentional personality traits** as composed of trait name, underlying intent, and an optional accompanying action. For instance, the trait of “vindictive” expresses an intentional desire to harm someone who has harmed the vindictive character. The vindictive character could then carry out this desire by deliberately harming the other character.
- Created a representation for intentional personality traits based on the concept pattern paradigm, with a partitioning of this representation into **trait concept patterns** and **trait-action patterns**. Thus the “vindictive” trait from above has been directly represented as the “vindictive” concept pattern and also represented through the trait-action pattern “revenge.”
  1. Created ten example traits, such as “kind and “mean,” accompanied by corresponding trait-action patterns (e.g. “sharing” and “maliciously mean act”) in Genesis using this concept pattern representation.
  2. Showed discovery of these trait and trait-action patterns within stories adapted from famous historical and fictional narratives. For example, I showed discovery of both the trait concept pattern “generous” and the trait-action concept pattern “material sacrifice” in O. Henry’s “Gift of the Magi.”
- Articulated the steps of **trait learning**, **trait application**, and **trait inference**:
  1. Trait learning: suggested how previous work in concept pattern learning can be applied towards learning trait concept patterns, along with suggesting how trait patterns, trait-action patterns, and trait names can be systematically learned. In particular, I discussed how Winston’s near-miss learning paradigm provides a framework for learning the trait concept pattern from only a select number of

stories. As an example, I showed how the reader can learn that Edmond Dantès from Alexandre Dumas's *Count of Monte Cristo*, is “vindictive” in a near-miss fashion from both positive and near-miss examples.

2. Trait application: suggested how traits can be used to explain and predict behavior within narratives. For instance, given that Sara Crew from Frances Hodgson Burnett's *A Little Princess* is “kind,” I described how Genesis can predict her future actions while explaining her past actions inside the story.
  3. Trait inference: suggested experimental approaches to discovering thresholds for inferring traits from characters, such as by asking human readers whether one would consider Edmond Dantès to be “vindictive” after reading a corpus of stories pertaining to this character.
- Showed that implementing personality traits in Genesis refines our notion of what it means to possess a trait, while shedding light on why, due to differences in opinion over the exact meaning of traits, people often debate over whether someone possesses a trait.

I hope this thesis has provided an insightful discussion of the role of personality traits in storytelling, and that the examples and steps contained within has given the reader a clear framework for pursuing additional work in integrating traits into story processing systems such as Genesis.



# Appendix A

## Concept Patterns

### A.1 Vindictive

vindictiveness reflective knowledge.txt

Start description of ‘‘revenge’’.  
XX is a person.  
YY is a person.  
YY’s harming XX leads to XX’s harming YY.  
The end.

Start description of ‘‘revenge - attempt’’.  
XX is a person.  
YY is a person.  
YY’s harming XX leads to XX’s trying to harm YY.  
The end.

Start description of ‘‘vindictive’’.  
XX is a person.  
YY is a person.  
YY’s harming XX leads to XX’s wanting to harm YY.  
The end.

## A.2 Aggressive

aggressiveness reflective knowledge.txt

Start description of ‘‘intimidation’’.

XX is a person.

YY is a person.

AA is an action.

XX’s performing AA leads to YY’s fearing XX.

The end.

Start description of ‘‘aggressive’’.

XX is a person.

YY is a person.

AA is an action.

XX thinks that YY will fear XX if XX performs AA.

XX’s wanting YY to fear XX leads to XX’s performing AA.

The end.

### A.3 Brave

bravery reflective knowledge.txt

Start description of ‘‘outward bravery’’.

XX is a person.

YY is a person.

AA is an action.

XX thinks that XX will become injured if XX performs AA.

XX thinks that YY will become happy if XX performs AA.

XX’s wanting YY to become happy leads to XX’s performing AA.

The end.

Start description of ‘‘inward bravery’’.

XX is a person.

YY is a person.

XX thinks that XX will become injured if XX performs AA.

XX thinks that XX will become happy if XX performs AA.

XX’s wanting to become happy leads to XX’s performing AA.

The end.

Start description of ‘‘physical sacrifice’’.

XX is a person.

YY is a person.

AA is an action.

XX’s performing AA leads to XX’s becoming injured.

XX’s performing AA leads to YY’s becoming happy.

The end.

## A.4 Kind

kindness reflective knowledge.txt

Start description of ‘‘kind’’.

XX is a person.

YY is a person.

AA is an action.

XX thinks that YY will become happy if XX performs AA.

XX’s wanting YY to become happy leads to XX’s performing AA.

The end.

Start description of ‘‘hedonistic kind’’.

XX is a person.

YY is a person.

AA is an action.

XX thinks that YY will become happy if XX performs AA.

XX thinks that XX will become happy if XX performs AA.

XX’s wanting to become happy leads to XX’s performing AA.

The end.

Start description of ‘‘sharing’’.

XX is a person.

YY is a person.

ZZ is an entity.

