From Fabula to Fabulous

Using Discourse Structure Relations to Separate Paragraphs in Automatically Generated Stories

Douwe Terluin

July 25, 2008

Contents

Chapter 1

Introduction

> At exactly that moment, his eyes and mouth began slowly to open in a sort of wonder, and slowly he raised his head and became still, absolutely motionless, gazing at the wall opposite with this look that was more perhaps of astonishment than of wonder, but quite fixed now, unmoving an remaining thus for forty, fifty, sixty seconds. Then gradually (the head still motionless), a subtle change spreading over the face, astonishment becoming pleasure, very slight at first, only around the corners of the mouth, increasing gradually, spreading out until at last the whole face as open wide and shining with extreme delight. It was the first time Adolph Knipe had smiled in many, many months
> 'Of course,' he said, speaking aloud, 'it's completely ridiculous.' Again he smiled, raising his upper lip and baring his teeth in a queerly sensual manner. 'It's a delicious idea, but so impracticable it doesn't really bear thinking about at all.'
> From then on, Adolp Knipe began to think about nothing else. The idea fascinated him enormously, at first because it gave him a promise - however remote - of revenging himself in a most devilish manner upon his greatest enemies. From this angle alone, he toyed idly with it for perhaps ten or fifteen minutes; then all at once he found himself examining it quite seriously as a practical possibility. He took paper and made some prelimiary notes. But he didn't get far. He found himself, almost immediately, up against the old truth that a machine, however ingenious, is incapable of original thought. It can handle no problems except those that resolve themselves into mathematical terms - problems that contain one, and only one, correct answer.
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 593

Artificial Intelligence (AI) is a research field best known for its attempts to duplicate intelligent human behaviour. It was considered a victory for AI when world champion chess player Kasparov was defeated by Deep Blue, a computer system designed by IBM. A victory of Machine over Man, providing a glimpse into a future where AI would best mankind at everything.

Whether Artificial Intelligence can ever be considered 'real' intelligence is a debate not likely to end soon. Through technological advances in computer science, we are slowly

reaching the point where a computer has the computing power to process the same amount of information a brain does. Advances in engineering would allow for the construction of a machine as complex as the human body. It is conceivable to combine the two and envision a machine capable of performing intelligent human behavior. But is this enough?

Most people would say no. It is hard to imagine a computer being as capable and *aware* as a human being is. We can accept that a computer can excel at chess, because chess is essentially just math. And life is not.

The Virtual Storyteller project intrigues me because creative behavior like making up stories is not easily described as 'just math'. But if we can create a computer program that makes up stories and possibly interacts with people, we would have clear evidence such behavior can be described as math. I believe that the Virtual Storyteller and other projects like it, can show that perhaps life is just math.

1.1 The original virtual storyteller

The Virtual Storyteller is a multi-agent system which generates stories. The original program was developed by Sander Faas [8], and extended by Sander Rensen [18], both students at the University of Twente. The project is supervised by Mariet Theune.

For this system a character-based approach was chosen, which means that for each character in the story there is a separate agent representing that character. Each of these agents keeps track of its goals and beliefs and plans its actions accordingly [21].

The main advantage to this approach is character-believability [14]. Each of the agents determines which actions will realize its goals, given its current knowledge. The reader will now be able to get an idea why the character is behaving in a certain manner and might identify the goal being pursued more easily.

A drawback of character-based systems is the tendency to generate boring stories. The goals that the characters have could very well have nothing to do with each other, generating no interaction between the characters. Furthermore, the characters do not follow any script or storyline, which will mean they aren't steering the story towards some climax or any ending at all. Soap-operas seem to fare pretty well with never-ending stories, entertaining people with a lot of interaction and conflicting purposes. Here, however, it is not a desired result.

The second version of the Virtual Storyteller was created by Sander Rensen in 2004 [18]. It is capable of generating a limited variety of stories. Within a simulated world different characters interact to create the story, but the characters' actions are partly decided at random. The generated stories are plausible which shows that the concept is sound. However, the simulated world model only includes two characters and four locations and the characters have access to a very limited amount of actions, making each story generated by the Virtual Storyteller more or less the same.

Also, the second version of the Virtual Storyteller has a very limited grasp on natural language. The program presents the events generated by the characters by converting individual events into natural language. This results in a story composed in a uniform grammatical structure. Although this is sufficient to let the receiver of the story know what happened, it can hardly be called entertaining.

A new version of the Virtual Storyteller is being implemented, which is based on the original structure but expands each aspect. A more complex simulated world inhabited by increasingly intelligent and autonomous agents in order to create more complex stories. Several techniques from the field of Natural Language Generation have been implemented in order to present the story in a more natural and more entertaining way.

1.2 This thesis

In this section I will summarize the thesis by providing a short description of each chapter.

In order to generate stories automatically, we first need to examine stories. Chapter 2 will provide an introduction to the field of *narratology* and define the terms used in this thesis. In chapter 3 I will turn to the field of *automated story generation* in order to give an overview of the different accomplishments.

Chapter 4 takes a closer look at the original version and the design of the current version of the Virtual Storyteller. I will discuss the design of the agent structure and what tasks the different agents perform. In the same chapter, an overview of various other research projects is provided in order to show the current status of the implementation. One the one hand, this will show that a lot of work is being done on the various agents responsible for creating a knowlegde structure containing events on which a story could be based. This structure is called *fabula*. On the other hand we will see that a lot of work has been done on creating the agent responsible for converting the story into natural language. One of the main concerns of this thesis is that there is still a big gap between the fabula and the input needed to generate natural language.

In chapter 5 I will describe the fabula in more detail, in order to clearly define what information the story generator provides us and how this information is stored. Next, in chapter 6 I will examine the structure that is necessary to create natural language, namely *discourse structure trees*. The set of discourse relations will be defined and I'll examine how relevant information can be found in the fabula. But this still leaves us with the question of how a discourse structure tree can be constructed from the fabula facts, in such a way that it assists in creating a story.

Chapter 7 on the theory of paragraphs will provide information about the function of paragraphs in texts and will put forth an argument on the relation between paragraph separation and well formed discourse structure trees. I will attempt to validate my method of creating a discourse structure tree based on a fabula, by showing that it can be used for automated paragraph separation. In the end, I can only comment on the effectiveness of the combination of both my method of creating discourse structure trees and my method

of automatically generating paragraphs, but cannot make any claims on the correctness of either method separatly.

Chapter 8 describes the results of an experiment performed to increase insight in the way people would separate text into paragraphs.

Chapter 9 revisits the list of rhetorical relations proposed in chapter 6 and determines how these relations can be identified in a fabula. Next, in chapter 10 I put forward my method of constructing a discourse structure tree using these automatically added rhetorical relations.

In chapter 11 I will evaluate my definitions and methods, by creating a fabula structure and applying my algorithms to that fabula.

Finally, I will summarize this thesis in chapter 12 and discuss what this research means to the Virtual Storyteller project. The last chapter describes projects and questions that have arisen during my research, and which could be useful to further the Virtual Storyteller.

Chapter 2

Narratology

> Then suddenly, he was struck by a powerful but simple little truth, and it was this: *that English grammar is governed by rules that are almost mathematical in their strictness!* Given the words, and given the sense of what is to be said, then there is only one correct order in which those words can be arranged.
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 594

In the story *The Great Automatic Grammatizator* by Roald Dahl, two mathematicians have developed the fastest electronic calculating machine in the world. They were so ordered by the government, in order to satisfy the ever-increasing need of science for rapid mathematical calculation.

One of them, Rudolf Knipe, is unhappy. Even though he has played an important part in designing and creating this computer, which is a momentous achievement, he is unhappy because he is unable to succeed at his true passion: writing. But perhaps, his mathematical genius can somehow assist him in fulfilling his dream. It occurs to Knipe that the machine he designed can be made to create stories.

At first he dismisses the idea because he feels that machines are incapable of being creative. However, the realization that grammar is governed by rules gives him new hope that his goal is attainable. He goes on to find similar rules for and regularities in the plots of stories published in various magazines, which prove to be sufficient to let the machine produce stories.

If we, like Knipe, want to find the rules that govern narrative, to further our understanding of what comprises a narrative and how one should go about automatically generating one, we should turn to research done in the field of Narratology.

2.1 Narrative layers

Narrative is commonly associated with fiction, mostly used to describe the plot or events occurring within a book or a movie. The term 'story' is used interchangeably to represent this. Within the field of narratology different definitions of these terms are used. I will

briefly discuss *narrative layers* (also known as *story layers*) in the following section to clarify the differences between narrative, story and discourse. In chapter 5 these narrative layers will be examined more closely.

2.1.1 Fabula layer

Within the field of narratology, a *narrative* is defined as 'representation of events' [15]. So in order for a narrative to exist, there must have occurred an event or series of events for the narrative to be about. These events are contained within the *fabula layer*. This layer contains *at least* all events that are represented within the narrative. In order to provide an example of a fabula, I present the following narrative:

> Adolph Knipe went home by bus to his two-room appartment. He threw his coat on the sofa, poured himself a drink of whisky, and sat down in front of the typewriter that was on the table. ([6], *page 593*)

It is easy to identify some of the events that occurred here. In our storyworld, the person known as Adolph Knipe has moved from some location which is not his home, to a location which is, and has done so by bus. Having arrived there, he throws his coat on the sofa and proceeds to pour a drink. Next, he moves over to the table and sits down.

Now consider how much more information the fabula layer for this particular example could contain. Going home by bus could also be described as waiting at the busstop, getting into the bus, paying the fare, searching for a seat and talking to another passenger. The level of detail of the fabula is dependent on the degree to which events are dissected. Consider the following segment from Dahl's story:

> Adolph Knipe took a sip of whisky, tasting the malty-bitter flavour, feeling the trickle of cold liquid as it travelled down his throat and settled in the top of his stomach, cool at first, then spreading and becoming warm, making a little area of warmness in the gut. ([6], *page 593*)

In order to describe the consequences of taking a sip of whisky with this amount of detail, the fabula needs to contain information about taking a sip of whisky in at least this amount of detail. In chapter 5 I will discuss in what form this information is stored within the fabula layer.

Besides omitting events from the fabula in the narrative by describing them with less detail, there is another way information from the fabula does not occur directly in the narrative. In a previous example I would venture that no reader assumed that "He threw his coat on the sofa" implied that Rudolph Knipe threw himself on the sofa, while wearing his coat. The assumption is made that he took off his coat before throwing it on the sofa. The events that *are* presented to the reader seem to imply (or at least, lead the reader to infer) that other events have occurred. Such events can not be missing from the fabula, because they represent a step that was taken in the story world and it has enabled the events that followed it. Of course one cannot claim that all events that can be inferred by all readers are present in the fabula, because the fabula contains only one series of events: the one that actually occurred in the story world.

2.1.2 Story layer

Most narratives aren't simple descriptions of an event; especially when multiple events are represented, the narrative will provide information on how these events are related. This information is part of the *story layer*, which is also referred to as the *plot* of a narrative.

In Dahl's story, Mr Bohlen tells Rudolph to take a two week holiday. The next two pages describe how Rudolph Knipe goes home, conceives of the idea of building a storytelling machine and does a lot preliminary research, after which he returns to Mr Bohlen's office. The reader receives no information on how Mr Bohlen spent those two weeks. We can only assume that it wasn't important enough to be part of the story, but we don't assume that Mr Bohlen did nothing (or didn't exist altogether) in this time.

The Story layer contains the information on how relevant events are and what perspective would be most interesting.

2.1.3 Presentation layer

The third and final layer of a narrative contains the information about how the story is presented. In most cases, this last layer will contain the exact text, but the medium for a story will not always be text. I have therefore opted to name this layer the *presentation layer*. In the same vain, it is difficult to aptly describe the person the narrative is presented to. 'Reader' implies that the narrative is presented as text, which might not be the case. Therefore I will use the term *n*arratee.

It is in the presentation layer that we find the difference between story and narrative; a story does not include the information about the presentation. A narrative is the way we perceive a story, but the story itself can only be inferred. According to Chapman (cited in [1]) the difference between story and narrative is a difference between two kinds of time and two kinds of order. It gives rise to what Chatman has called the "chrono-logic" of narrative: Narrative entails movement through time not only "externally" (the duration of the novel, film, play) but also "internally" (the duration of the sequence of events that constitute the plot). The first is called *discourse*, the second *story*.

If the chosen medium is text, the presentation layer contains text and thus information on the discourse structure. So, to summarize, *narrative* is the representation of events, consisting of *story* and *narrative discourse*; story is an *event* or sequence of events (the *action*), and narrative discourse is those events as represented [1].

2.2 Defining narrative

In this section I will take a closer look at what qualities an utterance should have, for it to be considered a narrative.

2.2.1 The minimum

It is important to understand that the term *narrative* encompasses a much larger set of representations than written text only. According to Porter (cited in [1]), narrative refers to *all events in which a medium is used to present an account of events.* Consider the following sentence:

> The little girl fell down

This could be considered a narrative, informing the reader about the little girl falling. Contrast this to:

> The little girl is called Mary

In this example, no event is described and thus is not considered a narrative.

However, some scholars feel that describing only one event is not enough to constitute a narrative (Barthes, Rimmon-Kenan, cited by [1]) They feel that at least two events should be described and that these events should succeed each other in time. Others believe that these events should also be causally related (Bal, Bordwell, Richardson, cited by [1]). Onega requires an even stricter definition: a narrative is a representation of a series of events *meaningfully* connected in a *temporal and causal* way [15]. Porter argues that narratives which describe single events are the basic building blocks for constructing more complex narratives, and should therefore be considered narratives themselves. Prince agrees, noting that a single representation can easily imply the occurrence of multiple events. For example "Mary got an A in history" indicates that Mary spent the day at school yesterday and her teacher informed her that she got an A on her history test. It also suggests that, some time before hearing she got an A, she had to take this test.

2.2.2 Framed narrative

In the previous paragraph we determined that "Mary got an A in history" constitutes a narrative. Consider now that narratives may contain dialogue or speech acts and these speech acts can constitute a narrative on their own. These narratives contained within a larger narrative are called *framed narratives.* Consider the following expansion on the previous example:

> When Mary got home she went to her mother and said: "I got an A in history"

In our example, Mary herself tells her mother about the grade she got. The sentence "I got an A in history" represents an event, so is considered a narrative. The example as a whole therefore represents a sequence of events, one of which is the presentation of a narrative.

2.2.3 Narrativity

The discussion on what is minimally required for a narrative is ongoing. As with many aspects concerning narrative, there is no definitive test that can tell us to what degree narrativity is present [1]. Onega introduces the terms *narrow* and *wide* to account for the different definitions of narrative, while Abbott coins the phrase *narrativity* to describe the degree in which an utterance is seen as a narrative [1].

But not only the minimum of what constitutes a narrative is debatable. Sometimes longer and more complex works aren't as easily recognized as narratives even if they contain narrative building blocks. John Bunyan's Grace Abounding (1666) and T.S. Eliot's The Waste Land (1922) are full of narratives and micro-narratives, yet they do not seem to have the cumulative effect of narrative [1].

In my thesis I shall consider each utterance describing at least one event a narrative. If a narrative describes more than one event, it has greater narrativity. Similarly, if a greater number of causal and temporal relations between the described events are present, the narrative has a greater narrativity.

By regarding it as such, it is possible to use narrativity as a measure to grade the stories produced by automated story generators.

2.3 Temporal aspects of narrative

Previously I've touched on the subject of time and the different role that time plays within narrative. In this section, I will elaborate on these concepts and also show how they are connected with human traits.

2.3.1 Clock time and narrative time

According to Porter, narrative is the principal way in which our species organizes its understanding of time [1]. Narrative predates other ways to organize time and express it. Earlier methods of telling time are in fact similar to narratives, depending on celestial events like the passage of the sun or the succession of the seasons. Here I would like to introduce two different kinds of time: *clock time* and *narrative time*.

Clock time is the time that actually passes while the events described in the narrative are taking place, or *internal time* as Chapman put it (Chapman sited by [1]). In our example of Mary getting home and telling her mother about her history grade, we can imagine that it might take Mary a minute or so to find her mother and another couple of seconds to actually utter the sentence "I got an A in history".

Narrative time (or *external time*) indicates the time that the narrative takes up. The amount of minutes it takes to read this text could be considered its narrative time. This implies that, because reading speed varies per person, narrative time is dependent on the reader. It is therefore more common to express narrative time not in a time measurement, but in a measurement of volume (number of symbols or words).

2.3.2 Narrative perception

When discussing narrative layers earlier, I explained a lot of the events that aren't present in the narrative *directly* are still implied. The tendency of people to infer these events extends beyond filling in the blanks between narrated events. As Porter put it [1], "narrative is so much a part of the way we apprehend the world in time that it is virtually

built into the way we see". This describes the human tendency to interpret everything as narrative. Even when presented a picture, people construct a small piece of narrative to represent it, using narrative templates and narrative formulas to fill in the blanks or encompass the image within a sequence of events.

Some artists use narrative templates to draw the recipient attention, by presenting pictures which are not easily described as a narrative, forcing the onlookers to actively engage in understanding the picture or narrative. Using elements which complicate the construction of a describing narrative is called *narrative* jamming. One can imagine this is what causes intrigue and suspense in books like murder mysteries.

2.4 Importance of narrative

Whatever perceptions are presented, we seek to grasp them in time as well as space. Narrative provides us with this understanding, so it seems reasonable to assume that our narrative perception is continuously adding context to our perceptions. Narrative is the universal tool for knowing as well as telling, for absorbing knowledge as well as expressing it [1].

Because narrative seems so abundant and unavoidable, and seems to play such a large part in how people understand the world around them, some theorist place it next to language itself as *the* distinctive human trait. Children as young as three show narrative capability, when they start putting verbs together with nouns (Porter, in [1]). The earliest childhood memories that are retained by adults occur at this time, which might indicate that the ability to remember is dependent on the capacity for narrative.

This chapter has provided the tools needed for a basic understanding of narratives and what might be needed to create them. I've briefly touched on the subject of different narrative layer, which will be explained more thoroughly in chapter 5. Also, I've discussed that a narrative minimally describes an action. Such an action might even be implied or inferred. Remember that a fabula layer needs to contain all the facts of the narrative, but the narrative must omit some of these facts in order create a presentable narrative. Furthermore narrativity was introduced as a measure for grading stories, and explained narrativity in terms of causal and temporal relations contained within the narrative.

Next, I will discuss different approaches to the automation of storytelling and provide an overview of different automated story generators.

Chapter 3

Automated story generation

> 'One thing I don't quite understand, Knipe. Where do the plots come from? The machine can't possibly invent plots.'
> 'We feed those in, sir. That's no problem at all. Everyone has plots. There's three or four hundred of them written down in that folder there on your left. Feed them straight into the "plot-memory" section of the machine.'
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 600

It would seem that I am not the only one who believes that even complex human abilities, like the ability to create narrative , aren't outside the reach of artificial intelligence. In this chapter I will discuss some aspects of Automated Story Generation and provide an overview of different Automated Story Generators. But first I have an issue to address.

The name Automated Story Generation seems to imply that the only aspect of the narrative that is generated automatically is the *story* (or plot). It makes sense to assume that in this case 'story' refers to an entire narrative and not just to the creation of a plot. The automated storytellers I will discuss in this chapter employ some way of presenting the story they created. This implies that a presentation layer has been added (a barren one at least).

With that out of the way, let us now discuss two different approaches to automated story generation.

3.1 Approaches to automated storytelling

In this section I will discuss different approaches to automated story generation, more specifically the difference between the *author-centric* and the *character-centric* approach. Examples of story telling programs will be discussed in the next section.

3.1.1 Author-centric approach

The author-centric approach to automated storytelling aims to create a storytelling program by emulating a human author. This top-down approach to story generating assumes that a narrative starts as a global concept (or intention). This concept can be

deconstructed into parts, based on either content or presentation. This same process of deconstruction can then be applied to these parts.

For instance, *The Great Automatic Grammatizator* by Roahl Dahl could have started as an idea to create a story about a computer program that can generate stories. The story could then be said to consist of two parts: *how such a machine would be created* and *what influence such a machine would have*. Both of these parts can be further deconstructed.

When generating stories, a program (or person) needs different rules on how to deconstruct different concepts. A set of such rules is called a *story grammar*. One of the more important work done is the research by the Russian folklorist Vladimir Propp[17], who created a classification system for Russian folktales. Propp claims that the plot is the most important aspect of these folktales, noting that folktales are regarded as similar if the contain the same order and content of plot elements.

The main advantage to the author-centric approach is that the narratives created are very likely to have strong story arcs, and in general this approach generates well-formed plots.

A downside to this top-down approach is that it is difficult to maintain the characters' consistency. Actions the characters choose to take are motivated from a story-perspective, which raises issues concerning *character believability*.

3.1.2 Character-centric approach

The character-centric approach relies on autonomous agents to represent characters. These agents interact with each other in some way to create a virtual sequence of events. The assumption underlying this approach is that from these interaction narrative *arises*. Aylett coins the phrase *emergent narrative* [2] to describe a narrative created in such a way.

Unlike author-centric systems, character-centric systems enjoy great character believability because it is relatively easy for each agent to retain consistency in behavior (or to behave in predictable or sensible ways).

Unfortunately, it is hard to ensure the generation of a decent plot within a character-centric system. Each of the agents is unaware of the 'greater purpose' they serve, namely to entertain the narratee.

3.2 An overview of automated storytelling programs

I will briefly discuss the following storytelling systems:

- Automated Novel Writer by Sheldon Klein (1973) [11]
- TALE-SPIN by James R. Meehan (1977) [14]
- Universe by Michael Lebowitz (1985) (Lebowitz, cited by [18])

- Oz Project by Bates (1992) [3]
- Minstrel by Scott Turner (1994) (Turner, cited by [18])
- ConTour by Davenport and Murtaugh (1997) [7]
- Teatrix by Prada et al.(1999) [16]
- STORYBOOK by Callaway and Lester (2001) [5]

This list is no way meant to be complete, but merely aims to provide an idea of the developments made in the field of automated storygeneration. For each of these systems, I will supply a short overview of the capabilities, focusing on the discourse structure and presentation of the stories.

3.2.1 Automated Novel Writer

One of the earliest examples of an automated story generating system is the Automated Novel Writer, created by Sheldon Klein in 1973 [11]. The program is designed to generate murder mysteries, but focuses on generating proper sentences. Because of the focus on natural language generation, the Automated Novel Writer creates relatively long stories of up to 2100 words. See appendix A.1 for an output example.

The Automated Novel Writer mostly relies on the author-centric approach to generate its narratives. The program contains different rules describing how events affect characters and these events combine to create a plot in which one of six available murder events occurs. Because of this rigid structure describing the scenarios, the stories generated by the Automated Novel Writer show very little variation.

It is interesting to note that the Automated Novel Writer does have some character-centric aspect. Each character is assigned a numerical value for a number of traits. The program makes choices on which events occur based on these character attributes (and past events).

3.2.2 TALE-SPIN

In 1977 James R. Meehan presented the interactive story generating program called TALE-SPIN. In his own words, TALE-SPIN is a program that simulates rational behavior by characters in the world [14]. Each of characters are given certain goals and the program is equipped with a problem solver that produces subgoals which will accomplish the main goal. The characters then take the actions necessary to accomplish these (sub)goals.

One of the most important aspects of the TALE-SPIN program is that each character can have one of thirteen different emotional states and the choices these characters make are dependent on their current emotional state. This system allows characters to display their emotions through the choices they make.

TALE-SPIN also involves the user by presenting choices about the plot. This interactive setup results in different outcomes, but this does not seem to be enough. The stories that are generated are said to be 'pointless' and do not take into account which actions would improve the story. Faas ([8]) notes that one of the main problems with TALE-SPIN is that it lacks the capacity to generate complex sentences. See A.2 for an output example.

3.2.3 Universe

Universe, created in 1985 by Michael Lebowitz, is a storytelling program that creates plots for a never-ending soap-opera. Similar to the Automated Novel Writer, characters in Universe have numerical values assigned to a set of traits. But also like the Automated Novel Writer, Universe uses the author-centric approach to automated storytelling.

Universe constructs the plot of the story using *plot fragments*, scenes or events that could take place in the story world. These fragments are defined by several aspects: a list of necessary characters, constraints on the starting situation and the effects on the characters and the situation if the event described in the plotfragment were to occur.

At any given time during the creation of a story, Universe can determine which of the plotfragments could be applied (or which events could occur). It simply checks if the current situation contains the list of necessary characters and satisfies all situational constraints.

In order to choose which plotfragment *should* be applied, it determines which of the applicable plotfragments effects the current situation to best match the goalsituation, derived from the author's goal.

By the writer's own admission, the program's performance is highly dependent on the quality and quantity of plot fragments. In order to address this problem, the program should be capable of creating new fragments in order to avoid repetitiveness.

3.2.4 Oz Project

Bates notes that traditional storytelling media draw much of their emotional power from characters and story [3]. In order for interactive entertainment to have a similar impact, it too must contain rich characters. The Oz project (1992) was created in order to show the validity of various concepts to create such characters.

The most important aspect of creating an interactive environment with believable characters, is to "suspend disbelief" [3]. Ideally, all the characters should be intelligent, emotional, behaving creatures but this is not easy to accomplish. This would be the end goal of basic research in artificial intelligence. The first step, according to Bates, is to create agents that do not actively destroy the illusion of intelligent and emotional behavior. To this end, character agents have been created that contain a basic emotional model, a set of behaviors they can display and a set of features they have.

The focus on interactivity causes the Oz project to feel more like a game than an automated story generator. It does employ a drama-manager to examine the virtual world and actions in order to direct the narrative once a narrative begins to emerge.

3.2.5 Minstrel

Minstrel is a program that generates stories about King Arthur and the Knights of the Round Table. The program was developed by Scott Turner in 1994.

Minstrel is a character-based story generator in which the character agents (or actors) have their own goals. What distinguishes Minstrel from other automated story generators is that a meta-goal is provided to the author. This meta-goal contains some kind of theme or moral and the author tries to incorporate this in the narrative.

The author is a problem-solving agent which uses case-based reasoning. The program's episodic memory contains solutions to problems it has created in the past. The author agent then tries to match existing solutions to the current problem in order to find a solution for the current problem. In order to do this, Minstrel transforms the current problem into a more general description of the situation, recalls an existing solution and determines whether that solution can be adapted to fit the current situation. These methods are called the Transform Recall Adept methods (TRAMs).

3.2.6 ConTour

ConTour is a storyteller system designed by Davenport and Murtaugh [7]. This system's basic unit is a video fragment, which should be combined to form a documentary or narrative. In the words of Davenport, "this prototype system doesn't simulate the story object itself; instead, by selecting the next story element based on the context of what proceeded [sic] it, they simulate the very process of storytelling and story understanding" [7].

Inspired by the concept of relational databases, each of the video fragments is labelled with a variety of keywords. Because of this variety, video segments can be related to each other in different ways (i.e. through different keywords). But in order to understand why this is important, we must first take a look at what is involved when creating a documentary.

Davenport describes the "traditional" process of making a documentary film as consisting of the following steps:

- The filmmakers collect raw material. This includes film footage, audio recordings, archive photographs and text articles.
- The raw material are organized in progressively larger chunks of narrative (shots, scenes and sequences).
- The finished sequences are edited together to create the final "cut" of the movie.

Davenport notes that of the total amount of raw material this process starts out with, only ten percent ends up in the final product. During this process of cutting away material, the product becomes more and more focused and determined [7].

3.2.7 Teatrix

In 1999, as part of the NIMIS project, Prada et al. developed TEATRIX, a learning environment wherein children can engage in interactive story-creation[16]. The main goal of this research was to merge acting, reading and writing into one single environment and to support those engaged with it. This helps children develop their notions of narrative.

Based on real world experience, the collaborative story creation process is divided in three phases: the story set-up, story creation and story writing.

In the story set-up phase a child chooses which scenes to include in a story. Scenes are pre-defined spatial locations with exits that connect to other scenes. Each scene also contains some decor objects. Next, characters are selected. Characters are defined by their name, their social stereotype (e.g. a little girl, an old lady) and their role. TEATRIX allows only six different roles, based on the work done by Propp ([17]): villain, hero, helper, magician, beloved one and beloved relatives. Based on the social stereotype and role, a set of possible actions is determined for each of the characters.

Children can now play the prepared story by selecting to 'become' one of the characters. At any time, the story can be initiated by the child who created the set-up. Software agents will control any characters that have not been chosen by one of the children.

The children interact with the characters (other children and/or software agents) through their own character within the TEATRIX environment. They can talk to each other and manipulate objects. Characters are controlled by selecting actions from their action lists.

When children interact within the TEATRIX environment, there is no guarantee that a coherent narrative will emerge. To address this problem, Prada et al. included the possibility for one of the participants to assume the role of a director. This director has no physical presence within the system, but can see everything that goes on at different scenes and can instruct (or even assume control of) other characters. It is important to note that TEATRIX is intended to engage children in story-creation, and focus less on the quality of the stories.

Each story performed within the TEATRIX environment is stored as a movie and can be reviewed by other children. They can edit the movie and leave notes, describing or criticizing what they have seen.

3.2.8 STORYBOOK

Callaway and Lester set out to create an program able to combine the accomplishments in natural language generation and automated story generation. They note that in both fields extensive research has been done, but that until now little effort has been applied to

filling the substantial gap between narrative plans produced by story generators and the requirements of natural language generating systems [5]. The AUTHOR narrative prose generation architecture has been implemented in STORYBOOK, which contains narrative planning, sentence planning, a discourse history, lexical choice, revision, a full-scale lexicon and a surface realizer. An output example can be found in appendix A.3.

The narrative planner creates a high-level story specification (the information from fabula layer and story layer) which is then structured into paragraph and sentence sized chunks. A discourse history analysis is performed to determine indefinite references and pronominalizations, followed by a lexical choice analysis to increase variety among concepts and event relations. Next, STORYBOOK maps actors and events to semantic/syntactic roles in full linguistic deep structures. It then revises paragraph-sized groups of deep structures via aggregation and reordering. Lastly, a surface realization is performed which produces narrative prose.

The narrative planner is, at the moment, only capable of creating two different high-level stories about Little Red Riding Hood. This simplified planner was created to test the other parts of STORYBOOK. During these tests various parts of the program were applied or skipped, which showed that discourse history and the revision components are extremely important features. Lexicalization was valued less, mostly because the test panel believed a great lexical diversity would be too difficult for children to read.

3.3 Importance of automated story generation

As I discussed the previous chapter, creating narrative is an important human trait. It follows that endeavours in automated story generation might tell us much about human nature. When (if) a computer program is able to create diverse narratives on a par with the work of human writers, it serves as an indication of what artificial intelligence is capable of. But these implications aside, it is interesting to look at possible applications of automated story generation to see if this research is merited.

To examine the importance of automated story generation, we must first examine the importance of story generation. *Entertainment* is the first thing that comes to mind when discussing the importance of stories. Whether the story is presented through a novel, a movie, a conversation or a computer game, it is the story that engages the mind of the narratee. *Education* also employs stories to present information in a way that that important facts are easily extracted and remembered. But why would we want to automate the process of creating narrative?

3.3.1 Why automate the creation of narrative?

As with any human task, an argument brought forth to promote automation is that it frees humans from performing that task. Industrial robots have taken over various tasks like welding, painting, and assembly work. Why not have creative programs write our books and scripts? Callaway notes that most of the research done in the field of natural language generation concerns nonfictional texts like technical manuals [5]. This tendency

can be explained by noting that it is an easier task to automate, but perhaps it is also a task humans desire to be rid off! And perhaps, there is a strong correlation between those two characteristics.

Next, there is the issue of time. The average non-literary fiction writer releases about two books a year. Though the market for books is not insatiable it is mostly limited by the writer's productivity and not by public demand. The cost of developing an automated story generating system that could create books of a 'selling' quality might be tremendous, but this investment would soon be compensated by the difference in production costs per book. It is clear that the costs of running a complex computer program do not compare to a writer's expenses for six months.

3.3.2 Interactive narrative

But most interesting is the advantage that automated story generation provides a way to create an interactive narrative experience.

Computer games of late have shown a shift in attention to storylines. Most of these storylines are fixed and are only used to tie different parts of the game together (*linear narrative*). Sometimes a game contains a certain number of branching alternatives (*branching narrative*). But if the technology exists to create an entire narrative automatically, these games would not be restricted to predefined options. The impact of the user on how the game and the storyline develops would be much greater, which in turn would engage the user to a far greater extend.

Interactivity provides the same added value for educational programs. Davenport remarks that sitting in front of a TV screen, you may appreciate an hour-long documentary; you may even find the story of interest; however, your ability to learn from the program is less than what it might be if you were actively engaged with it, able to control its shape and probe its contents [7].

I've shown various automated story generators, but now it is time to take a closer look at the one this research is intented to improve: The Virtual Storyteller. The next chapter will discuss the history and current state of the Virtual Storyteller.

Chapter 4

The Virtual Storyteller

> They grabbed the sheets and began to read. The first one they picked up started as follows: 'Aifkjmbsaoegweztpplnvoqudskigt&,-fuh-pekanvbertyuio-lkjhgfdsazxcvbnm,peruitrehdjkg mvnb, wmsuy ...' They looked at the others. The style was roughly similar in all of them. Mr Bohlen began to shout. The younger man tried to calm him down.
> 'It's all right, sir. Really it is. It only needs a little adjustment. We've got a connection wrong somewhere, that's all. You must remember, Mr Bohlen, there's over a million feet of wiring in this room. You can't expect everything to be right first time.'
> 'It'll never work,' Mr Bohlen said.
> 'Be patient, sir. Be patient.'
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 601

Creating a computer program that can create fairy-tales is a complex task, not to mention an ambitious one. In this chapter I will elaborate on the origin of this idea and describe the previous two versions. Next, I will provide an overview of the current status of the Virtual Storyteller.

4.1 Virtual Storyteller v1.0

The first version of the Virtual Storyteller was created by Sander Faas [8], during his graduation project carried out from June 2001 until June 2002. The goal of the Virtual Storyteller was to create a new agent for Virtual Music Center (a long term project of the Parlevink research group). Though recognized immediately as the ambitious task it has proven to be, Faas saw the importance of such a project. He chose to focus on the foundations of an interactive storyteller and designed an agent architecture. His aim was to combine existing software for natural language generation with rule-based reasoning and ontology handling.

An examination of several automated story generators led Faas to the following conclusions [8]:

- it is insufficient to randomly fill in a story grammar;

- it is insufficient to let the characters build the story without notion of story structure;
- most story generators are actually *plot* generators;
- Most story generators produce poor language;
- most previous work concerns story *generation*, but does not involve actual story *telling*.

Faas stressed the difference between a story generator and a story teller. A story teller should have capacity to not only *come up with* the story, but also *tell* it. In the following section I will describe the agent structure Faas designed and show how Faas incorporated the task of telling the story. Remember from chapter 2 that the Virtual Storyteller might employ methods other that text to present the story, so I will use the term *narratee* rather than reader.

4.1.1 Agent Structure

Faas designed an agent architecture based on an improvisational theatre structure called *Typewriter* (Joh81, cited by [8]). The *Typewriter* structure involves several actors, one of which assumes the role of a narrator. This actor directly addresses the audience and tells them a story. He also assumes the role of a director, creating a setting, introducing characters and keeping an eye on the story structure. The other actors act out different roles and contribute to the story in this way.

As Faas explains, the appeal of this structure is that the director and other actors have a *shared responsibility* to keep the story interesting and entertaining. The director may use such devices as addressing the audience or introducing new objects, while actors can control the characters directly and influence the story in this way.

Based on the *Typewriter* structure and the intent to provide *spoken* rather than written text, Faas created an agent structure containing Actor Agents, a Director Agent, a Narrator Agent and a Presentation Agent (see 4.1). I will discuss each of these agents and describe the tasks they perform in order to jointly create a narrative.

Director Agent

As in the *Typewriter* setting, the Director Agent controls the story. He might sketch an initial situation and can assign roles to the other actors. Most importantly, he must ensure that the scenes acted out by the other actors connect to the overall story line. He can accomplish this by providing additional information, which benefits the audience, the actors or both.

In order for the Director Agent to safeguard the overall story line, he must have some concept of what a story line is. Faas opted to integrate information about *story grammars* in the Director Agent.

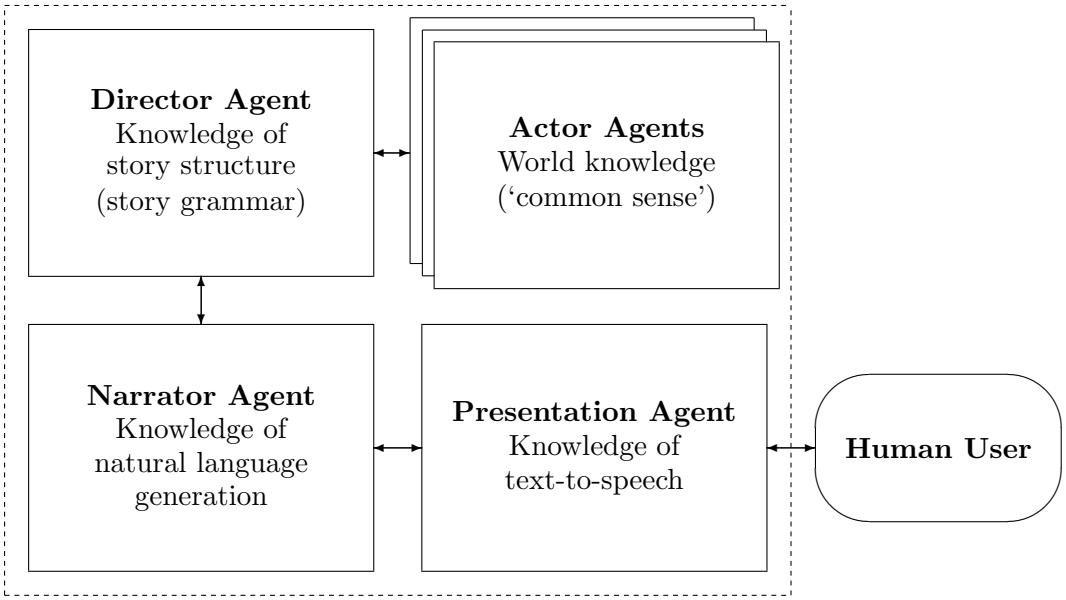


Figure 4.1: Structure of the Virtual Storyteller v1.0

Actor Agents

The Actors take their cues from the Director and act out their parts. They embody one (or more) of the characters and act out the scenes. Mostly, they are responsible for presenting a small part of the story in an entertaining way, but more experienced actors will be able to steer the entire story.

Narrator Agent

The Narrator Agent is responsible for converting the story to natural language. In an interactive storyteller, the Narrator Agent would also translate the user's input into a format which the Director Agent can understand.

Presentation Agent

The text created by the Narrator Agent is provided to the Presentation Agent, who converts it to spoken language and presents the story to the user. This is also the agent that handles the interaction with the user. Faas wanted the Presentation Agent to be more than a mere voice, perhaps presenting it as an (animated) object. A thusly *embodied* Presentation Agent would greatly enhance the user's experience.

4.1.2 Story Grammar

The Director Agent needs information on story grammars to be able to guide the story creation. In order to collect this information, Faas summarized several stories in order to analyse their overall plot outline (see Appendix B). Based on the summaries of these

well-known stories, Faas deviced a simple story grammar. A story always consists of a *begin*, a *middle* and an *end*. The beginning introduces the protagonist by providing a description and name. The middle consists of two parts: first, it introduces the *antagonist* and describes how it creates an undesired (or even harmful) situation for the protagonist. Secondly, it describes how the protagonist overcomes or escapes this situation. The ending is quite simple: it informs the narratee that the protagonist lives happily ever after.

4.1.3 Implementation

The first version of the Virtual Storyteller was implemented in Java, because Faas wanted it to remain compatible with another project, the Virtual Music Centre. To avoid having to construct an agent framework Faas used the Java Agent Development Environment (JADE). To incorporate the rule based reasoning Faas turned to Jess, which is a simple java-based reasoner based on Clips. For more information on the software used in this version, I will refer you the appendices in Faas' thesis [8].

4.1.4 Results

Considering the severe limitations of the first version, Faas made his goal to recreate a simple story. The story could not contain more than one character and the character should have a single goal (provided by the Director Agent). His aim was to create a program that could create stories based on this simple story, meeting the following requirements:

- the setting and the character should be initiated by the Director Agent;
- the same goal should not always result in the same plot;
- the output must be usable as input for the Narrator Agent.

Faas selected the following story:

> Once upon a time there lived a dwarf named Plop. Plop was in the forest. He was hungry. Plop knew that there was an apple at home. Plop walked to his house and went inside. He picked up the apple and ate it. Dwarf Plop lived happily ever after.

Eventually the first version of the Virtual Storyteller was able to produce the following output:

> Once upon a time there was a dwarf. He was called Plop. Plop was in the forest. A apple was in the house [sic]. Plop was hungry. Plop walked to the house. He entered the house. He picked up the apple. He ate the apple. Plop lived happily ever after.

Faas implemented the Presentation Agent using standard text-to-speech software and a visualisation provided by MSAgent, resulting in a cute little Merlin telling the story about the dwarf plop (see figure 4.2).

Figure 4.2: The little Merlin telling the story

4.2 Virtual Storyteller 2.0

In 2003 Sander Rensen selected for this master thesis the goal to expand upon the first version of the Virtual Storyteller [18]. He wanted to focus on generating plots and increasing the range of the generated stories. In the following section, we will point out the shortcomings that Rensen found and in what way he improved upon the original version.