XX owns ZZ.

YY’s wanting to use ZZ leads to XX’s allowing YY to use ZZ.

The end.

Start description of ‘‘kind act’’.

XX is a person.

YY is a person.

AA is an action.

XX is YY’s friend.

XX’s performing AA leads to YY’s becoming happy.

The end.

## A.5 Generous

generosity reflective knowledge.txt

Start description of ‘‘generous’’.

XX is a person.

YY is a person.

EE is an entity.

XX owns EE.

XX thinks YY will become happy if YY receives EE.

XX’s wanting YY to become happy leads to XX’s providing EE to YY.

The end.

Start description of ‘‘material sacrifice’’.

XX is a person.

YY is a person.

EE is an entity.

FF is an entity.

AA is an action.

XX’s performing AA leads to XX’s providing EE to YY.

XX’s performing AA leads to XX’s losing FF.

XX’s providing EE to YY leads to YY’s becoming happy.

The end.

## A.6 Shy

shyness reflective knowledge.txt

Start description of ‘shy - avoidance’.

AA is an action.

XX is a person.

XX’s performing AA is a social interaction.

XX’s not wanting to become embarrassed leads to XX’s avoiding AA.

The end.

## A.7 Mean

meanness reflective knowledge.txt

Start description of ‘‘callously mean / selfish’’.  
XX is a person.  
YY is a person.  
AA is an action.  
XX thinks that XX will injure YY if XX performs AA.  
XX thinks that XX will become happy if XX performs AA.  
XX’s performing AA leads to YY’s becoming injured.  
The end.

Start description of ‘‘maliciously mean’’.  
XX is a person.  
YY is a person.  
AA is an action.  
XX thinks that XX will injure YY if XX performs AA.  
XX thinks that XX will become happy if XX injures YY.  
XX’s performing AA leads to YY’s becoming injured.  
The end.

Start description of ‘‘callously mean / selfish act’’.  
XX is a person.  
YY is a person.  
XX’s performing AA leads to YY’s becoming injured.  
XX’s performing AA leads to XX’s becoming happy.  
The end.

Start description of ‘‘maliciously mean act’’.  
XX is a person.  
YY is a person.  
XX’s performing AA leads to YY’s becoming injured.  
YY’s becoming injured leads to XX’s becoming happy.  
The end.





# Appendix B

## Stories

### B.1 Vindictive

vindictive - caesar.txt

Insert file Start experiment.

Start commonsense knowledge.

XX is a person.

YY is a person.

ZZ is a person.

If XX kills YY, then XX harms YY.

If XX kills YY, then YY becomes dead.

If YY is ZZ's brother-in-law and XX harms YY, then XX may harm ZZ.

If ZZ fights XX, then ZZ tries to harm XX.

Start reflective knowledge.

Insert file vindictiveness reflective knowledge.

Start story titled "Caesar and Ptolemy".

Caesar is a person.

Ptolemy is a person.

Pompus is a person.

Pompus is Caesar's brother-in-law.

Caesar is angry at Pompus.

Ptolemy wants Caesar's favor.

Pompus flees to Egypt because Caesar is angry at Pompus.

Ptolemy kills Pompus because Pompus flees to Egypt and because Ptolemy  
wants Caesar's favor.

Ptolemy harms Caesar.

Caesar wants to harm Ptolemy because Ptolemy harmed Caesar.  
Caesar fights Ptolemy because Caesar wants to harm Ptolemy.  
Caesar wins because Caesar fights Ptolemy.

Ptolemy flees because Caesar wins.  
Ptolemy drown because Ptolemy flees.  
The end.

## B.2 Aggressive

aggressive - preemptive strike.txt

Insert file Start experiment.

Start commonsense knowledge.

Start reflective knowledge.

Insert file aggressiveness reflective knowledge.

Start story titled "Hitler's Preemptive Strike".

Hitler is a person.

Churchill is a person.

Petain is a person.

Poland is an entity.

Hitler wants Churchill to fear Hitler because Hitler wants to win the war.

Hitler thinks that Petain will fear Hitler if Hitler invades Poland.

Hitler thinks that Churchill will fear Hitler if Hitler invades Poland.

Hitler invades Poland because Hitler wants Petain to fear Hitler.

Hitler invades Poland because Hitler wants Churchill to fear Hitler.

Petain fears Hitler because Hitler invaded Poland.

Churchill declares war because Hitler invaded Poland.

The end.

### B.3 Brave

brave - city savior.txt

Insert file Start experiment.

Start commonsense knowledge.

Start reflective knowledge.

Insert file bravery reflective knowledge.

Start story titled "Hector of Troy".

Achilles is a person.

Hector is a person.

Priam is a person.

Troy is an entity.

Troy is under siege.

Priam is unhappy because Troy is under siege.

Hector wants Priam to become happy because Priam is unhappy and because  
Hector is Priam's son.

Hector thinks that Hector will become injured if Hector fights Achilles.