4.2.1 Agent Structure 2.0

Rensen examined the original agent structure as developed by Faas and came to four conclusions regarding the Virtual Storyteller's capacity:

- A story needs structure, but there needs to be room for creativity within this structure;
- A story needs lifelike, believable characters;
- A story needs to adhere to one central theme;
- A narrative text should be well-written and not seem 'mechanical'.

Each of these conclusions inform changes made to the different agents. We will discuss these changes in the following sections.

Director Agent 2.0

Faas had designed the Director Agent to be in charge of the story and it was its responsibility to ensure that the story becomes interesting. Rensen concluded that 'the Director doesn't know whether the information it relates is relevant or interesting' [18] and felt this had to do with the limited story grammar Faas employed. How can the Director Agent determine if a story is interesting, when it has but a vague idea of what a story is? How Rensen expanded the story grammar will be discussed in section 5.2.2.

Actor Agents 2.0

Rensen noted that the first version of the Virtual Storyteller only included one character. Though a single Actor Agent can be considered a proof of concept, Faas felt that it was important to add at least one other character to the story, which means he needed to add an Actor Agent to the system. This would necessitate the system to handle interaction between characters.

In order to involve the Director in the middle part of the story, or to provide the Director with a way to influence the actions taken by Actor Agents. The system Rensen devised was to have Actor Agents *obtain permission for their actions* from the Director Agent. Actor Agents would themselves plan actions they wanted to undertake, but in order to execute any action they would have to request it from the Director Agent. If the request is granted, the action is executed. But the Director agent can deny the request, thus providing the Director Agent a way to block actions that disrupt the story.

The most important addition to the Actor Agents was to grant the Actor Agents *emotions*. Rensen discusses several existing emotional models for agents, selecting the OCC model as the basis for his own. In order for the characters to express emotions, it was necessary for the agents to let their emotions influence their choice of actions. For instance, if a princess hears a strange noise, she might run away in terror if she was frightened already, but she might investigate if confident enough.

Narrator Agent 2.0

Rensen concluded that the language generated by the first Virtual Storyteller was mechanical, but unfortunately had little time to expand the Narrator Agent. The only addition was an expansion of the Narrator's vocabulary, in order to deal with new items and actions added to the rest of the system.

Prestentation Agent 2.0

As we mentioned at the start of this section, Rensen was mostly interested in improving the plots of the stories generated by the Virtual Storyteller. As can then be expected, Rensen did not expand on the way the stories were presented. The little Merlin continued to be the face of the Virtual Storyteller.

4.2.2 Story Grammar 2.0

The simple story grammar Faas designed was insufficient according to Rensen. The 'begin-middle-end structure', though simple and efficient, did not guarantee interesting stories. Moreover, the way the story grammar was implemented gave the Director Agent very little influence. The beginning and end part were very static, while the middle part was generated completely independent of the Director Agent's control.

Rensen discusses several different story grammars, some of which are deemed too simple like the one Faas originally employed, or too complex, like Propp's famous story grammar based on Russian folk-tales (Propp, cited by [18]). In search of a workable compromise,

Rensen selects the work of Greimas, who simplified Propp's original story grammar, to base his own story grammar on. For a detailed description, I refer you to Rensen's thesis [18]. Here, it suffices to say that relatively simple rules or *episodes* are defined. A sequence of these episodes should guarantee that the story builds to a climax and rounds down.

4.2.3 Implementation 2.0

Rensen expanded on the original program, adding his additional modules to the Java-based Jade structure. In order to implement the more complex changes, Rensen turned to *ontologies*. Ontologies, according to Rensen, are a useful tool in the design of a rule-based system, allowing for easier generalization and greater application of rules. Four ontologies were designed: a world ontology, an action ontology, a predicate ontology and one ontology to encompass the system.

4.2.4 Results 2.0

Rensen built a single set of episodic information and created two characters: the princess Amalia and the villain Brutus. Amalia was given the goal to kill Brutus, while Brutus was provided with the goal to capture the princess and lock her in his castle. Here, we present two examples of stories generated by the second version of the Virtual Storyteller:

> Er was eens een prinses. Ze heette Amalia. Ze bevond zich in het kleine bos.
> Er was eens een schurk. Zijn naam was Brutus. De schurk bevond zich in de
> de woestijn2.
> Er ligt een zwaard in de bergen.
> Er ligt een zwaard in het grote bos.
> Amalia loopt naar de woestijn1.
> Brutus loopt naar de woestijn1.
> Amalia ervaart angst ten opzichte van Brutus vanwege de volgende actie:
> Amalia ziet Brutus.
> Amalia loopt naar het grote bos.
> Brutus loopt naar het grote bos.
> Amalia pakt zwaard op.
> Brutus ervaart angst ten opzichte van Amalia vanwege de volgende actie:
> Amalia pakt zwaard op.
> Brutus schopt de mens.
> Amalia steekt de mens neer.
> en ze leefde nog lang en gelukkig!!!

> Er was eens een prinses. Ze heette Amalia. Ze bevond zich in Het kleine bos.
> Er was eens een schurk. Zijn naam was Brutus. De schurk bevond zich in de
> de woestijn2. Er ligt een zwaard in de bergen.
> Er ligt een zwaard in het grote bos.
> Amalia loopt naar de woestijn1.
> Brutus loopt naar de woestijn1.
> Amalia ervaart angst ten opzichte van Brutus vanwege de volgende actie:
> Amalia ziet Brutus.

Amalia loopt naar het kleine bos.
Brutus loopt naar het kleine bos.
Amalia schreeuwt.
Brutus slaat de mens.
Amalia ervaart angst ten opzichte van Brutus vanwege de volgende actie:
Amalia ziet Brutus.
Amalia schopt de mens.
Brutus schopt de mens.
Amalia schreeuwt.
Brutus pakt de mens op.
Amalia schreeuwt.
Brutus ervaart hoop ten opzichte van Amalia vanwege de volgende actie:
Brutus pakt de mens op.
Brutus loopt naar de woestijn1.
Amalia schreeuwt.
Brutus ervaart hoop ten opzichte van Amalia vanwege de volgende actie:
Brutus pakt de mens op.
Brutus slaat de mens.
Amalia schreeuwt.
Brutus gaat het kasteel binnen.
Amalia schreeuwt.
Brutus neemt in het kasteel de mens gevangen.
en de mensen spraken jaren later nog over deze treurnis.

Rensen concludes that both stories seem to adhere to the structure embedded in the episodic information, while still leaving room for variation and interaction between the two characters. He admits that there is a lot of room for improvement, especially where the natural language generation is concerned. Further expansion of possible actions would also greatly increase the number of variation in the generated stories.

Next, I will discuss the current version of the Virtual Storyteller and show it to be even more ambitious than the first two versions.

4.3 Virtual Storyteller 3.0

The current version of the Virtual Story is of course based on the previous two versions, but it has been completely rebuilt from scratch. It is the work of Swartjes [20] and Uijlings [23], which was expanded upon by Hielkema [9] and Slabbers [19]. In the following sections I will discuss the fruits of their labor.

4.3.1 Virtual world

The previous two implementations of the Virtual Storyteller both focused on the behavior of the agents and the overall narrative. Uijlings realized that in order to expand the number of possible narratives and expand the actions agents could perform, it would be necessary to enrich the environment in which the story takes place.

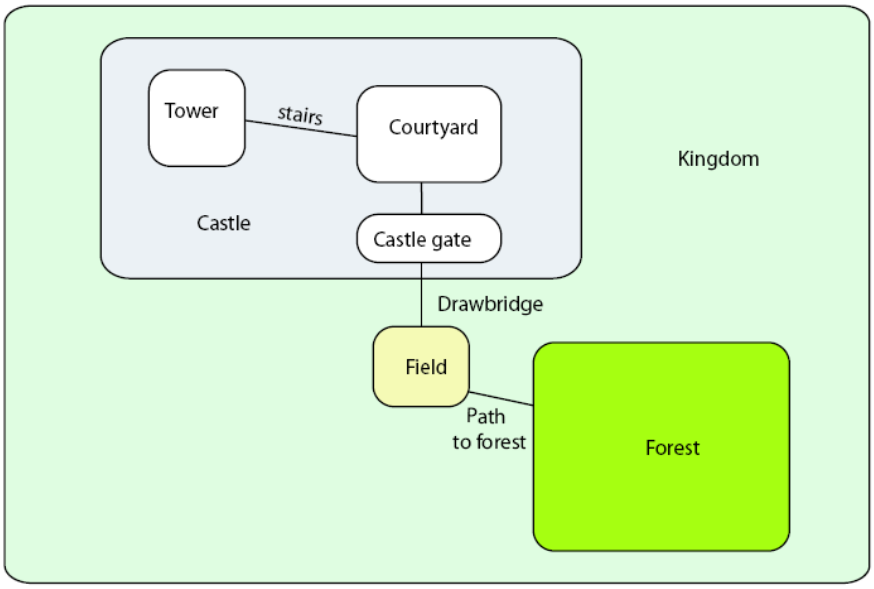


Figure 4.3: An example of part of a virtual world

Rensen employed a simple map, consisting of seven locations, and the characters were able to move from one location to the other. But this is insufficient to represent to effects of other actions. We can imagine the princess running away from the villain without ever leaving the forest. We require the forest to contain multiple locations. This argument applies to all kinds of locations, so it is necessary to create a hierarchy to define locations.

Next, the number of different locations would also have to be expanded. In order to handle a lot of different kinds of locations (and objects, for that matter) Uijlings, in collaboration with Swartjes and Oinonen, designed an ontology of possible locations. By using an ontology it becomes possible to add a lot of different types of locations without having to explicitly define them. Figure 4.3 shows a small part of the kind of world that can now be represented. In a similar fashion they could now define all types of objects (or *entities*) that might be part of a story world. This ontology was named the Story World Core and is the basis for the Virtual Storyteller's common knowledge about objects and attributes. Figure 4.4 shows part of the ontology that was created. It shows that, for instance, a river is a waterway. All waterways are transit-ways, which are defined as regions. For more detail on the development of this ontology, I will refer you to the theses of Uijlings [23] and Swartjes [20].

4.3.2 Actions

In order for agents to manipulate the virtual world, they need to be able to perform actions. This also requires us to define actions and how they affect the world.

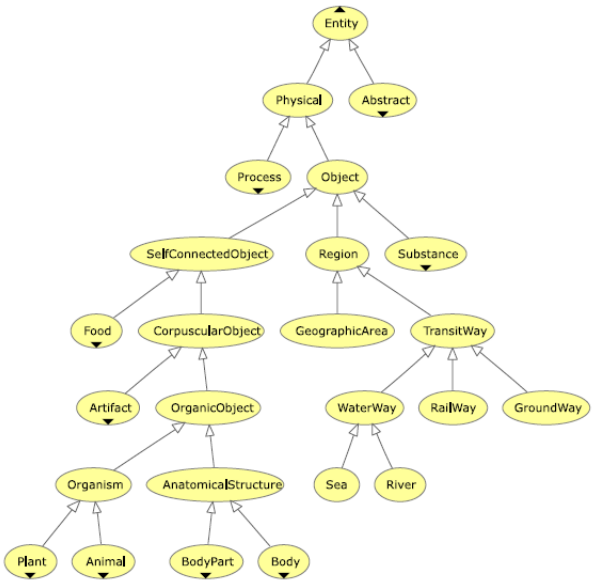


Figure 4.4: Part of the entity ontology

Effects of actions

In order to determine the effects of actions, we regard actions as operators on the state of the virtual world. For each of the actions, we need to define how this action alters this state. Briefly, we rejoin the main characters from Roald Dahl's story as they first operate their automatic grammatizator:

> Knipe smiled an pressed the selector button marked *Reader's Digest*. Then he pulled the switch, and again the strange, exciting, humming sound filled the room. One page of typescript flew out of the slot into the basket.
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 602

Supposing we would want to recreate this story with the Virtual Storyteller, we would not only need to define what a button is, but also what it means to press a button. What has changed in the virtual world after the button has been pressed? In this case it activates the machine, which in turn prints a page, which is then expelled from the machine. But in order to accommodate various narratives an action should not defined so specifically. Instead of defining the action "Pressing the button on the automatic grammatizator" and its effects, we should define 'pressing' an object which would cause the object, in this case a button, to be become pressed. We want characters to be able to press only certain objects though. We wouldn't want a valiant knight to attempt to press a dragon, hoping the dragon will become pressed.

Preconditions of actions

In order to ensure that actions can only be taken when they would be possible in a real situation, we need to formulate constraints for each action. To 'press' an object, you need

a character who is able to press and an object which is 'pressable'. By defining an object's attributes within the story world core ontology and defining preconditions and effects of a variety of actions, our characters become able to perform these actions within the virtual world. We know how the virtual world looks before and after the action takes place. But what happens in between? Is there something in between these two world states, or are actions taken instantaneously? What problems arise if we consider actions instantaneous and what problems arise if we don't?

Scale of actions

It isn't difficult to view Knipe pressing a button as a basic unit of a story, but we could imagine Mr Bohlen wanting to make some adjustments before starting the machine. How is he to stop Knipe from pressing the button if that action is instantaneous? The problem becomes more apparent when we look at another action:

> Adolph Knipe went home by bus to his two-room apartment.

Clearly, Knipe undertakes some kind of movement from work to his apartment, for which the preconditions are that Knipe is able to move and located at his place of employment. The effect would be that Knipe now is located at his home instead, perhaps a bit more tired than he was before taking the action. But what if something was to happen to him on the way home?

This problem could be solved by sticking to instantaneous actions, but defining only the most basic of actions. 'Moving' would be defined as several 'taking a step', which could described as 'lifting one foot and putting in front of the other'. 'Lifting a foot' could be described by 'lifting a foot one inch' several times. It is clear that this kind of definition is unbounded at the bottom and makes the entire system highly complex.

If actions were defined less specifically, we need them to take time. Readers would expect 'going home by bus' to take at least a couple of minutes, perhaps an hour. But now we are left with the problem of how to define actions as non-instantaneous and interruptable.

Intereffects of actions

Uijlings and Swartjes decided on adding an additional element to the definitions of actions: *intereffects*. Figure 4.5 shows the timeline for an action, starting with an interruptable duration. During this time, the action has not had any effects on the virtual world and any changes made to the world (by other agents) might make the preconditions untrue. If this is the case, the action fails. However, if the characters action is still viable, intereffect are applied to the world. This would prevent awkward situations in which actions are hindered long after they have started. Both the interruptable and the total duration are defined per action.

To summarize, action are defined by defined preconditions, intereffects and effects. Uijlings created a taxonomy of possible actions, so that once again different actions can easily added. But how do all these changes to the system affect the agents and the agent structure?

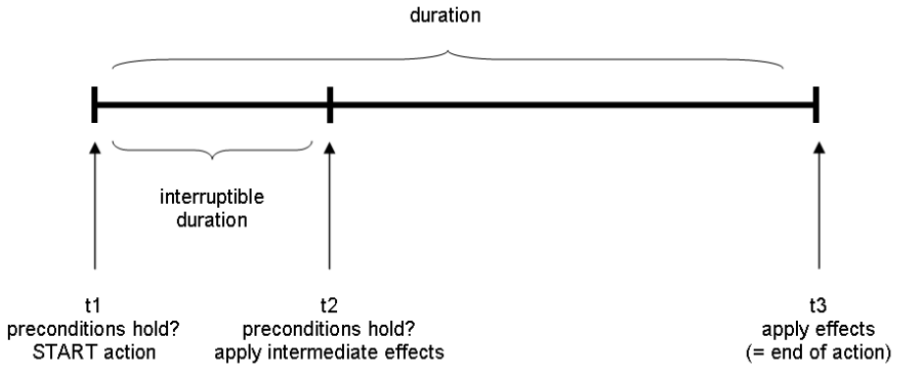


Figure 4.5: Representation of an action

4.3.3 Agent structure 3.0

The agent structure has been slightly redesigned. In the previous two versions of the Virtual Storyteller a Director Agent was used to ensure the created narrative would be interesting. This Director agent would also maintain the state of the virtual world in which the narrative takes place. These are now considered to be two separate tasks. The Director agent has been replaced by a Plot agent and a World Agent. The new agent structure is shown in figure 4.6. I will again discuss each of the agents that comprise the Virtual Storyteller.

World Agent 3.0

The World Agent contains a representation of the virtual story world as it actually is. Figure 4.7 shows the internal flow of the World Agent; it receives from the Plot Agent the actions that character agents undertake and the events that the Plot agents would like to occur. Each of these requests are handled by the Action & Event scheduler which uses the Action database and the Event database to determine what effect different actions and events will have. The effects are applied to the virtual world and an update of the world is sent back. It is through this feedback that the Character Agents and Plot Agent will receive information on the results of their actions, so that they might plan their next move. Notice that the Plot agent can also directly request world changes. I will come back to this when discussing the Plot Agent.

Plot Agent 3.0

The role this agent plays within the Virtual Storyteller is to ensure the plot becomes interesting and entertaining. The Plot agent doesn't accomplish this by telling character agents what to do. Recall that we let the character agents determine on their own which actions they wish to perform in order to maintain character believability and hope that a narrative thusly emerges. Swartjes proposes four ways in which the Plot Agent can direct the story.

- Generate events to interfere with the plans of characters.
  For instance, if the princess decides to flee, but it would be in the interest of the plot

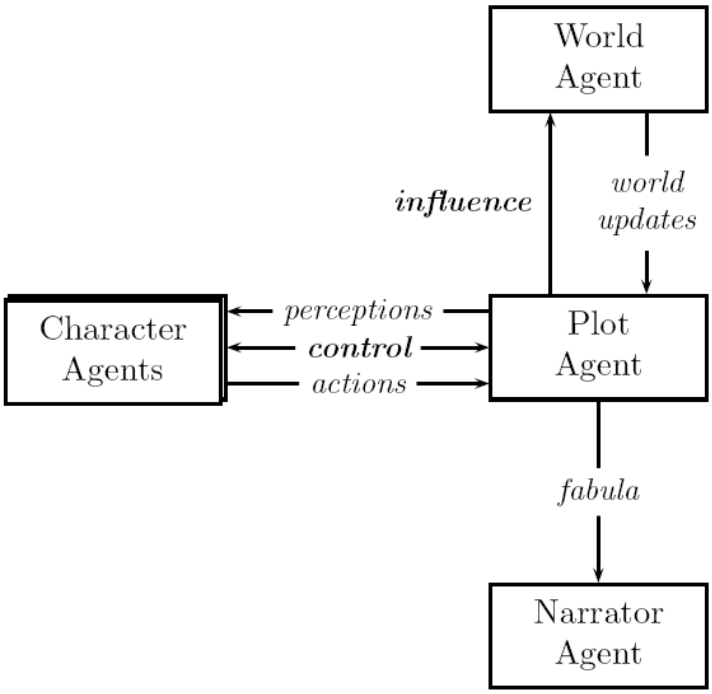


Figure 4.6: Agent structure of the Virtual Storyteller

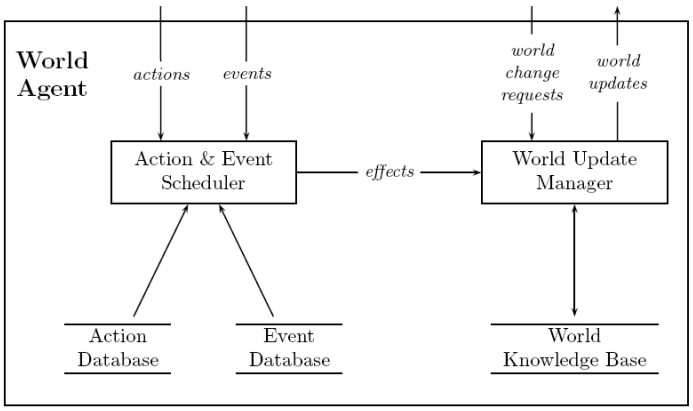


Figure 4.7: Design of the World Agent

to not allow that, the Plot Agent could generate an event like a falling tree. This tree now bars the way and forces the princess to come up with a different action. This method preserves the autonomy of the character agents, but might create disbelief with respect to the virtual world itself.

- Influencing the perceptions of the characters.
  Suppose the hero comes to save the princes from the clutches of the evil villain. A fierce struggle ensues in which one of the two warriors is cut down. The Plot Agent can now prevent the princess of immediately seeing who is the victor, thus increasing her fear and overall suspense. Swartjes warns that changing perceptions, though effective, can very easily become unbelievable and should be used scarcely.

- Changing the setting.
  Changing the setting is mostly used in the start of the plot, to create a surrounding and some starting environment. 'Once upon a time there was nothing' doesn't seem like the right way to start a fairy-tale. But during plot development, Swartjes argues that the character agents should be allowed to improvise and suggest changes to settings. If a character is hungry, the might suggest to the Plot agent to place an apple tree in the forest so that they might walk to it and take an apple. It is the Plot Agent's responsibility to honor only requests which are consistent with the virtual world and the events that have taken place so far. The sudden appearance should not make past events seem unrealistic.

- Directing the characters by suggesting goals or actions
  Swartjes admits that this is not the cleanest of solutions. It transfers the task of maintain character believability from the Character Agents to the Plot Agent. Although this is somewhat opposite of the concept of emergent narrative, it remains a powerful tool in directing a plot.

Figure 4.8 shows the design of the Plot Agent. On the right, we see the Story Builder. This module is responsible for storing all the relevant information in the knowledge structure called the fabula. In chapter 5 this knowledge structure will be examined more closely. For a more detailed explanation of all the different modules, I will refer to Swartjes' master thesis [20].

Character Agents 3.0

The redesign of both the virtual world and the greatly expanded ontology of possible actions neccessitates far more complex character agents than the previous two versions of the Virtual Storyteller. They now require a new way to select one of the many possible actions while taking the complex world into account. But in order to truly let the character agents become more lifelike and believable, they require more complex goals.

Recall from the previous version of the Virtual Storyteller that the two characters, princess Amalia and the villain Brutus, both had a static goal they wished to achieve. It is easy to see that the choice of these 'starting goals' determines most of the narrative. In order for the character agent to become more realistic, they would have to able to create their own goals, alter existing goals and adopt new goals. How these goals are represented in the fabula will be discussed in chapter 5.

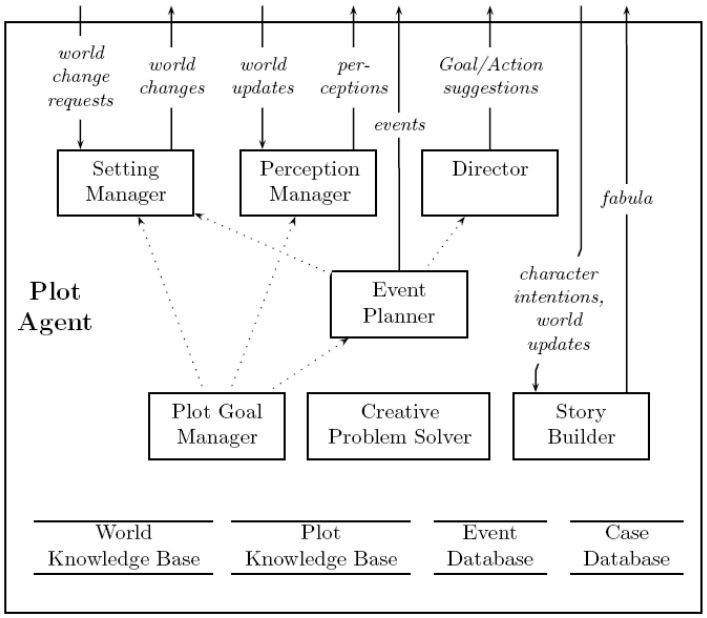


Figure 4.8: Design of the Plot Agent

The next step would be to devise a new way to plan actions based on these complex goals. This requires the character agent to match the effects of actions to the goals it is pursuing. But in our complex virtual world, a series of actions might be required to reach a goal. Moreover, the goal itself might be too complex to match. Thus, the agents need a method to dissect their goals into subgoals. For each of these subgoals the agent can determine if there is a set of actions they can perform which is likely to bring about that goal.

Feedback from the world agent in the form of perceptions should not only motivate character agents to review their planned action sequence, but they should also revaluate their subgoals and perhaps even main goal.

At the moment, the Character Agent software is still being developed. Several different people are looking into creating techniques that might produce lifelike and creative character agent behavior. We need solutions for the the discussed planning algorithms, starting with simple goal-based action planning and ending up, for instance, with integrated case-based reasoning. The character agents could be equipped with their own World agent to maintain their private (and possibly false) state of the virtual world. They probably should be fitted with an emotional model to allow for emotional reasoning like the second version of the Virtual Storyteller employed.

Narrator Agent 3.0

In May of 2005 Feikje Hielkema finished her work on the Virtual Storyteller. At that time, the current version (the third) of the Virtual Storyteller only existed on paper. The form of the information structure that the fabula layer would have wasn't decided on and so

Hielkema had to make assumptions about the information that would be available to her. She outlines a architecture for the language generating part of the Presentation agent, based on a architecture of a natural language generation system (see figure 4.9).

Hielkema's thesis [9] describes a natural language generating system capable of combining several 'short sentences' or clauses into one sentence. This is called *syntactic aggregation* (see the example below). In her thesis she focuses on one of the components of the Presentation Agent, namely a Surface Realizer, capable of producing varied and concise sentences which use different cue words. In order to achieve this, discourse structure relations were used to denote the connection between clauses. A set of cue words was ordered within a taxonomy in order to correctly apply these cue words when combining clauses.

Aside from performing syntactic aggregation, Hielkema wanted to take another step towards creating entertaining narrative. She identified repetition of facts as important factor in readers' experience, namely that unnecessary repetition gives a text an artificial feel. Hielkema expanded the program to perform *ellipsis*, which means to omit repetitive parts of a sentence. She differentiates between five kinds (taken from [9]):

- Conjunction reduction: "Diana entered the desert. Diana saw Brutus" becomes "Diana entered the desert and saw Brutus"
- Right Node Raising: "Diana entered the desert. Brutus left the desert" becomes "Diana entered and Brutus left the desert"
- Gapping: "Diana left the desert. Brutus left the forest" becomes "Diana left the desert and Brutus the forest"
- Stripping: "Diana left the desert. Brutus left the desert" becomes "Diana left the desert and Brutus too"
- Coordinating one constituent: "Diana entered the desert. Brutus entered the desert." becomes "Diana and Brutus entered the desert."

Hielkema has not performed an extensive evaluation of the texts generated by the implemented Surface Realizer, but "feels certain that any narrative, and most other texts as well, are improved by (syntactic) aggregation" [9]. She outlines that there is a considerable amount of work to be done on various aspects of the Surface Realizer (for instance, implementing the Clause Planner and Content Planner) but concludes that a distinctive step forward has been taken towards generating narrative that is grammatical, varied and concise. Shown are two quotes taken from Hielkema's thesis, to show the improvement in the output.

> Diana goes to the desert.
> Brutus goes to the desert.
> Diana is afraid of Brutus.
> Diana goes to the forest.
> Brutus goes to the forest.

Diana and Brutus go to the desert.
Diana goes to the forest, because she is afraid of Brutus.
Therefore Brutus goes to the forest too.

In 2006 Slabbers expanded upon the Narrator model proposed by Hielkema, opting to better incorporate the natural language generation into the story creation. Slabbers designed the 'pipeline' (figure 4.10) for the Narrator Agent and identified several tasks the agent should be able to perform:

- Content determination: deciding what information should be included in the output document, so selecting the relevant information from the non-linguistic input.
- Document structuring: deciding how chunks of content selected in the previous step should be grouped and how the chunks should be related using rhetorical relations.
- Lexicalization: choosing which specific words will be used to express the content selected by the content determination component.
- Referring expression generation: deciding what expressions should be used to refer to the entities used in the output text.
- Aggregation: deciding how the structure created by the component responsible for content determination and document structuring should be mapped onto linguistic structures such as sentences and paragraphs.
- Linguistic realization: converting abstract representations of sentences into real text, by determining the order of the words in the sentence.
- Structure realization: adding mark-up symbols which can be understood by the document presentation component such that this component will present the text in a correctly formatted form.

In order to determine what needed to be done, Slabbers analyzed four fairy-tales. She examined these narratives with respect to fabula elements, sentence and paragraph structures, referring expressions and the applicability of discourse structure relations.

Fabula elements Slabbers examined whether it was possible to identify the different fabula elements in (parts of sentences). I will discuss the fabula elements and this analysis in detail in chapter 5.

Sentence and paragraph structures Slabbers examines the list of rhetorical relations employed by Hielkema and tests if this is sufficient to describe the structure of individual sentences. She recognizes the need to identify a second *elaboration* relation which she calls elaboration2. The list of rhetorical relations she deems sufficient is as follows:

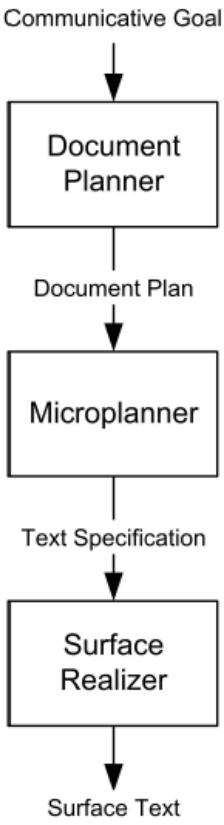


Figure 4.9: Architecture of a Natural language generation system

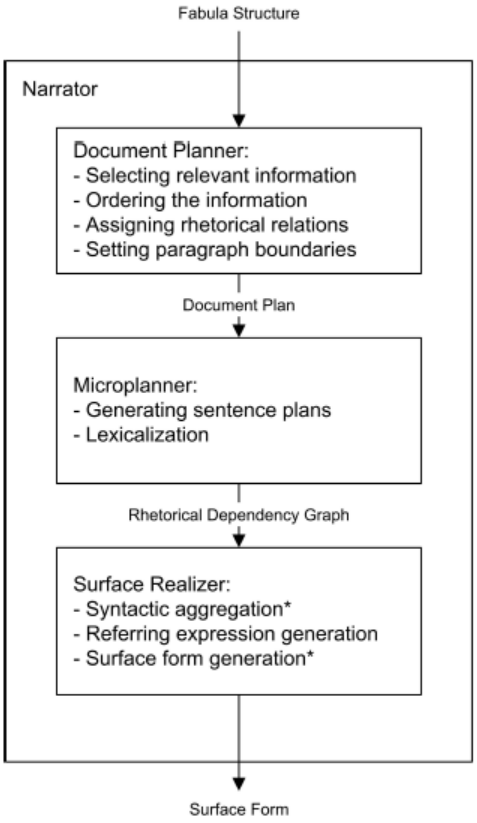


Figure 4.10: Global architecture of a Narrator Agent

- Additive: An additive relation simply enumerates the two related clauses. This is a quite general definition, but was introduced by Hielkema to enable ellipsis. It is my view that this simplification, though effective, will create issues that need to be addressed.
- Cause: A causal relation defines one clause as the cause of another (the other clause becoming the consequence).
- Purpose: The purpose relation resembles the causal relation, with the noted difference that the cause is an action performed to purposefully bring about the consequence.
- Contrast: The contrast relation combines two clauses which contradict each other.
- Temporal: Slabbers defines a temporal relation as a relation that is used to combine clauses which happen at the same time or consecutively. This seems to be a simplification of the possible temporal relations that could be deduced from a fabula.
- Condition: The condition relation combines two clauses, one of which is the condition which needs to be met in order for the other clause to be true.
- Elaboration: The second clause provides additional information about the first clause.
- Elaboration2: Like the elaboration relation, with the noted difference that the second clause provides additional information about *an entity* in the first clause.

The list of rhetorical relations is based on the work of Mann and Thompson (**??**), but contains some simplified or generalized definitions. In both the research done by Hielkema and the research done by Slabbers it was uncertain what the fabula structure would look like. They determined the list sufficient to generates narratives focussing on the task of creating the text. I feel that a more complete and complex list of rhetorical relations is necessary, in order for the Virtual Storyteller to be able to expand its abilities. I will discuss this viewpoint further in chapter 6.

Referring expressions Slabbers analyzed the different fairy-tales in order to determine which referring expressions are used, but also *when* they are used. She identified several characteristics on which she bases rules for automatically adding referring expressions. This is beyond the scope of this research, so I will refer you to Slabber's thesis ([19]).

Discourse structure relations After analyzing whether rhetorical relations are sufficient to describe the structure among clauses, Slabbers tests whether the relations are sufficient to describe the structure among sentences. As noted above, chapter 6 will put forth a more complete list of rhetorical relations. If we ensure that this list encompasses the list used by Slabbers, we can safely assume that if hers is, our list is also sufficient.

As you can see, I've chosen to go into the work done on the Narrator Agent with a bit more detail than employed heretofore. This is because it is between the generation of the fabula and the creation of the narrative that we find the biggest gap in the project. In chapter 5 I will discuss the fabula in more detail, and chapter 6 will provide more information on rhetorical relations, which should provide us with the information needed to bridge this gap. But first, it is left to me to describe the Presentation Agent.

Presentation Agent 3.0

With all the work that still needs to be done on the Character agents, the Plot agent and the Narrator agent, it will come as no surprise that for now the basic text-to-speech software developed by Faas remains sufficient. Within the Virtual Storyteller research group suggestions have been made to further examine animation to present the story, as well as expanding on text-to-speech software by varying pitch and rhythm. Although there is plenty of room for exciting improvements here, for now Merlin will have to do.

4.4 Focus of this research

As this chapter has shown, with time the Virtual Storyteller has become increasingly complex. The overall structure has been well defined so that different researchers can simultaneously work on improving different parts of the Virtual Storyteller.

A difficulty that arises when developing such complex programs is the problem of maintaining functionality without being able to test the program as a whole. Several parts of the Virtual Storyteller were developed simultaneously, while some parts have not been developed at all. One of the aims of this research is to fill the great gap between having a fabula that contains a narrative and having a structure to which the natural language devices designed by Hielkema and Slabbers can be applied.

We are concerned with what information is contained within the fabula structure, what information we need to know in order to create a structure that can be used to create the presentational form, and how we can make that information available. The following chapters will address each of these steps, starting with a closer look at the narrative structure that the Plot Agent creates, followed by a examination of discourse structures.

Chapter 5

Implementation of a Narrative Structure

> This time it began. 'Fewpeopleyetknowthatarevolutionarynewcurehasbeendiscoveredwhichmaywellbringpermanentrelieftosufferersofthemostdreaded diseaseofourtime...' And so on. 'It's gibberish!' Mr Bohlen shouted. 'No sir, it's fine. Can't you see? It's simply that she's not breaking up the words. That's an easy adjustment. But the story's there. Look, Mr Bohlen, look! It's all there except that the words are joined together.' And indeed it was.
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 602

Within the simulated world of the Virtual Storyteller the Character Agents interact and, guided by the Plot Agent, a story emerges. Next, the virtual storyteller must determine the story it needs to tell and how best to tell it. The final step is presenting the story to the user. These are complex and ambitious tasks which require different kinds of knowledge to be stored and require ways to manipulate that knowledge. This complexity places high demands on the implementation of the *narrative structure*. Note that I use the term narrative structure not to refer to the organisation of a narrative, but to refer to a data structure which *contains* all information necessary for the narrative.

In this chapter we will discuss our implementation of a narrative structure. Recall from chapter 2 about narratology that we discern three narrative layers: the fabula layer, the story layer and the presentation layer. In the following sections we will discuss each of these layers and show how they are implemented in the Virtual Storyteller.

5.1 The Fabula Layer

Within the Fabula layer all knowledge about the story-world is represented. It contains which objects and locations occur in the story, which actions were performed and which events occurred. Each of these actions and perceptions are time-coded, so that a sequence of events can be deduced. The princess can first pick up an apple and then eat it. (One could consider objects and environments time-coded as well, but they are probably a bit more durable.)

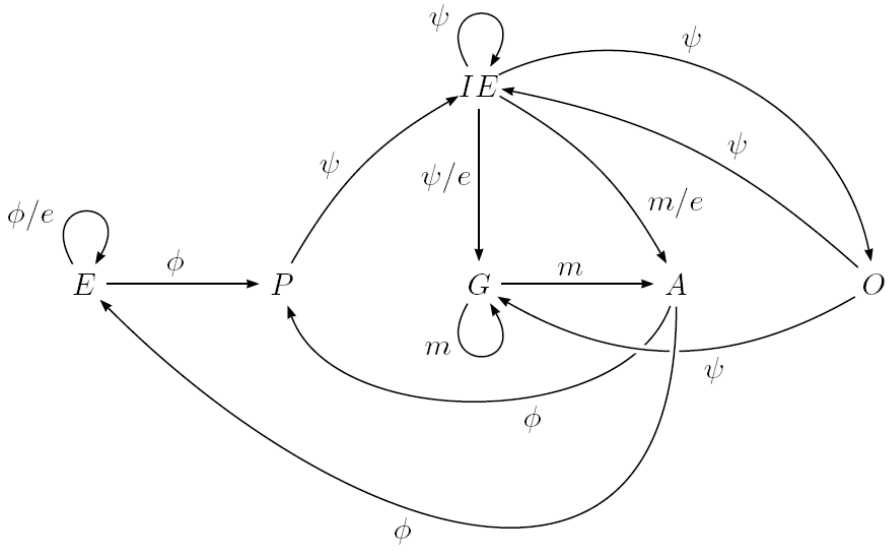


Figure 5.1: Fabula structure, taken from [20]

The fabula also contains knowledge about relations between actions, perceptions and/or internal states of characters. For instance, we can represent causalities between all events using internal states and perceptions [20]. We can not only represent a hungry princess picking up an apple and eating it, but we can represent a hungry princess whose hunger *motivates* her to look for food. Next, the perception of the apple causes the princes to adjust her goal from looking for something to eat to obtaining the apple. This goal motivates taking certain actions, like walking towards the table and picking up the apple.

5.1.1 Implementation

Design of a narrative structure that contains the information that the simulation generates, is no easy task. Several members of the Virtual Storyteller research contributed suggestions and input, but the brunt of the work has been performed by Swartjes ([20]). Most of the following definitions have been taken from his thesis, but bear repeating here.

Fabula elements

The fabula structure is a causal network based on the General Transition Network Model designed by Trabasso et al. (Trabasso1989, cite by [20]). This model is a formal model for story analysis, and can be used to define the structure of a narrative. 5.1 shows the model designed by Swartjes, containing the different types of element that may be present, with all the possible relations which can hold between these elements. First I will elaborate on the different elements, followed by a description of the relations.

- A goal (**G**) is a state or action a character is basing its actions on. I will provide a more complete description of the types of goals shortly.
- An outcome (**O**) is the result of a goal, which can be positive, negative or neutral. Outcome elements don't contain any other information, but only serve to keep track

of the state of a goal.

- An internal element (**IE**) is a single emotion, cognition or belief. It denotes a single element of a character agent's internal state. Note that internal elements, like hunger or sadness, are by definition internal. Acts like crying are actions motivated by internal elements, and it is through these actions that characters can make their internal elements perceivable.
- An action (**A**) represents an action performed by a character. Remember from chapter 4 that actions are considered operators on the virtual world. They have preconditions that must be met, after which inter-effects are applied. If the preconditions still hold after the 'interruptable period', the effects of the action are applied to the virtual world.
- An event (**E**) represents an action performed by 'the world'. To be more precise, it refers to changes to the virtual world which are not the intended effect of character actions.
- A perception (**P**) is the element that enables changes in the virtual world to influence characters. In order to react to a change, they must first perceive it.

Types of goals In order to create believable characters, the character agents must be able to pursue different kinds of goal. Swartjes compiled a list of different types of goals, based on the research done by Trabasso et al. The goals refer to a state of the virtual world, to a certain object or to a certain activity.

to attain a state	to attain an object	to make an activity possible
to maintain a state	to keep an object	to maintain an activity possible
to leave a state	to lose an object	to stop doing an activity
to avoid a state	to avoid an object	to avoid an activity

Note that goals referring to an object can be seen as a referring to a particular state. For instance, if the Princess has the goal of keeping the rose given to her by her true love, this can be represented by the Princess having a goal to maintain the state of the virtual world in which she owns that rose.

Fabula ontology For an overview of the fabula ontology, see 5.2. Now that we defined the different fabula elements, we turn to the possible relations between these elements.

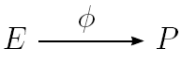
Fabula relations

In order to connect the fabula elements in meaningful ways, we need to define the possible relations between the fabula elements. The definitions and examples below are taken from Swartjes ([20]), with added comments.

Figure 5.2: The fabula ontology

Event physically causes Event An Event can cause another Event, for instance a tree that falls on the bridge causes the bridge to break. Or Sleeping Beauty pricking herself on the spinning wheel causes her to fall asleep. Generating this causality requires either a certain kind of physical reasoning to really model how an Event causes another Event, or it requires a library containing examples of how certain Events physically cause other Events.

Event enables Event An Event ontology has not been designed yet, but Events will probably have preconditions and effects in the same way that Actions do. An Event E_1 enables an Event E_2 if the effects of E_1 match or overlap with the preconditions of E_2.