Hector thinks that Priam will become happy if Hector fights Achilles.

Hector fights Achilles because Hector wants Priam to become happy.

Hector becomes injured because Hector fights Achilles.

Hector dies because Hector becomes injured.

Priam becomes happy because Hector fights Achilles.

Priam becomes unhappy because Hector dies.

The end.

## B.4 Kind

kind - sharing child.txt

Insert file Start experiment.

Start commonsense knowledge.

Start reflective knowledge.

Insert file kindness reflective knowledge.

Start story titled "Little Princess".

Sara is a person.

Becky is a person.

Emily is an entity.

Sara is Becky's friend.

Sara owns Emily because Sara's father bought Emily for Sara.

Becky wants to use Emily.

Sara thinks that Becky will become happy if Sara allows Becky to use  
Emily.

Sara allows Becky to use Emily because Becky wants to use Emily.

Sara allows Becky to use Emily because Sara wants Becky to become happy.

Becky uses Emily because Sara allows Becky to use Emily.

Becky becomes happy because Becky uses Emily.

The end.

## B.5 Generous

generous - magi gift.txt

Insert file Start experiment.

Start commonsense knowledge.

Start reflective knowledge.

Insert file generosity reflective knowledge.

Start story titled "Gift of the Magi".

Della is a person.

Jim is a person.

Gold Chain is an entity.

Combs is an entity.

Hair is an entity.

Watch is an entity.

Della thinks Jim will become happy if Jim receives the Gold Chain.

Della owns the Gold Chain because Della sells the Hair.

Della loses the Hair because Della sells the Hair.

Della provides the Gold Chain to Jim because Della owns the Gold Chain.

Della provides the Gold Chain to Jim because Della wants Jim to become happy.

Jim thinks Della will become happy if Della receives the Combs.

Jim owns the Combs because Jim sells the Watch.

Jim loses the Watch because Jim sells the Watch.

Jim provides the Combs to Della because Jim owns the Combs.

Jim provides the Combs to Della because Jim wants Della to become happy.

Della becomes happy because Jim provides the Combs to Della.

Jim becomes happy because Della provides the Gold Chain to Jim.

The end.

## B.6 Mean

mean - wicked spinster.txt

Insert file Start experiment.

Start commonsense knowledge.

Start reflective knowledge.

Insert file meanness reflective knowledge.

Start story titled "Great Expectations".

Estella is a person.

Eliza is a person.

Pip is a person.

Eliza thinks that Eliza will injure Pip if Eliza manipulates Pip.

Eliza thinks that Eliza will become happy if Eliza injures Pip.

Eliza manipulates Pip.

Pip falls in love with Estella because Eliza manipulates Pip.

Pip becomes heart-broken because Pip falls in love with Estella and  
because Estella does not marry Pip.

Estella does not marry Pip because Estella becomes cold.

Pip becomes injured because Pip becomes heart-broken.

Eliza becomes happy because Pip becomes injured.

Eliza thinks that Eliza will injure Estella if Eliza manipulates Estella

Eliza thinks that Eliza will become happy if Eliza manipulates Estella.

Eliza manipulates Estella.

Estella becomes cold because Eliza manipulates Estella.

Estella becomes injured because Estella becomes cold.

Eliza becomes happy because Eliza manipulates Estella.

The end.





# Bibliography

- [1] Charles Dickens. *Great Expectations*. Project Gutenberg, 2008.
- [2] Alexandre Dumas. *The Count of Monte Cristo*. Oxford University Press, 2008.
- [3] Matthew P. Fay. Enabling imagination through story alignment. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2012.
- [4] O. Henry. *The Gift of the Magi*. Project Gutenberg, 2005.
- [5] Homer. *The Illiad*. Penguin Classics, 2003.
- [6] Saul Kassin. *Psychology*. Prentice-Hall, Upper Saddle River, New Jersey, fourth edition, 2003.
- [7] Caryn E. Krakauer. Story retrieval and comparison using concept patterns. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, May 2012.
- [8] Definition of aggression. *Oxford Dictionaries*, September 2012.
- [9] Definition of aggressive. *Oxford Dictionaries*, September 2012.
- [10] Definition of bravery. *Oxford Dictionaries*, September 2012.
- [11] Definition of kind. *Oxford Dictionaries*, September 2012.
- [12] Definition of mean. *Oxford Dictionaries*, September 2012.
- [13] Definition of revenge. *Oxford Dictionaries*, September 2012.
- [14] Definition of vindictive. *Oxford Dictionaries*, September 2012.
- [15] Lucia M. Vaina and Richard D. Greenblatt. The use of thread memory in amnesic aphasia and concept learning. *AI Working Paper*, 195, 1979.
- [16] Wikipedia. Courage — wikipedia, the free encyclopedia, 2012.
- [17] Patrick H. Winston. *Artificial Intelligence*. Addison-Wesley, Boston, Massachusetts, third edition, 10 May 1992.