Event physically causes Perception The Plot Agent determines what characters perceive. When an Event happens, and the Plot Agent sends perceptions of world changes due to this Event to a character, a connection of physical causality can be made.

$$P \xrightarrow{\psi} IE$$

Perception psychologically causes Internal Element The character will appraise the incoming Perceptions using a certain model of Event appraisal. This causes an automatic (psychological) connection between a Perception and an Internal Element that is the result of the Event appraisal.

Internal Element psychologically causes Internal Element This connection arises from further processing of Internal Elements by the character. It arises due to cognitive processes, for instance beliefs or cognitions can lead to emotions and emotions can lead to other emotions.

$$IE \xrightarrow{\psi} G$$

Internal Element psychologically causes Goal The characters need a way to determine Goals based on their beliefs, desires and emotions, done by the Character Agents. It leads to a psychological causation for the Goal, which is an important element for character believability.

$$IE \xrightarrow{e} G$$

Internal Element enables Goal This mainly applies to BeliefElements. When a character believes something to be true, a Goal can be enabled by this belief. For instance, when Lovely believes the lamb has escaped, she can adopt a Goal to go outside and find it.

$$IE \xrightarrow{m} A$$

Internal Element motivates Action Actions that are causally connected to Internal Elements are emotional Actions that are not goal-driven, as discussed in Section **??**. These are Actions like crying and screaming, that are directly caused by an Internal Element and not by a strategic attempt to fulfill a Goal.

$$IE \xrightarrow{e} A$$

Internal Element enables Action This again mainly applies to BeliefElements. When a character believes something to be true, an Action can be enabled by this belief because the belief matches the Action's preconditions. For instance, when Lovely believes there are stairs that lead to the tower, this enables her to walk up these stairs. Whether there actually *are* stairs, and whether the Action succeeds, is out of her control.

$$IE \xrightarrow{\psi} O$$

Internal Element psychologically causes Outcome When a Goal is about attaining or stopping an Internal Element, this Goal will be resolved due to a new Internal Element, which leads to an Outcome caused by this Internal Element. When a character believes that his plan failed or succeeded, this will also lead to an Outcome, namely a positive Outcome when the plan succeeded, and a negative one when the plan failed. When a character stops wanting the Goal, this will lead to a neutral Outcome.

Goal motivates Goal A Goal G_1 motivates another Goal G_2 when G_2 is a sub-goal of G_1, selected for instance based on a library of solutions to problems. For instance, Princess Lovely's goal of wanting the lamb back could motivate a sub-goal of finding the lamb. The library will have to contain examples of goal-subgoal structures.

$$G \xrightarrow{m} A$$

Goal motivates Action Due to the Character Agent's planning algorithm, Actions arise that are goal-driven. When such an Action is planned or executed, it is motivated by a Goal and a motivational causality can be made.

$$A \xrightarrow{\phi} P$$

Action physically causes Perception Characters should be enabled to perceive Actions. When an Action is perceived by a character, a physical causality occurs. Such a causality is also made when world changes are perceived that are the result of an Action.

$$A \xrightarrow{\phi} E$$

Action physically causes Event This is how an unforeseen effect of an Action can be modeled. For instance, when Princess Lovely crosses the castle bridge, it can cause the bridge to collapse. When there is no Event, she ends up on the other side, the normal effect of the Action.

$$O \xrightarrow{\psi} G$$

Outcome psychologically causes Goal When a failed attempt leads to the reinstatement of the Goal, a new episode starts with the same Goal as before, but probably with a different attempt, if the Character Agent is smart.

$$O \xrightarrow{\psi} IE$$

Outcome psychologically causes Internal Element Appraisal of goal success/failure.

Properties of fabula elements

By defining the properties fabula elements can have, we can store all necessary information in the fabula structure. Following the causal relations described above, the ontology will define the properties motivates (m), enables (e), phi_causes (ϕ) and psy_causes (ψ). Also, each fabula element will be annotated with an integer indicating the moment in time when the element was created. This time moment enables a temporal ordering. Actions and Events also contain the time they finished. Lastly, elements contain a property linking them to a character agent.

5.2 The Story Layer

Given that some believable characters witnessed and experienced some fascinating events, we should determine which parts of the fabula comprise the actual story. Which events need to be told and in which order? From whose perspective?

The Story layer contains part of the Fabula layer, more specifically the part that is deemed relevant for this story, and some additional information, further clarifying the *content* and *structure* of the story. All information in this layer should be independent from the type of medium chosen to present the story.

For example, the Virtual Storyteller selects a Storyline Goal so that the story has focus. The Story layer contains all facts which are considered relevant to that goal and contains additional information stating who the protagonist is, based on this goal.

5.2.1 Implementation

Both Hielkema [9] and Slabbers [19] developed the architecture for the Narrator Agent that would be able to fit in the overall architecture for the Virtual Storyteller. Unfortunately, because other parts of the Virtual Storyteller were still unfinished, they were forced to implement their designs in separate applications, in order to provide a proof of concept. In this research I will expand on their designs rather than on their implementations.

5.3 The Presentation Layer

The Presentation layer contains the knowledge necessary to present the story defined in the Story layer to a recipient. (Note that this structure implies that the Fabula and Story layer are independent of the medium used to present the story.)

The Virtual Storyteller's aim is to present the stories as text, so the Presentation layer will contain knowledge about the specific wording and phraseology. In order to structure the generated text, the Presentation layer holds information on narrative discourse structures.

5.3.1 Implementation

As we have seen, the first two versions of the Virtual Storyteller simply convert each fact from the Story layer into Dutch sentences, in chronological order, generating an accurate but prosaic and boring text. Hielkema [9] and Slabbers [19] greatly improved the natural language generation capabilities of the current version, which were outlined in chapter 4.

5.4 Conclusion

The work done by Swartjes has provided us with a structure to capture the fabula layer, using ontologies, rdf and xml to create a widely applicable structure. The work done by Hielkema and Slabbers has shown that in theory it is possible to transform the information

contained within the fabula layer to a interesting text. Now, it is our turn to recreate the work done by Slabbers in the same general way, allowing for more improvements to be later added to different parts of the process.

In the next chapter, we will take a closer look at discourse structures and discourse structure relations.

Chapter 6

Discourse Structure Relations

This chapter introduces the concept of discourse structure and the notion that narratives are coherent because of this structure. First I will examine the characteristics of discourse structures and then address some issues concerning discourse structure relations. I will proceed to present a taxonomy of discourse structure relations to be applied within the Virtual Storyteller.

6.1 Discourse structure and coherence

When analyzing coherent narrative, it is clear that it exhibits some kind of internal structure. Whether the unit of analysis is morphophonemic, a clause, a sentence or a paragraph, units cluster together in specific ways to form larger units [10]. It has been generally assumed that the presence of rhetorical or semantic relationships between segments provides coherence to a discourse. In order to analyze a narrative Hovy makes the following simplifying assumptions:

- A narrative is a structured collection of clauses.
- The clauses are grouped into segments on intentional, semantic, and other grounds.
- The nesting of segments to form larger segments provides the discourse structure.
- At the top level, the discourse is governed by a single node.
- At the leaves, the basic elements are single grammatical clauses.

Hovy explains that these simplifications might be too confining; the bottom three assumptions imply that coherent narratives are represented as a tree structure. Hovy thinks it possible that there are coherent narratives which can't be represented as a simple tree structure. Hovy names the model developed by Trabasso as an example, which refers to the same research that Swartjes has used as a basis to develop the fabula layer model (see chapter 5 on narrative structure). In essence, the Virtual Storyteller translates a complex fabula structure into a tree-shaped structure.

Though not proven here, one can intuitively assume that a narrative that *does* adhere to these simplifications provided above is guaranteed to be coherent.

Although limited in this research to the clause level, Hovy states that he believes the relations he described pertain to both the subclause and the macrostructure levels [10]. Therefore we feel justified to employ Hovy's list to structure our stories.

6.2 Discourse structure relations

If we assume that all narratives can be described as a tree structure of meaningfully related segments, it should be our next endeavor to determine which relations are used to construct narrative. Again we turn to the work done by Hovy [10], who has examined a large amount of research pertaining to discourse structure relations.

Before we can try to answer which relations are necessary to construct a structure within any narrative, it is prudent to discuss this set of relations in more general terms. For instance, is there even such a fixed set? How large must this set be? Do different types of text require different sets of relations?

6.2.1 Parsimonious versus profligate

Hovy describes two views on the problem of determining the number of relations needed to describe English discourse in a discourse structure.

The *profligate position* is held by researchers like Toulmin, Quirk & Greenbaum, Halliday, Hobbs, Mann & Thompson, Knott & Dale, Sanders, Redeker, Asher, and Schank & Abelson. They have produced lists of intersegment relations, which typically contain between five and thirty relations. They feel that there is a closed set of relations that can be found to satisfy all demands that arise from text analysis and natural language generation.

On the other hand there is the *parsimonious position*, which argues there is no closed set of relations that would satisfy the demands. The closer you examine intersegment relationships, the more variability you encounter, until you are forced to account for semantics [10]. Grosz and Sidner propose to avoid the semantic effects on the structure by defining only two types of relations: *dominance* and *satisfaction-precedence* which hold intentional meaning. The only distinction that these two relations offer, is to categorize the relation between segment as being 'equal' (on the same level) or subordinate with respect to the the narrator's intentions.

The parsimonious position may be satisfactory when there's only a limited analysis depth required. Hovy feels, however, that two relations are not sufficient for discourse processing. Some relations between segments are purely semantic, so discarding the semantic aspects to discourse relations would cause problems, especially considering natural language generation endeavors. For example, adding cue phrases is highly dependent on semantic information.

Another issue with the parsimonious position relates to text analysis systems. Within these systems, it is necessary to use semantic knowledge of the segments to provide adequate interpretations of the text.

The arguments against the parsimonious position provided in the paragraphs above exclude this approach for practical application within natural language generation and automated text analysis. So we are left with the task of constructing a closed set of discourse structure relations that will satisfy the demands that are now enforced by text analysis and natural language generation, but can also make a claim to completeness, in order to provide a list that will satisfy demands yet to arise.

Different researchers from different fields have taken to creating such lists, with different applications in mind. It will come as no surprise that these proposed lists display a great diversity. In order to utilize such a list within the Virtual Storyteller, we could examine these lists to determine which is best suited for the tasks.

The problem with this approach is that the Virtual Storyteller is for the most part a natural language generating program and will probably achieve good results using a list originated within the field of NLG (for instance, Mann & Thompson). On the other hand, the Virtual Storyteller aspires to create narratives, which must adhere to rules originating from text and discourse analysis. Therefore, we need a list of relations that considers these, and other relevant fields of research.

6.2.2 Hovy's taxonomy

Hovy presented a taxonomy of different discourse relations, collated from the works of various narratologists [10]. A complete overview of this taxonomy can be found in appendix D.

As Hovy explains, there is no definite proof that this taxonomy is 'correct', but he mentions that the same can be argued for concepts like 'noun' and 'verb'. Correctness is not the issue here, but meeting requirements is. He does express that a certain trust can be derived by considering that Hovy combined the works from approximately thirty researchers from various fields, combining over 350 different relations.

Utilizing a taxonomy to structure these relations has an additional benefit towards the various endeavors in which this taxonomy might be applied: depending on the task, an appropriate level of detail can be chosen. Hovy admits to the problems that arise from using a taxonomy which is unbounded toward the bottom. This places researchers on the slippery slope toward having to deal with the full complexity of semantics [10]. For instance, Hovy identifies *elaboration* as a rhetorical relation, and names five different kinds of elaboration. One of them is *ElabObject*, which he defines as consisting of *ObjectAttribute* and *ObjectFunction*. At this piont, Hovy doesn't provide any further specifications, but we can imagine defining ObjectAttribute as ObjectDimensions, ObjectColors and several others. ObjectDimensions in turn can be defined by ObjectHeight, ObjectWeight etc. The further we go 'down' in the more narrow the relations are defined. Hovy feels that this problem can be circumvented by making the correct choices concerning the level of detail.

6.3 My taxonomy of discourse structure relations

Here, I will present my own taxonomy of discourse structure relations, based on the taxonomy presented by Hovy. Hovy has chosen to identify three main groups: *semantic*, *interpersonal* and *presentational* relations.

Semantic relations are relations based on the content of the clauses. Both clauses are actual storyworld events or descriptions. Interpersonal relations are used to infuse the writer's opinion into the text, and used to address the reader directly. Presentational relations are used to improve the readability, based only on syntax.

Because I am mostly interested in the structure that is brought about by the information held within the fabula, I will leave the interpersonal and presentational relations aside and focus on the semantic relations.

For each semantic rhetorical relation I will provide a description and indicate its position in the taxonomy under 'location'. For the leaves I will provide some examples and name typical discourse markers or *cue phrases*.

In each of the descriptions I will use the terms *nucleus* (N) and *satellite* (S) to indicate the two elements that the rhetorical relation combines. The nucleus is the main part of the combination, while the satellite is related to the main part in some way, depending on the rhetorical relation used. In the case of an elaboration, for instance, the satellite is elaborating the nucleus, instead of the other way around. In the examples provides below, I will show the satellite in italics.

6.3.1 Elaboration

First I will discuss the different relations that denote *elaborations* or clarifications of some sort. The satellite should not introduce new objects [13]. Hovy identifies combinations in which the satellite elaborates on a certain object contained in the nucleus (ElabObject), elaborates on a part of the nucleus (ElabPart), provides a more general or a more specific description (ElabGenerality), identifies the nucleus (Identification) or provides a summary or conclusion (Restatement). Below, I will provide short definitions for each of these relations.

Name	ElabObject
Location	Elaboration - ElabOject
Description	The satellite provides additional information about an object present (or inferentially accessible [13]) in the nucleus. This additional information concerns either an attribute of the object or the object's function
Consists of	ObjectAttribute ObjectFunction

Name	ObjectAttribute
Location	Elaboration - ElabOject - ObjectAttribute
Description	The satellite presents additional detail about an object which is presented in the nucleus (or inferentially accessible). The additional information about the object must provide an attribute of this object.
Examples	The house, *an imposing building*, dominated the street. The king wanted to rule the world. *He was a vicious man.*
Cue words	in fact, actually, indeed, at least, i.e., viz.

Name	ObjectFunction
Location	Elaboration - ElabOject - ObjectFunction
Description	The satellite presents additional detail about an object which is presented in the nucleus (or inferentially accessible). The additional information about the object must provide an function this object can perform.
Examples	The cat lay in her lap, *purring.* The boys ran around, clearly *enjoying themselves.*
Cue words	clearly

Name	ElabPart
Location	Elaboration - ElabPart
Description	The satellite provides additional information about a set, process or object that is present in the nucleus. It provides additional information concerning either a member of the set, a step in the process or a part of the object.
Consists of	Set-Member Process-Step Whole-Part

Name	Set-Member
Location	Elaboration - ElabPart - Set-Member
Description	The nucleus presents a set of objects or concepts and the satellite presents one of the members of this set.
Examples	The children were starving. *The eldest hadn't eaten in a week.* The soldiers marched on, *some of them singing as they went.*
Cue phrases	One of them, the eldest, the first

Name	Process-Step
Location	Elaboration - ElabPart - Process-Step
Description	The nucleus refers to a process and the satellite elaborates on a single step of that process.
Examples	Mom baked a cake. *Applying the frosting was her favorite part.* Ralph enjoyed flying, *even checking in at the airport.*
Cue phrases	including

Name	Whole-Part
Location	Elaboration - ElabPart - Whole-Part
Description	The nucleus refers to a object, of which a part is presented in the satellite.
Examples	The car was running perfectly, *the engine humming steadily.* The house, (*the hallway in particular), was very tidy.*
Cue phrases	specifically, in particular

Name	ElabGenerality
Location	Elaboration - ElabGenerality
Description	The nucleus contains a general description or reference to an object or abstraction, and the satellite identifies this object or provides an instantiation of the abstraction.
Consists of	Genl-Specific Abstr-Instance

Name	Genl-Specific
Location	Elaboration - ElabGenerality - Genl-Specific
Description	The nucleus contains a generality, which the satellite specifies.
Examples	His excuses, *say the breakdown of his car*, never seemed plausible. The children like the animals, *particularly the monkeys*
Cue phrases	one time, for example, for instance, in particular, e.g.

Name	Abstr-Instance
Location	Elaboration - ElabGenerality - Abstr-Instance
Description	The nucleus contains an abstraction, and the satellite provides an instantiation of that abstraction.
Examples	We used to have races - *we used to have relays.* His existence, *specifically his presence here*, was quite annoying.
Cue phrases	for example, for instance, in particular, specifically, e.g.

Name	Identification
Location	Elaboration - Identification
Description	The nucleus contains an object and the satellite provides additional information about the identity of that object. This can be done by either providing more specific or more general information about the subject.
Examples	*The company commander, who was Captain Madison*, assembled his men and announced their mission. Timur was a 14th century military leader. *He was also known as Tamerlane.*
Cue phrases	namely, that is to say, also known as

Name	Restatement
Location	Elaboration - Restatement
Description	The satellite provides a rewording of the content of the nucleus. Restatements are used to summarize the content of the nucleus or to present a conclusion.
Consists of	Summary Conclusion

Name	Summary
Location	Elaboration - Summary
Description	The satellite provides the same content as the nucleus, but presents a shorter version.
Examples	That clock doesn't go; *it's not working.* Each argument was fatal to the other; *both could not be true.*
Cue phrases	or (rather), in other words, that is to say, i.e., in brief, and now

Name	Conclusion
Location	Elaboration - Conclusion
Description	The satellite provides the same content as the nucleus, but mentions explicitly what is left implicit in the nucleus.
Examples	After forgetting his pasport, losing his suitcase and missing his plane, *we may conclude that Richard's holiday was off to a bad start.*
Cue phrases	I will sum up by saying, I shall conclude by saying, If so, if not, that implies, you can conclude from that

6.3.2 Circumstance

The second type of semantic relation Hovy identifies are *circumstances*. The satellite could indicate where the nucleus takes place (Location), when it takes place (Time) and what might be taking place at the same time (ParallelEvent). Hovy also identifies Means, Manner and Instrument, but after consulting the sources used by Hovy there appear to be two different kinds of relations: the satellite indicates how the event in the nucleus takes place (Manner) or with what it takes place (Instrument). The circumstance presented in the satellite must of course be present, and not yet unrealized [13].

Name	Location
Location	Circumstance - Location
Description	The satellite provides additional information about the location where the nucleus takes place. This not only includes literally identifying the location, but also describes the position, direction, movement and passage of the events or objects in the nucleus.
Examples	Alice looked up, *and there stood the Queen in front of them.*
Cue phrases	then, so, for, but, yet, still, and then, and there, and thus, and so, and yet, at that time, soon afterwards, till then, in that case, in that way, as far as, where, wherever, everywhere, like, by (means of).

Name	Time
Location	Circumstance - Time
Description	The satellite provides information about the time at which the nucleus occurs.
Examples	The churchbells rang *at precisely twelve o'clock.*
Cue phrases	meanwhile, subsequently, previously, now, and then, at that time, soon afterwards, till then, precisely

Name	Means
Location	Circumstance - Manner
Description	The satellite describes the manner in which the action contained in the nucleus is performed.
Examples	Keep on subtracting the difference, *and in that way you will arrive at the correct figure.*
Cue phrases	and thus, and in that way

Name	Instrument
Location	Circumstance - Instrument
Description	The satellite provides the instrument(s) used in the nucleus.
Examples	He hit the intruder as hard as he could, *wielding the umbrella as a bat.*
Cue phrases	and thus, and in that way

Name	ParallelEvent
Location	Circumstance - ParallelEvent
Description	Both nucleus and satellite occur at exactly the same time and contain verbs (in order to be an event).
Examples	The waiter dropped his tray. *At that time, the princess made her getaway.*
Cue phrases	meanwhile, now, at that time, precisely

6.3.3 Sequence

Mann and Thompson describe sequential relations as *multi-nuclear*, meaning the relation combines more than two clauses [13]. In this research however, we will assume that the sequential relations only combine two clauses. A sequence of more than two clauses consists of multiple sequences containing two clauses. For instance, a sequence consisting of four clauses (c1,c2,c3 and c4) will be represented by a set of three relations (sequence(c1,c2), sequence(c2,c3) and sequence(3,4)). How these relations are represented in the tree, depends on which of the clauses is the nucleus.

Hovy describes different kinds of *sequences*. The satellite could follow or precede the nucleus in time (SeqTemporal), in location (SeqSpatial) or in rank (SeqOrdinal).

Name	SeqTemporal
Location	Sequence - SeqTemporal
Description	The satellite occurs (directly) after the nucleus.
Examples	The prince kissed the princess, *and then set her upon his horse.*
Cue phrases	and then, then, and so, soon afterwards

Name	SeqSpatial
Location	Sequence - SeqSpatial
Description	A spatial relationship between clauses indicates that they are located next to each other and can only be applied when an order of some sort can be imposed. This means that the content of the clause must have a location in the storyworld.
Examples	The supermarket, *the one right next to the bakery* is the only one for miles.
Cue phrases	was located, can be found next to

Name	SeqOrdinal
Location	Sequence - SeqOrdinal
Description	I'll define ordinal [1] sequences as clauses that are not spatially or temporally related, but do have a specific (well-known) order.
Examples	China has the largest population, *followed by India.*
Cue phrases	followed by,

6.3.4 Cause/Result

Hovy identifies two different aspects of causal relations: *volitionality* and *whether the cause or the effect is more important for the story.*

First of all he differentiates between causal relations which employ a *volitional* action. Consider the following example:

> "The Princess spit in the villain's face, because she hated him."

The action (spitting) is motivated by hate. Even though one could consider actions taken in the heat of a moment as not volitional, we consider this a volitional action. In contrast, consider:

> "Because the arrow pierced his heart, the prince died"

Dying is not an action the prince *chooses* to take after the arrow pierces his heart. Of course it isn't always easy to determine if an action is volitional or not.

Secondly, there is the issue of the main part of the combination. You can imagine the cause being more important for the story (and thus being the nucleus) because this cause also motivates other actions. Or perhaps the effect is the more important clause, being the cause or motivation for yet more actions.

Considering both *volitionality* and *importance of cause and effect* provides us with four different causal relations, which will be defined below. That leaves one type of causal relation undiscussed, in which the motivation for the action is not currently present but relates to a future goal (Purpose).

Name	C/RVol
Location	Cause/Result - C/RVol
Description	A volitional action (or else a situation that could have arisen from a volitional action [13]) is present in either clause (nucleus or satellite), the other clause contains a motivation or reason
Consists of	VolCause VolResult

Name	VolCause
Location	Cause/Result - C/VolCause - VolCause
Description	A volitional action (or else a situation that could have arisen from a volitional action [13]) is present in the nucleus, the satellite contains a motivation or reason
Examples	Alice said nothing, *because she didn't want another argument.* Alice was standing with her hands ready, *for she was any moment expecting him to fall.*
Cue phrases	(and) so; and + therefore; for; (because), as, since, in case, seeing that, considering, with, through, by, at, as a result of, because of, in case of

Name	VolResult
Location	Cause/Result - C/VolCause - VolResult
Description	A volitional action (or else a situation that could have arisen from a volitional action [13]) is present in the satellite, and the nucleus contains a motivation or reason
Examples	Alice didn't want to begin another argument, *so she said nothing.* Alice was expecting him to fall, *so stood with her hands ready.*
Cue phrases	(and) so; and + therefore; for; (because), as, since, in case, seeing that, considering, with, through, by, at, as a result of, because of, in case of

Name	C/RNonVol
Location	Cause/Result - C/RNonVol
Description	A causal relation other than a volitional action is present in either clause (nucleus or satellite), and the other clause contains a motivation or reason.
Consists of	NonVolCause NonVolResult

Name	NonVolCause
Location	Cause/Result - C/RNonVol - NonVolCause
Description	A causal relation other than a volitional action is present in the nucleus, and a motivation or reason can be found in the satellite.
Examples	The prince bled to death, *because the arrow pierced his heart.* The tree flourished, *as a result of the frequent rains.*
Cue phrases	(and) so; and + therefore; for; (because), as, since, in case, considering, with, through, by, at, as a result of, because of, in case of

Name	NonVolResult
Location	Cause/Result - C/RNonVol - NonVolResult
Description	A causal relation other than a volitional action is present in the satellite, and a motivation or reason can be found in the nucleus.
Examples	The bridge hadn't been repaired in years, *and finally collapsed.*
Cue phrases	(and) so; and + therefore; for; (because), as, since, in case, seeing that, considering, with, through, by, at, as a result of, because of, in case of

Name	Purpose
Location	Cause/Result - Purpose
Description	The satellite motivates the action presented in the nucleus, but does so referring to a goal being pursued.
Examples	The prince rode off in the direction of the village, *because he wanted to mary the farmer's daughter.*
Cue phrases	in order that, so that, in order, so as, for, for the sake of, with the aim of, for fear of

6.3.5 GeneralCondition

If the nucleus describes a general rule, the satellite can be used to identify a condition that needs to be met (Condition) or an exception to this rule (Exception).

Name	Condition
Location	GeneralCondition - Condition
Description	The nucleus contains a general rule or an unrealized situation, and the satellite describes a condition that has to be met in order for the rule to apply or the situation to be valid.
Examples	*The ends of his mouth might meet behind*, and then I don't know what would happen to his head. The two brothers will fight to the death, *if they ever meet again.*
Cue phrases	if, provided that, as long as, (and) then, (and) in that case, unless, (or) otherwise, or else

Name	Exception
Location	GeneralCondition - Exception
Description	The nucleus contains a general rule or an unrealized situation, and the satellite describes an exception that prevents that situation from becoming a reality or prevents that rule from applying.
Examples	The king will surely win the day, *unless he loses his head.*
Cue phrases	unless

6.3.6 Comparative

Comparative relations are used by writers to relate the events they are currently describing to similar events, in order to stress their similarities (Equative or Analogy) or to stress their differences (Contrast). We can also use a comparative relation to combine possible outcomes with a current action (Otherwise).

Name	Equative
Location	Comparative - Equative
Description	The nucleus contains an object, about which an attribute or function is presented, or which performs an action. The satellite contains a different object with that same attribute or function, or performing that same action.
Examples	She was a fair maiden, *just like her mother once was.*
Cue phrases	like, while, when

Name	Contrast
Location	Comparative - Contrast
Description	The nucleus contains an object, about which an attribute or function is presented, or which performs an action. The satellite contains a different object without that attribute or function, or not performing that same action.
Examples	The eldest princess was kind to everyone, *unlike her younger sister.*
Cue phrases	unlike, in contrast to

Name	Otherwise
Location	Comparative - Otherwise
Description	The nucleus presents an unrealized situation that if it were to become true, would prevent the unrealized situation presented in the satellite from coming true.
Examples	The prince must find the princess in time, *or she will surely perish.*
Cue phrases	otherwise, or else

Name	Comparison
Location	Comparative - Comparison
Description	The nucleus is compared to the satellite, in order to note similarities and/or differences. Both the differences and the similarities are important for the story.
Examples	The little girl fled the village, *like her mother did twenty years earlier.*
Cue phrases	and + similarly; (and) so, thus, as, as if, like, the way

Name	Analogy
Location	Comparative - Analogy
Description	The satellite presents an object with similar functions to an object present in the nucleus.
Examples	The prince fought *like a lion.*
Cue phrases	like, as if

6.4 Conclusion

Now that we have taken the list of rhetorical relations that Hovy provided and added descriptions for each of them, it is time to see how we can add these relations to the fabula. In the next chapter I will discuss some general aspects of adding rhetorical relations to the fabula, and compose rules for each of the rhetorical relations.

Chapter 7

Theory on Paragraphs

In chapter 4 I explained how the Virtual Storyteller operates and how it creates the knowledge structure called the fabula which contains all information about events and objects in our virtual world. As mentioned at the end of chapter 6, the goal of this research is to bridge the gap between the fabula, discussed in chapter 5 and a discourse structure, discussed in chapter 6.

The problem arises that there is no given method to validate a method which creates a discourse structure based on a fabula. This results from the larger issue which ails the Virtual Storyteller project: the only way to evaluate stories, is by letting people evaluate them. But a the moment several of the parts needed to create actual text are not completed, which leaves us with the difficult task of devising another method of testing the different parts. In our case, we must assume that a 'sufficient' fabula will be created. Furthermore, we must assume that a 'well-formed' discourse structure will be correctly converted into a narrative. This leaves us with the difficult task of evaluating a method to create a 'well-formed' discourse structure based on a 'sufficient' fabula, without knowing what 'sufficient' and 'well-formed' actually mean. In order to attempt this task, I'll focus on a certain aspect of generated narrative: *automatically generated paragraph separation.*

In this chapter I will explain why I have chosen to implement automated paragraph separation in the Virtual Storyteller. In order to do this, I will first discuss what paragraphs are and what function they fulfill. Next, I will put forth arguments why paragraph separation is important and how this ability can be used to evaluate my method of constructing discourse structure based on a fabula.

7.1 What are paragraphs?

A *paragraph* is a unit of written narrative, containing any number of sentences. The paragraph boundaries are made visually apparent, at the least by starting the first sentence of a paragraph on a new line. Other visual clues are sometimes employed, like leaving a empty line before, indenting the first sentence, capitalizing the first words or enlarging the first letter of a new paragraph.

Often a paragraph is defined as a part of written text that 'contains a single semantic element' [25], 'is comprised of sentences which are mutually connected in a logical way' [24] or 'consists of a coherent group of sentences which together form an expression' [4].

First, I will briefly talk about the role paragraphs play in nonfictional text, in order to better grasp the role they might perform in fictional text. The main difference between these two types of narrative, is information about what the narrator wishes to achieve with the narrative.

7.1.1 Nonfictional text

Tiggeler [22] claims that within nonfictional text[1] a well-formed paragraph should hold to the following three constraints:

- *The paragraph is constructed logically.* The information is coherent and important information is found in the first or last sentence.
- *The paragraph addresses exactly one issue.* A paragraph contains a tight group of sentences that form one coherent concept.
- *The first sentence of a paragraph should introduce the subject of the paragraph.*

According to Tiggeler [22] the central aim of any text is either descriptive, explanatory, evaluating or advising. For each of these kinds of texts, Tiggelaar provides examples and outlines a sequence of subjects that need to be addressed. Considering that well-formed paragraphs should address exactly one issue (the second constraint proposed by Tiggeler), these sequences of subjects each provide a sequence of paragraphs.

Tiggeler explains that such templates as he provides need not be followed strictly; they assist writers in gathering and structuring their thoughts before committing them to paper. On the other hand, Tiggeler does mention that one should have a reason to deviate from the provided templates, to indicate that these templates should not be easily discarded (or perhaps to stress the usefulness of these templates).

Tiggeler puts forth a very prescriptive view on paragraph separation, which might not wholly apply to all nonfictional texts, let alone to fictional texts. Burger and de Jong [4] present a less strict view on paragraphs in their book. They propose the following constraints:

- The main sentence should be placed at an accessible location;
- Vary the length of your paragraphs;
- Distinguish between paragraph and paragraph-groups.

I will discuss each of these in turn.

[1] Tiggeler has written a prescriptive guide to writing 'zakelijke' text which translates poorly to English. The book contains examples of business correspondence, reports and proposals and even though nonfictional text encompasses more than text affiliated with business, 'nonfictional' seems a reasonable translation.

Location of the main sentence

Burger and de Jong agree with Tiggeler that the first sentence of a paragraph is a convenient place the most important sentence. Readers want to know what this paragraph is about as quickly as possible [4] so you shouldn't keep them waiting.

But sometimes, it can be a good idea to place the most important sentence at the end of a paragraph. For instance, if the paragraph is working its way to a conclusion, the readers become more engaged in trying to determine what the paragraph is leading up to.

As a third alternative, they suggest letting the second or third sentence be the most important one. The first sentence should now be used to indicate how this paragraph is connected to the rest of the text, or what its function is [4].

Varied paragraph length

The length of paragraphs is mostly dictated by their content, but varying the length is possible and desired [4]. Reader apprehension decreases when a paragraph is too long. Burger and de Jong propose the following rules: *A paragraph should not address more than one issue* and *If a paragraph would become too long, separate it in parts while maintaining coherency.*

A paragraph can never be too short. Burger and de Jong state that, though a lot of writers consider three sentences a minimum to constitute a paragraph, they feel that sometimes shorter paragraphs are warranted and can sometimes be as short as a single word.

Too many small paragraphs will make a text seem eroded and incoherent, but when used in moderation short paragraphs can increase the tempo of a text.

Paragraph and paragraph-groups

A paragraph always starts on a new line, sometimes also indicated by inserting a blank line and/or indenting the first sentence. Burger and de Jong also recognize the difference between paragraphs and paragraph-groups. If certain paragraphs are closely connected to each other, one can employ one paragraph marker (such as indentation) to distinguish them from each other, but employ another (mostly an empty line) to identify them as a set belonging together.

Though the difference between paragraphs and paragraph-groups is very interesting and might be very useful in structuring stories, it is beyond the scope of this research to focus on this difference.

Conclusions

From the constraints on the paragraphs and the strictness of the document templates, one could asses that one thing is paramount within writing nonfictional text: *clarity.* Not

using well-formed paragraphs is as detrimental to the readers' understanding of a text as using bad. punctuation

Nonfictional texts contain information that the writer wants to confer to the reader. The easier it is for the reader to extract *and remember* this information, the better the writer will have achieved his intention. If the information is presented in a clear and structured way, not only is it easier for the reader to *find* the important information, but also to identify the structure of the information. Identifying this structure makes it easier to remember the information.

Another argument why nonfictional texts should adhere to appropriate templates is that readers who read many of these texts will become familiar with the template. They will have an easier time extracting information from texts that fit a familiar template.

Paragraph separation's main purpose in nonfictional texts is to clarify the text's internal structure, in order to facilitate the reader's apprehension of the information and its structure. To summarize the constraints presented above, a clarifying paragraph structure would adhere to the following rules:

- The information in a paragraph is coherent;
- The main sentence should be at an accessible location, usually at the beginning or at the end of a paragraph;
- A paragraph addresses exactly one issue;
- If a paragraph is too long, divide it into several smaller paragraphs;
- Vary paragraph length.

Constraints like these provide criteria on which to grade automatically generated paragraph separation. But I have only considered nonfictional text. To what extent does this conclusion apply to fictional texts? Which additional constraints arise and which ones will be ineffective?

7.1.2 Fictional text

For any nonfictional text, we could claim that the writer's intention can be considered to *convey information*. This is of course a generalization and many other motivations could apply, but I've argued in the previous subsection that conveying information is an important enough aspect to defend considering paragraph separation solely as a method to improve this conveyance.

Similarly, we could claim that for any fictional text, the writer's intention is to *entertain the reader*. Again, the information contained within the narrative should be conveyed with clarity in order for the reader to have a complete understanding. By definition, entertaining information causes entertainment, if conveyed correctly.

This view has one obvious oversight: the entertainment provided by reading fictional texts isn't just derived or determined by the information contained in the text; the way the story is presented is of vital importance. This implies that the presentation of the story would determine how well the entertainment contained within the story is conveyed *whilst* being entertaining itself. We can imagine that there are situations where a writer would have to sacrifice clarity for presentational entertainment. Likewise, there could be situations in which the clearest presentation and presentational entertainment coincide. It is beyond the scope of this thesis to examine both types and dissect their interaction fully. Therefore, I've chosen to focus on clarity, noting that it is important to account for a later inclusion of other factors.

7.2 Structure and paragraph separation

To summarize, paragraphs are used, amongst other reasons, to clarify the structure of a text. If we would want to develop a method to automatically apply paragraph separation to automatically generated stories which clarify their structure, we must address two issues: *How can we automatically detect the structure of an automatically generated story* and *how does paragraph separation relate to this structure*?

7.3 Well-formed discourse structure

At the start of the chapter I explained that it is difficult to evaluate any method of creating a discourse structure based on a fabula. In this chapter I've shown that well-formed paragraphs should clarify the structure of a narrative and determined that we will focus on this aspect of paragraph separation. But how will this guide the constructing of a well-formed discourse structure?

The assumption I put forward is that *a well-formed discourse structure is necessary for good paragraph separation.* This means that if a well-formed paragraph structure can be generated automatically, the underlying discourse structure is also well-formed. If, however, the generated paragraph structure isn't well-formed, we could blame either the method of separating paragraphs or the discourse structure itself.

In addition to generating a discourse structure based on the fabula, I will also develop a method of applying paragraph separation to the discourse structure. In a way, this doesn't reduce the problem of evaluating the discourse structure, but only adds the burden of paragraph separation. I would argue that it is at least easier to determine if the discourse structure is well-formed *with regard to paragraph separation* than it is to determine if the discourse structure is wholly well-formed.

It is then left to each of us to wonder if a discourse structure which is well-formed with regard to paragraph separation, combined with my assumption that good paragraph separation necessitates a well-formed discourse structure, implies that such a discourse structure is indeed well-formed.

7.4 Conclusion

We would like the Virtual Storyteller to employ the use of paragraphs in order to clarify the structure of the story it is conveying. In order for the Virtual Storyteller to do this, it must have some idea of *how* the story is structured. It becomes the main goal of this research to develop a way to automatically determine the structure of a narrative, when only provided information on a fabula level.

The facts on fabula level seem to resemble clauses; the basic units of a text. In order to structure a fabula as a text, it seems reasonable to define relations between fabula facts as we would define relations between clauses with discourse relations. Hovy notes that 'several theories of interclausal relations have been quite productive in suggesting new and powerful ways to plan coherent paragraphs of text automatically from information stored in computers in various non-linguistic ways' [10].

In chapter 9 I will discuss my method for identifying discourse relations between clauses, and chapter 10 will describe my method of using these relations to construct a well-formed discourse structure. But first I will examine how people apply paragraph separation and how (or if) this corresponds to the lessons learned in this chapter.

Chapter 8

Separating Paragraphs

The goal of this research is to have the Virtual Storyteller separate paragraphs in a way that makes sense. In the previous chapter we discussed the formal side of the function of paragraphs and how the focus of this research will be to use paragraph separation to enhance the clarity of a story. Remember that we set out to evaluate the ability to separate paragraph, in the assumption that this reflects on the quality of the method for generating discourse structures.

Having determined what the Virtual Storyteller should care about as a writer and which rules seem to govern paragraph separation, it is interesting to see how people separate paragraphs. Does following the rules always produce unique results? Do people follow these rules at all? And how can the Virtual Storyteller perform the task of separating paragraphs?

In this chapter we will discuss an experiment with which we have tried to uncover how people separate paragraphs. We will use the results of this study to validate the rules on paragraph separation discussed in the previous chapter.

8.1 How do people separate paragraphs?

As described in chapter 2, narratives are everywhere. We can therefore assume that people in general have encountered a vast quantity of narratives in their lifetime and have formed ideas about how a narrative is structured. More specifically, literate people have encountered a great amount of texts and have linked aspects of the presentation to the kind of discourse it presents and the kinds of content it contains. For example, if a text is adorned with an illuminated letter we do not expect the text to provide stock-information. Likewise, we would not look through The Financial Times in hopes of finding an appropriate bedtime story.

An experiment was conducted in this research project to provide insight into the way people separate paragraphs. We are interested in finding out *where* someone would place paragraph boundaries and how much these decisions vary between people. Most importantly, what are the principles on which people do this and how can we use this knowledge in the virtual storyteller.

8.2 Method

We approached several subjects by email. Attached to the email was the fairy-tale "The Wicked King" (Figure 8.2) but the story was edited so that each sentence started on a new line to obfuscate the original paragraph structure. The story was selected because it is relatively short and fairly unknown. Of the six subjects that participated, two are students at the University of Groningen and four are university graduates. Each of them was between the ages twenty-five and thirty. None of them had a background in linguistics.

De Slechte Vorst

Er was eens een vorst die de hele wereld onder zijn heerschappij wilde brengen. En niet door wijsheid, goedheid, eerlijkheid, rechtvaardigheid en vrede, maar door oorlog te voeren. Hij was zo slecht dat hij eerder een schepsel van de duivel dan van God leek te zijn.

Te vuur en te zwaard overwonnen zijn soldaten de buurlanden; er stroomden rivieren van bloed en tranen, en van vele trotse steden bleven slechts smeulende puinhopen over. De wrede huurlingen van de vorst slachten zelfs hele kuddes vee af; honden en katten die zich niet snel genoeg uit de voeten maakten, werden door hen aan het zwaard geregen.

De slechte vorst was meedogenloos. Hij was uitsluitend genteresseerd in een zo groot mogelijke oorlogsbuit uit de paleizen, kerken en kloosters die hij overwonnen had. Zijn schatkamer stroomde over met goud, edelstenen, parels en juwelen: zelfs gouden deurknoppen liet hij loswrikken! Met al deze schatten bouwden zijn architecten schitterende paleizen, burchten en kastelen, zodat iedereen die ze zag van angst voor hem vervuld zou raken en zou buigen voor zijn almacht. Maar daar was de vorst nog niet tevreden mee. 'Ik wil de machtigste vorst op aarde worden!' zei hij iedere dag.

Toen hij alle omringende landen veroverd had, liet hij hun koningen vastbinden achter zijn gouden triomfwagen, en hij liet ze vertonen in zijn hoofdstad alsof het wilde dieren waren. Hij genoot van de bewondering van zijn onderdanen.

Na vele veldtochten stonden zijn legers aan de oever van de zee: er was geen enkel land op aarde dat nog niet veroverd was. De vorst besloot nu dat op ieder stadsplein een standbeeld van hem moest worden opgericht. Maar ook dat was hem nog niet voldoende eer. Hij besloot dat zijn standbeeld ook moest worden opgericht in alle kathedralen, kerken, moskeen en synagogen van de overwonnen landen. De priesters verzetten zich als enige tegen dit voorstel. 'U heerst over de mensen in vele landen, machtige vorst, maar u bent niet de heerser over de hemel. God is nog altijd machtiger dan u. We kunnen uw bevel niet opvolgen: er kunnen geen standbeelden van u worden opgericht in onze godshuizen!'

> De vorst bedreigde hen met de marteldood, maar de priesters lieten zich hierdoor niet afschrikken.
>
> Nu had de vorst nog maar één doel voor ogen: hij zou ook de hemel veroveren. Hij begon onmiddellijk met de voorbereidingen. Zijn architecten maakten bouwtekeningen voor een luchtschip, dat voortgetrokken zou worden door duizend adelaars: daarmee zou hij de hemel bestormen. Zijn wapenmeesters maakten kanonnen die zo groot waren dat ze het hemelse paleis in puin konden schieten.
>
> Zeven jaar lang was het leger in de weer met de voorbereidingen voor de aanval. Toen was alles klaar: een groot luchtschip voor de legertroepen, en een kleiner luchtschip voor de vorst, die zelf de aanval wilde inzetten.
>
> De militaire kapel speelde. De maarschalken staken lange redevoeringen af voor de aanval werd geopend. De adelaars die het merkwaardige luchtschip zouden voorttrekken zaten nog in hun kooien.
>
> De soldaten stelden zich in een mijlenlange rij op om plaats te nemen in hun luchtschip. Ook de vorst zelf, gehuld in zijn vergulde harnas, steeg aan boord. Hij stond op de brug, klaar om het eerste bevel te geven.
>
> Plotseling kwam er een wolk muggen aanzwermen: alle muggen stortten zich op de vorst en zoemden oorverdovend. De machtige vorst trok zijn zwaard en sloeg om zich heen, maar hij sloeg alleen maar in de lucht.
>
> Een van de muggen wist zijn helm binnen te dringen en stak de meedogenloze vorst in zijn oor. De vorst begon om hulp te roepen, sprong uit zijn luchtschip en zette het op een lopen in de richting van het bos. Ondertussen rukte hij alle onderdelen van zijn wapenuitrusting los: poedelnaakt rende hij weg, voor het oog van al zijn legertroepen; en aldoor werd hij omstuwd door een horde boze muggen.
>
> De vorst verdween in het bos, en werd nooit meer teruggezien. Het hele leger lachte en slaakte een zucht van opluchting. De machtige vorst die de hemel wilde bestormen ging roemloos ten onder!

The test subjects were asked to read the story and insert blank lines where they thought it was appropriate to start a new paragraph. They were asked to submit their answer through email and were invited to accompany their submission with comments about the story and the task. Each of the subjects was told to take all the time they needed, but were instructed that there were no right or wrong answers and that they need not spend a lot of time deliberating. The subjects were not told to which end they should employ paragraph separation.

	Subject #1	Subject #2	Subject #3	Subject #4	Subject #5	Subject #6
1	1	1	1	1	1	1
2	1	1	1	1	1	1
3	1	2	1	1	1	1
4	1	2	1	2	1	2
5	1	2	1	2	1	2
6	1	2	2	2	2	2
7	2	3	2	2	2	2
8	2	3	2	2	2	2
9	2	3	2	2	2	2
10	3	4	3	3	2	2
11	3	4	3	3	3	2
12	3	4	3	3	3	3
13	3	5	3	3	3	3
14	4	5	4	4	3	3
15	5	5	4	4	3	3
16	5	5	4	5	3	3
17	5	6	4	5	3	3
18	6	6	5	5	4	3
19	6	6	5	5	4	3
20	6	6	5	5	4	3
21	6	6	5	5	4	3
22	7	7	5	5	4	3
23	8	7	6	6	4	4
24	8	7	6	6	5	4
25	8	7	6	6	5	4
26	8	7	6	6	5	4
27	9	7	6	6	5	4
28	9	7	7	6	6	4
29	9	8	7	7	6	4
30	9	8	7	7	6	4
31	9	8	7	7	6	4
32	9	8	7	7	6	4
33	9	8	7	7	6	4
34	10	8	7	7	6	4
35	10	9	8	8	7	5
36	10	9	8	8	7	5
37	10	9	8	8	7	5
38	10	9	8	8	7	5
39	11	9	8	8	7	5
40	11	9	9	9	7	5
41	11	10	9	9	7	5
42	12	10	9	9	7	5

Figure 8.1: Six subjects separating the story “The Wicked King” in paragraphs

Note I only removed the blank lines, removed indentation that indicated the paragraph boundaries and started every sentence on a new line. Any textual clues or discourse markers that might indicate paragraph boundaries remained, and the subjects’ decisions might very well be influenced by them.

8.3 Results

Figure 8.1 shows an overview of the results. For each of the lines of the story, the figure shows to which paragraph the subject feels that line belongs. The subjects have been ordered by number of paragraphs used, sorted from high (12) to low (5).

On average a paragraph consists of about 4,85 sentences. Subject #1 separated the text into twelve paragraphs, on average 3,5 sentences per paragraph. Subject # 2 used ten

Paragraph containing sentences:	Subjects in agreement
1 - 3	4 + 6
1 - 5	3 + 5
7 - 9	1 + 2
10 - 13	1 + 3 + 4
28 - 34	3 + 5
29 - 34	2 + 4
35 - 39	3 + 4
35 - 42	5 + 6
40 - 42	3+ 4

Figure 8.2: Agreement on paragraph separation in the story "The Wicked King"

paragraphs, averaging 4,2 sentences per paragraph. Subject # 3 and # 4 both separated the story into nine paragraphs, averaging 4,7 sentences per paragraph. Subject # 5 used seven paragraphs averaging six sentences per paragraph and subject #6 only used five paragraphs, averaging 8,4 sentences per paragraph.

There is no paragraph boundary identified by all of the subjects and there are only a few paragraphs that at least two subjects are in complete agreement about. These are shown in figure 8.2.

These figures show that the six subjects have a quite different view on how this story should be separated into paragraphs. It would seem that there is not a lot of agreement when considering entire paragraphs. Therefore, let us take a closer look at the similarities and differences on the level of sentences.

8.4 Analysis on sentence level

In Figure 8.3 the results of the six subjects have been combined in one table. For each pair of sentences the number denotes how many of the subjects put these sentences in the same paragraph. The figure shows this score for each sentence and several of its surrounding sentences, ranging from four sentences back (-4) to eleven forward (+11).

In the following sections several aspects of the results will be discussed. We are interested in finding out whether the results match the rules presented in the previous chapter and whether other rules might emerge. These are the five rules selected in the previous chapter:

- The information in a paragraph is coherent;
- The main sentence should be at an accessible location, usually at the beginning or at the end of a paragraph;
- A paragraph addresses exactly one issue;
- If a paragraph is too long, divide it into several smaller paragraphs;
- Vary paragraph length.

	-4	-3	-2	-1		+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11
1					-	6	5	3	3	1	-		-	-	-	-
2				6	-	5	3	3	1	-	-	-	-	-	-	-
3			5	5	-	4	4	2	-	-	-	-	-	-	-	-
4		3	3	4	-	6	4	2	2	2	1	1	-	-	-	-
5	3	3	4	6	-	4	2	2	2	1	1	-	-	-	-	-
6	1	2	4	4	-	4	4	4	2	1	-	-	-	-	-	-
7	-	2	2	4	-	6	6	2	1	-	-	-	-	-	-	-
8	2	2	4	6	-	6	2	1	-	-	-	-	-	-	-	-
9	2	4	6	6	-	2	1	-	-	-	-	-	-	-	-	-
10	2	2	2	2	-	5	4	3	-	-	-	-	-	-	-	-
11	1	1	1	5	-	5	4	1	1	1	1	-	-	-	-	-
12	-	-	4	5	-	5	2	2	2	2	1	1	1	1	1	-
13	-	3	4	5	-	3	3	3	2	1	1	1	1	1	-	-
14	-	1	2	3	-	5	4	3	1	1	1	1	1	-	-	-
15	1	2	3	5	-	5	4	1	1	1	1	1	-	-	-	-
16	2	3	4	5	-	5	2	2	2	2	2	-	-	-	-	-
17	2	3	4	5	-	3	3	3	3	2	-	-	-	-	-	-
18	1	1	2	3	-	6	6	6	4	1	-	-	-	-	-	-
19	1	2	3	6	-	6	6	4	1	-	-	-	-	-	-	-
20	2	3	6	6	-	6	4	1	-	-	-	-	-	-	-	-
21	3	6	6	6	-	4	1	-	-	-	-	-	-	-	-	-
22	4	4	4	4	-	2	1	1	1	1	1	-	-	-	-	-
23	1	1	1	2	-	5	5	5	4	3	1	1	1	1	1	1
24	-	-	1	5	-	6	6	4	3	1	1	1	1	1	1	-
25	-	1	5	6	-	6	5	3	1	1	1	1	1	1	-	-
26	1	5	6	6	-	5	3	1	1	1	1	1	1	-	-	-
27	4	5	5	5	-	4	2	2	2	2	2	1	-	-	-	-
28	3	3	3	4	-	4	4	4	4	4	3	-	-	-	-	-
29	1	1	2	4	-	6	6	6	6	5	-	-	-	-	-	-
30	1	2	4	6	-	6	6	6	5	-	-	-	-	-	-	-
31	2	4	6	6	-	6	6	5	-	-	-	-	-	-	-	-
32	4	6	6	6	-	6	5	-	-	-	-	-	-	-	-	
33	6	6	6	6	-	5	-	-	-	-	-	-	-	-		
34	5	5	5	5	-	1	1	1	1	-	-	-	-			
35	-	-	-	1	-	6	6	6	5	3	2	2				
36	-	-	1	6	-	6	6	5	3	2	2					
37	-	1	6	6	-	6	5	3	2	2						
38	1	6	6	6	-	5	3	2	2							
39	5	5	5	5	-	4	3	2								
40	3	3	3	4	-	5	4									
41	2	2	3	5	-	5										
42	2	2	4	5	-											

Figure 8.3: Figure : Combined results

We have already seen that the subjects varied paragraph length, and none of the paragraphs seem extraordinary long. This experiment provides no information on main sentences, but we might examine sentences at accessible locations to see if we come across something of interest. Our main focus remains to discover whether we can find similarities and differences between the subjects on a sentence level.

8.4.1 Inseparable sets of sentences

First, one of the most obvious observations will be discussed: some of the sentences aren't separated by any of the subjects. This would seem to imply that in each of these cases there is overwhelming, or at the very least compelling, evidence to link the two (or more) sentences together and that a paragraph boundary should not occur between them.

1)	Er was eens een vorst die de hele wereld onder zijn heerschappij wilde brengen.
2)	En niet door wijsheid, goedheid, eerlijkheid, rechtvaardigheid en vrede, maar door oorlog te voeren.

The reluctance of the subjects to separate these sentences can be attributed to the word "And" that starts the second line. "And" is usually not found at the start of a sentence, but as a conjunction joining several sentence parts. Another explanation might be that line 2 explains *how* the King wants to subjugate the world.

4)	Te vuur en te zwaard overwonnen zijn soldaten de buurlanden; er stroomden rivieren van bloed en tranen, en van vele trotse steden bleven slechts smeulende puinhopen over.
5)	De wrede huurlingen van de vorst slachtten zelfs hele kuddes vee af; honden en katten die zich niet snel genoeg uit de voeten maakten, werden door hen aan het zwaard geregen.

These two sentences seem to be part of the same description; namely how the Wicked King's army conquered neighboring countries. Moreover, line 5 seems to elaborate on how brutally the King's mercenaries achieved the victories mentioned in line 4. It is important to note that line 4 contains the only reference to these bloody battles, which makes it the only line that line 5 could refer to.

7)	Hij was uitsluitend geïnteresseerd in een zo groot mogelijke oorlogsbuit uit de paleizen, kerken en kloosters die hij overwonnen had.
8)	Zijn schatkamer stroomde over met goud, edelstenen, parels en juwelen: zelfs gouden deurknoppen liet hij loswrikken!
9)	Met al deze schatten bouwden zijn architecten schitterende paleizen, burchten en kastelen, zodat iedereen die ze zag van angst voor hem vervuld zou raken en zou buigen voor zijn almacht.

These three sentences seem to be connected by subject: the treasures and spoils of war. While line 7 describes the King's obsessions with treasure, it is line 8 which elaborates on how this obsession shows in his behavior and line 9 reveals the reason for this obsession.

18)	De priesters verzetten zich als enige tegen dit voorstel.
19)	'U heerst over de mensen in vele landen, machtige vorst, maar u bent niet de heerser over de hemel.
20)	God is nog altijd machtiger dan u.
21)	We kunnen uw bevel niet opvolgen: er kunnen geen standbeelden van u worden opgericht in onze godshuizen!'

Once again, the link in these sentences is the subject. Lines 18 through 21 describe how the priests react to the King's decree presented in sentence 17. Lines 19 through 21 are easily defined as inseparable, because they are part of one utterance, presumably voicing the collective opinion of all priests. The quotation marks are a clear indication and this could explain why these sentences have not been separated.

24)	Hij begon onmiddellijk met de voorbereidingen.
25)	Zijn architecten maakten bouwtekeningen voor een luchtschip, dat voortgetrokken zou worden door duizend adelaars: daarmee zou hij de hemel bestormen.
26)	Zijn wapenmeesters maakten kanonnen die zo groot waren dat ze het hemelse paleis in puin konden schieten.

Line 24 tells us that the King immediately starts preparing for battle, and the following two lines elaborate on what these preparations consist of. As with lines 4 and 5, the elaborations could not easily have referred to any other line in the text.

29)	De militaire kapel speelde.
30)	De maarschalken staken lange redevoeringen af voor de aanval werd geopend.
31)	De adelaars die het merkwaardige luchtschip zouden voorttrekken zaten nog in hun kooien.
32)	De soldaten stelden zich in een mijlenlange rij op om plaats te nemen in hun luchtschip.
33)	Ook de vorst zelf, gehuld in zijn vergulde harnas, steeg aan boord.

Lines 29 through 33 describe the setting for the last scene. It is interesting to note that, though some actions are described, these actions aren't sequential. Lines 29 through 33 describe several actions going on simultaneously.

35)	Plotseling kwam er een wolk muggen aanzwermen: alle muggen stortten zich op de vorst en zoemden oorverdovend.
36)	De machtige vorst trok zijn zwaard en sloeg om zich heen, maar hij sloeg alleen maar in de lucht.
37)	Een van de muggen wist zijn helm binnen te dringen en stak de meedogenloze vorst in zijn oor.
38)	De vorst begon om hulp te roepen, sprong uit zijn luchtschip en zette het op een lopen in de richting van het bos.

Line 35 introduces the mosquitoes and describes their attack on the King. The consequences of this attack are described in the following sentences.

In this small set of examples it is easy to find a single issue for the set of inseparable sentences to be about, but it is interesting to see that in almost every set we find that the first sentence provides the main content, and additional sentences seem to elaborate on that first sentence. We consider the sixth example (lines 29 through 33) an exception. Here each sentence describes an event, and these events occur simultaneously.

8.4.2 Separable pairs of sentences

After providing an overview of sentences that are never separated, it is fruitful to examine the results 'separated pairs': adjacent sentences that are always separated. A quick inspection shows that no such sets exist. It is still interesting to examine 'separable pairs': pairs of adjacent sentences which only one or two subjects placed in the same paragraph. Recalling our main question of inseparable sets, we may now wonder: do separable pairs address more than one issue?

There is only one pair of sentences that five of the six subjects separated:

34)	Hij stond op de brug, klaar om het eerste bevel te geven.
35)	Plotseling kwam er een wolk muggen aanzwermen: alle muggen stortten zich op de vorst en zoemden oorverdovend.

Line 34 is the last sentence that describes what the preparations had led to, while line 35 describes the start of how the King's plans start to fall apart. In this case, line 34 seems to be a conclusion of the previous scene. It would be logical to consider the 'preparations' and the 'falling apart' (or even 'the mosquito attack') are different issues.

The story contains only two pairs of sentences that four of the six subjects separated. Here follows the first such pair:

9)	Met al deze schatten bouwden zijn architecten schitterende paleizen, burchten en kastelen, zodat iedereen die ze zag van angst voor hem vervuld zou raken en zou buigen voor zijn almacht.
10)	Maar daar was de vorst nog niet tevreden mee.

Line 9 describes the results of the Wicked King's behavior and presents the setting of the story; it serves to show how wicked the Wicked King really is. The next sentence starts with "But" to indicate that a new argument is going to be put forward to show that the wickedness of the king was even greater. The marker "But" could be responsible for creating a new paragraph here, but the subjects who didn't interject a new paragraph between these lines have indicated a new paragraph soon after line 10 (once between line 10 and line 11, once between line 11 and line 12). This would seem to indicate that 'the King's treasures' and 'the King's wish to be worshipped' are separate issues.

Frequency					
sentence	first	last	sentence	first	last
1	6	0	22	2	4
2	0	1	23	4	1
3	1	2	24	1	0
4	2	0	25	0	0
5	0	2	26	0	1
6	2	2	27	1	2
7	2	0	28	2	2
8	0	0	29	2	0
9	0	4	30	0	0
10	4	1	31	0	0
11	1	1	32	0	0
12	1	1	33	0	1
13	1	3	34	1	5
14	3	1	35	5	0
15	1	1	36	0	0
16	1	1	37	0	0
17	1	3	38	0	2
18	3	0	39	2	1
19	0	0	40	1	1
20	0	0	41	1	1
21	0	2	42	1	6

Figure 8.4: Showing the number of subjects who selected a sentence as the first or last sentence of a paragraph

22)	De vorst bedreigde hen met de marteldood, maar de priesters lieten zich hierdoor niet afschrikken.
23)	Nu had de vorst nog maar één doel voor ogen: hij zou ook de hemel veroveren.

Line 22 concludes the conflict that occurred between the Wicked King and the priests. Even when threatened with torture and death, they won't renounce their faith. The next line (23) presents the conclusion that the king has drawn from this conflict: he must conquer Heaven. Once again, the subjects who didn't choose to begin a new paragraph between lines 22 and 23, start one soon after this line.

8.4.3 Main sentences

In chapter 7 the location of main sentences was discussed; mostly at the beginning of a paragraph, and sometimes at the end. It is hard to define what a main sentence is, but it is interesting to see how many times sentences in our story are selected as the first or last sentence of a paragraph. Figure 8.4 shows how many times a sentence was chosen to be the first or the last sentence of a paragraph.

It shall come as no surprise that the first and the last sentence of the story score high on being either a first or a last sentence of a paragraphs. But which other lines where often used to begin or end a paragraph? There are two sentences with a score of six, and three with a score of five. Let us examine them:

22)	De vorst bedreigde hen met de marteldood, maar de priesters lieten zich hierdoor niet afschrikken.

34)	Hij stond op de brug, klaar om het eerste bevel te geven.

10)	Maar daar was de vorst nog niet tevreden mee.

23)	Nu had de vorst nog maar n doel voor ogen: hij zou ook de hemel veroveren.

35)	Plotseling kwam er een wolk muggen aanzwermen: alle muggen stortten zich op de vorst en zoemden oorverdovend.

No study was done to determine which sentences in this story can be considered main sentences. Nevertheless each of the sentences in this list can be said to envelop surrounding sentences. Especially sentences 10, 23 and 35 describe major changes in the overall story-line. They seem related to the goal currently being pursued by the main character. Further study, in which test subjects are asked to identify main sentences in a text, could provide interesting information when combined with the paragraph information gathered here.

8.5 Conclusions

In this chapter we have tried to gain a better understanding on what motivates a paragraph structure. To this end we have performed a small experiment that has provided some insight in this matter.

The subjects each submitted a different paragraph structure showing that we can assume that there not one unique paragraph structure that can be considered correct. Despite differences in paragraph structure, sets of sentences exist which none of subjects separated. These sets were usually relate to a single subject, and often contain elaborations. Some pairs of adjacent sentences were often separated; we have seen that these refer to more than one event or issue.

It should be noted that a study of this size does not result in definite conclusions, but some of the results can be taken into account when implementing paragraph separation in the Virtual Storyteller. Specifically, we will focus on determining a paragraph structure with a single topic per paragraph. In the following chapter I will outline my method for automatically detecting discourse relations between fabula elements. Then, in chapter 10 I will use the discourse relations to construct a discourse structure which is well-formed with regard to paragraph separation. The lessons learned in this chapter will be used in each of these steps.

Chapter 9

Automatically adding rhetoric relations

Chapter 7 and chapter 8 have provided some insight in the rules that govern paragraph separation, which we will need in the next chapter. It is there that we construct a discourse structure tree based on the fabula. But before we can create this tree, we need to determine which rhetorical relations are possible between the facts contained in the fabula.

9.1 Adding rhetorical relations to the fabula

Below I've relisted the rhetorical relations listed in chapter 6. For each of them, I will determine if we can assign them to sets of fabula elements at this time. Note that a lot of relations mentioned in the original list, might not be readily available, or their application does not help toward combining fabula elements.

For those which do seem useful at this time, I will provide their descriptions and add what the nucleus and satellite should contain in order for the relation to apply. This allows us to implement the automated recognition of rhetorical relations in the fabula structure; the first step to automatically creating a well-formed discourse structure tree. But first I will explain the notation used in the definitions below.

9.1.1 Fabula elements

Chapter 5 described the fabula and provided descriptions of the different elements that can be produced by the virtual simulation. Once again, in bold I've shown the letters to represent them in the definition to follow.

- A goal (**G**) is a state or action a character is basing its actions on.
- An outcome (**O**) is the result of a goal, which can be positive, negative or neutral.
- An internal element (**IE**) is a single emotion, cognition or belief. It denotes a single element of a character agent's internal state.
- An action (**A**) represents an action performed by a character. Remember from chapter 4 that actions are considered operators on the virtual world. They have

preconditions that must be met, after which inter-effects are applied. If the preconditions still hold after the 'interruptable period', the effects of the action are applied to the virtual world.

- An event (**E**) represents an action performed by 'the world'. To be more precise, it refers to changes to the virtual world which are not the intended effect of character actions.
- A perception (**P**) is the element that enables changes in the virtual world to influence characters. In order to react to a change, they must first perceive it.

Note that I'll also refer to the objects (**Obj**) present in the storyworld. This includes characters (**Char**), locations (**Loc**) and other objects (**OtherObj**). Of course, these objects are mostly linked to fabula elements. For instance, a 'throw' action involves a character and an object being thrown.

9.1.2 Elaboration in the fabula

Remember that Elaboration consists of ElabObject, ElabPart and ElabGenerality. Each of these will be discussed in turn.

ElabObject

For each of the relations that elaborate on an object, we need to check the nucleus for any objects that are present, see which attributes and functions these objects can have, and see if the satellite contains one of the values of the attributes or functions in order for the relation to hold. The problem is that at the moment, the attributes an object can have and actions an object can perform are defined in the storyworld core, which is currently a standard object ontology. This ontology still needs to be altered to fit the Virtual Storyteller properly, and this work will most likely be done after the character agents are done.

For now, that means we have the option of identifying different rhetorical relations for each of these attributes and functions. For instance, ElabAttribute could consist of ElabAttrHeight, ElabAttrWidth and ElabAttrColor. Obviously, this list is far from complete, but seeing that the list of attributes and functions isn't definite, I would rather not provide such definitions for the rhetorical relations.

Name	ElabObject
Location	Elaboration - ElabOject
Description	The satellite provides additional information about an object present (or inferentially accessible [13]) in the nucleus. This additional information concerns either an attribute of the object or the object's function.
Consists of	ObjectAttribute ObjectFunction

Name	ObjectAttribute
Location	Elaboration - ElabOject - ObjectAttribute
Description	The satellite presents additional detail about an object which is presented in the nucleus (or inferentially accessible). The additional information about the object must provide an attribute of this object.
Nucleus Satellite	contains **Obj1** contains attribute of **Obj1**

Name	ObjectFunction
Location	Elaboration - ElabOject - ObjectFunction
Description	The satellite presents additional detail about an object which is presented in the nucleus (or inferentially accessible). The additional information about the object must provide an function this object can perform.
Nucleus Satellite	contains **Obj1** contains function of **Obj1**

ElabPart

For the ElabPart relation to hold true, the satellite must contain information either about a member of set, a step in a process or about a part of a whole. Each of these are not easy to detect in the fabula, because these concepts can be interpreted in different ways. Because we, as of yet, have no automatically generated fabula to work with, I feel it would be unproductive to specify each of these relations further. For now, the more general descriptions will have to suffice.

Name	ElabPart
Location	Elaboration - ElabPart
Description	The satellite provides additional information about a set, process or object that is present in the nucleus. It provides additional information concerning either a member of the set, a step in the process or a part of the object.
Consists of	Set-Member Process-Step Whole-Part

Name	Set-Member
Location	Elaboration - ElabPart - Set-Member
Description	The nucleus presents a set of objects or concepts and the satellite presents one of the members of this set.
Nucleus Satellite	contains **Obj1**, which is a set contains **Obj2**, which is a member of this set

Name	Process-Step
Location	Elaboration - ElabPart - Process-Step
Description	The nucleus refers to a process and the satellite elaborates on a single step of that process.
Nucleus Satellite	contains **A1**, which is a process contains **A2**, which is a part of this process

Name	Whole-Part
Location	Elaboration - ElabPart - Whole-Part
Description	The nucleus refers to a object, of which a part is presented in the satellite.
Nucleus Satellite	contains **Obj1** contains **Obj2**, which is a part of **Obj1**

ElabGenerality

Though it is clear that differences in abstraction or generality occur in text, we must note that the satellite doesn't provide information that wasn't present in the nucleus, if we examine it on the level of the fabula. Let us consider the following example:

> "The children laugh at the animals, *particularly the monkeys*."

On a fabula level both the nucleus can be represented as the action LaughingAt(Children, Animals). The fact presented in the satellite is the same, with monkeys instead of animals. But LaughingAt(Children, Monkeys) isn't a separate element in the fabula, but is contained within LaughingAt(Children, Animals). This relation cannot be found to exist between two different fabula elements, becasuse it relates an element to itself. Although this does hint a method for adding the use of these relations to the fabula, it doesn't help at creating a discourse tree.

For this reason, I propose to set aside the relations pertaining to generality. Further study should determine where these relations can be helpful and how they should be incorporated in discourse trees.

Identification

Like ElabGenerality, with Identification the satellite provides additional information that is already contained within the nucleus. Objects are identified uniquely in the fabula and need no further elaboration, which leads me to conclude that Identification is another relation that is not readily recognized in the fabula. Instead, the writer has some other motive for elaborating on an object's identity, or had good reason to obscure it to begin with.

Restatement

Restatements, either Summary or Conclusion, are not literally present in the fabula. Remember that the fabula contains only the facts that occurred in the virtual world, and these facts cannot occur twice. I would propose that in a later stage, perhaps when the

discourse structure tree has been fully constructed, methods can be developed to duplicate certain parts of the tree. These duplicated parts can then be connected using restatement (and generality) relations. For now we can conclude that the fabula alone does not provide the information to automatically add restatement relations.

9.1.3 Circumstance in the fabula

The different circumstances that Hovy identified are represented in the fabula in different ways. The location of an object is (at the moment) one of its attributes, which can be looked up in the storyworld core. Time can be determined by looking at the attributes of elements, except for objects. We seem to have no way to determine the Manner in which an action is performed. This too will probably become an attribute of actions in the future. Instrument is defined within the action and ParrallelEvent simply combines two actions or events which overlap.

It would seem the ParrallelEvent is the only circumstance relation that actually combines two fabula facts currently present in the fabula. Our main interest is in combining the available fabula elements, and not in adding new ones, so we'll leave all but ParrallelEvent for now.

Name	ParallelEvent
Location	Circumstance - ParallelEvent
Description	Both nucleus and satellite occur at exactly the same time and are either events or actions.
Nucleus Satellite	contains **A1** or **E1** contains **A2** or **E2**, where the starttimes are equal.

9.1.4 Sequence in the fabula

One of the reasons to set our sights on creating fairy-tales, is the perceived simplicity. Most fairy-tales contain only a few characters and easy to follow plots. And, once again I cautiously say mostly, are told in strict chronological order. This is why the SeqTemporal relation is very important; it determines the order in which the fabula facts are presented.

Unfortunately, SeqSpatial and SeqOrdinal are a lot harder to define than temporal sequences. Spatial relations could be determined for large geographical areas by looking up their relation to each other in the storyworld core, but at the moment all objects in a single location have no specified location in relation to each other - they are all just floating around in there. SeqOrdinal is even harder to determine, because it refers to a (well-known) order of objects or concepts. Because this kind of information is poorly supported in the current fabula design, I will limit this research to using only SeqTemporal.

Name	SeqTemporal
Location	Sequence - SeqTemporal
Description	The satellite occurs (directly) after the nucleus.
Nucleus Satellite	contains **A1** contains **A2** or **E2** with starttime within 'threshold' of starttime or endtime of **A1**
Nucleus Satellite	contains **E1** contains **A2** or **E2** with starttime within 'threshold' of starttime or endtime of **E1**

9.1.5 Cause/Result in the fabula

The four different kinds of causal relations are defined by two aspects: *volitionality* and *whether the cause or the effect is more important for the story.* But how do these two aspects relate to the fabula?

Volitionality

As discussed in chapter 6, it isn't easy to determine whether an action is volitional or not. In the example we discussed ("Because the arrow pierced his heart, the prince died") it is still unclear how the fabula representing these actions would look. Is the prince dying even a separate action? Unfortunately, these decisions have not been made as of yet. On the other hand, the fabula does provide a clear distinction in another way.

Within the fabula we distinguish between actions and events, where events are similar to actions except that they are used by the Plot Agent to steer the story. The motivation for events lies outside the fabula, and thus appear to be non-volitional.

For the remainder of this thesis, I will assume that volitional actions are synonymous with Actions, and non-volitional actions are synonymous with Events.

Cause or effect

Next, we are left with the task of determining which is more important for the story: the cause or the effect. Of course, when combining two clauses it is impossible to determine the importance without looking at the rest of the story. This argues that both possibilities should be available when we construct the discourse structure tree that governs the entire story.

For each pair of clauses that we combine with a causal relation (say **c1** and **c2**), it then becomes necessary to add both VolCause(c1,c2) and VolResult(c2,c1) to the fabula (or NonVolCause(c1,c2) and NonVolResult(c2,c1) if the clause contained an event).

Name	C/RVol
Location	Cause/Result - C/RVol
Description	A volitional action (or else a situation that could have arisen from a volitional action [13]) is present in either clause (nucleus or satellite), the other clause contains a motivation or reason
Consists of	VolCause VolResult

Name	VolCause
Location	Cause/Result - C/VolCause - VolCause
Description	A volitional action (or else a situation that could have arisen from a volitional action [13]) is present in the nucleus, the satellite contains a motivation or reason.
Nucleus Satellite	contains **A1** contains the motivation for **A1**

Name	VolResult
Location	Cause/Result - C/VolCause - VolResult
Description	A volitional action (or else a situation that could have arisen from a volitional action [13]) is present in the satellite, and the nucleus contains a motivation or reason.
Nucleus Satellite	contains the motivation for **A1** contains **A1**

Name	C/RNonVol
Location	Cause/Result - C/RNonVol
Description	An event is present in either clause (nucleus or satellite), and the other clause contains a motivation or reason.
Consists of	NonVolCause NonVolResult

Name	NonVolCause
Location	Cause/Result - C/RNonVol - NonVolCause
Description	An event is present in the nucleus, and a motivation or reason can be found in the satellite.
Nucleus Satellite	contains **E1** contains the motivation for **E1**

Name	NonVolResult
Location	Cause/Result - C/RNonVol - NonVolResult
Description	An event is present in the satellite, and a motivation or reason can be found in the nucleus.
Nucleus Satellite	contains the motivation for **E1** contains **E1**

Name	Purpose
Location	Cause/Result - Purpose
Description	The satellite motivates the action presented in the nucleus, but does so referring to a goal being pursued.
Nucleus Satellite	contains **A1** contains **G1**, which motivates **A1**

9.1.6 GeneralCondition in the fabula

The descriptions for GeneralCondition refer to general rules or unrealized situations. It is hard to determine how general rules are represented in the fabula, but it is clear that unrealized situations refer to goals. Conditions and exceptions can be related to subgoals to either meet or avoid, so the GeneralCondition relation always combines two goals.

Name	Condition
Location	GeneralCondition - Condition
Description	The nucleus contains a general rule or an unrealized situation, and the satellite describes a condition that has to be met in order for the rule to apply or the situation to be valid.
Nucleus Satellite	contains **G1** contains **G2**, which is a goal motivated by **G1**

Name	Exception
Location	GeneralCondition - Exception
Description	The nucleus contains a general rule or an unrealized situation, and the satellite describes an exception that prevents that situation from becoming a reality or prevents that rule from applying.
Nucleus Satellite	contains **G1** contains **G2**, which is to avoid **G3**, which is motivated by **G1**

9.1.7 Comparative in the fabula

Hovy defines five kinds of comparative relations: Equative, Contrast, Comparison, Otherwise and Analogy. With the Equative relation, two objects are related either on a shared attribute, function, or performing the same action. Attributes and functions can be found in the storyworld core, and actions are of course present as a fabula element. Contrast is

defined in a similar way. At this point, I feel that Comparison might have become unnecessary, but we might want to compare objects for cases in which no attribute, function or action is present.

Analogies are used to illustrate the function of an object by comparing it to another object. While it would be easy enough to compare each pair of clauses to see whether they contain objects with similar functions, the only objects available within a fabula are actual storyworld objects. When comparing two objects which are present, a comparative relation is more applicable.

In order to identify analogies, the fabula would have to include a collection of objects to compare the storyworld objects to. It would be interesting to see what such a collection would look like, or even how a text would look if only analogies to just one object are possible. Unfortunately, these musings are beyond the scope of this research.

Lastly, the Otherwise relation is a complex relation between two goals that seemed tied into goal determination. Seeing that the character agent is still being developed, I reluctantly choose to disregard Otherwise.

Name	Equative
Location	Comparative - Equative
Description	The nucleus contains an object, about which an attribute or function is presented, or which performs an action. The satellite contains a different object with that same attribute or function, or performing that same action.
Nucleus Satellite	contains **Obj1** and one of its attributes contains **Obj2** which has that same attribute
Nucleus Satellite	contains **Obj1** and one of its functions contains **Obj2** which can perform that same function
Nucleus Satellite	contains **A1** performed by **Obj1** contains **A1** performed by **Obj2**

Name	Contrast
Location	Comparative - Contrast
Description	The nucleus contains an object, about which an attribute or function is presented, or which performs an action. The satellite contains a different object without that attribute or function, or not performing that same action.
Nucleus Satellite	contains **Obj1** and one of its attributes contains **Obj2** which does not have that attribute
Nucleus Satellite	contains **Obj1** and one of its functions contains **Obj2** which can not perform that function
Nucleus Satellite	contains **A1** performed by **Obj1** contains **Obj2** not performing **A1**

Name	Comparison
Location	Comparative - Comparison
Description	The nucleus is compared to the satellite, in order to note similarities and/or differences. Both the differences and the similarities are important for the story
Nucleus Satellite	contains **Obj1** contains **Obj2**

9.2 Revisiting previous work

Remember from chapter 4 that Slabbers proposed a list of rhetorical relations which she deemed sufficient to construct discourse structures with. It is easy to see that I have incorporated *Cause*, *Purpose*, *Contrast*, *Temporal* (as SeqTemporal), *Condition* and both kinds of *Elaboration* (*Elaboration2* as ElabObject). But what about Slabbers' *Additive*?

9.2.1 Additive

An additive relation simply enumerates the two related clauses. This relation was introduced by Hielkema to enable ellipsis, but as Slabbers mentions is not part an 'official' rhetorical relation. I feel that this relation consists of two parts: Equative and Conjunction.

Equative is one of the types of relations that make up Comparative relations, according to Hovy. Though this relation is not sufficient to enable *all* possibilities for ellipsis, it should enable them where they are semantically sound. Consider the following example:

> "The prince enters the forest. The princess enters the forest."

Seeing that these two sentences are syntactically similar, we could apply ellipsis to create "The prince and princess enter the forest.". But is it always semantically correct to do this? What if the prince and princess in question both enter the forest at completely different ends? And never meet? The fabula that lies at the basis of these sentences would contain that kind of information, and if an equative relation were present you could safely apply ellipsis.

Hovy also defines Conjunction, as a LogicalRelation, which in turn is categorized as a textual relation. This relation is the one already hinted at in the last paragraph, and identifies syntactical opportunities for ellipsis.

In this thesis I've chosen to focus on semantical relations, but note that by supporting only equative relations and not conjunctions, I will miss some of occasions where ellipsis could be performed. On the other hand, this approach also prevents ellipsis from occurring where it might not be semantically applicable.

9.2.2 Conclusion

Chapter 6 provided us with a large framework of rhetorical relations that can be part of a discourse structure tree. As noted at the start of this chapter, we won't be able to use

some of them at this moment. In the end, we ended up with a list of relations barely longer than the list Slabbers used. So what have we gained?

Hovy compiled the list of relations by combining different sources from different fields. His main objective was to present a list that can make some claim to completeness. By adopting this structure, we can now hope to have a framework that will allow for expansion of the Virtual Storyteller language generating abilities. For each of the rhetorical relations left aside in this chapter, we can conceive of specific research telling us how to add this relation to the fabula automatically. And once the list has been implemented more fully, it should provide a sufficient basis to perform other linguistic tricks which improve the quality of generated text.

9.3 Next chapter

In the next chapter, we are going to use the rhetorical relations added to the fabula and construct a discourse structure tree. We will return to the lessons learned from analyzing both the theory of paragraphs and manmade paragraph separation.

Chapter 10

Creating a discourse tree

In this chapter we examine methods of using the rhetorical relations identified between fabula elements to create a discourse tree. This tree should be well-formed with respect to paragraph separation, so a method for identifying paragraphs in such a tree will also be described.

In order to provide a proof of concept for the methods proposed in this thesis, we will apply them to a handmade fabula in my evaluation (chapter 11).

But first I will make some observations and assumptions to help with constructing the discourse tree.

10.1 Important elements

I propose to use the number of times it could be used as the nucleus in a relation, as a measure for the importance of that fabula element for the narrative. In turn, this measure can be used to determine which element should be the nucleus if this hasn't been determined for a particular relation. In other words, if it is unclear which of two elements is the nucleus in a particular combination, we look to see which of the two elements is more important to the story by looking at other combinations those elements are part of.

10.2 Unconnected elements

By connecting elements one at a time, choosing the 'most important' one without looking ahead, we run the risk of leaving some elements unconnected from the discourse structure. Note that it is also possible that some elements were not connected in the first place! I feel that losing some of the elements is probably not going to be detrimental to the construction of the tree, but it is an issue that needs to be solved in the future.

10.3 Selecting the top node

I've proposed to determine the importance of a fabula element by counting the number of times it can function as the nucleus in a rhetorical relation. It would make sense to

use this method to find a starting point for constructing our tree. By placing the most important node at the top, we can be certain that most of the story will be able to fit underneath the root.

10.4 Creating the tree

Once we have selected the top element, the question becomes how to continue. Clearly the more important nodes should be connected first, so I would propose selecting from the fabula elements which could be a satellite to the top element, and connect the most important ones. Because we want to create a tree, and not a single chain, I propose to add several of the highest scoring fabula elements. At what level this threshold is set, depends heavily on the number of relations that have been implemented. For now, I propose to to add all the elements that fall within 10% of the highest scoring one. Once an element has been added in this way, we repeat the process in a breadth-first way, to evenly fill out the tree.

By recursively adding elements to the tree, each branch will come to a natural conclusion when it adds an element for which there are no more possible satellites.

10.5 Find paragraphs

Now that we've created this tree, it's time to determine the paragraph structure. Remember from chapter 7 that a paragraph has to be coherent. So every paragraph should be a discourse tree on its own. But how do we divide a discourse tree into several smaller ones?

The algorithm I've designed starts at the leaves of the discourse tree, and travels up until it reaches an element where it combines with another branch. At this point, we need to determine whether these two branches can combine into one paragraph, or that a boundary has been reached. In our example below, I will try out several methods to set this boundary and discuss its effects.

10.6 Conclusions

This chapter provides a simple algorithm to create a well-formed discourse tree based on a fabula, to which rhetorical relations have been added automatically. I decided to stick to a basic idea, simply because at the moment there is no easy way of testing and trying different options. The Virtual Storyteller is, as of yet, unable to create fabula structures. But more importantly, because the fabula structure used is designed solely for the Virtual Storyteller, and not much other work has been done that really compares, there is hardly any theory on which to base such an algorithm. And when both theory and ability to test and experiment are lacking, it is hard to fine-tune.

In the evaluation, I will take on the task of creating a fabula by hand, and executing the algorithm upon it. I hope this provides a proof of concept for the choices made in this chapter.

Chapter 11

Evaluation

> 'This idea,' Mr Bohlen's lower lip was saying, 'is very ingenious - I might almost say brilliant - and it only goes to confirm my opinion of your abilities, Knipe. But don't take it too seriously. After all, my boy, what possible use can it be to us? Who on earth wants a machine for writing stories? And where's the money in it, anyway? Just tell me that.'
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 596

11.1 An example: The Wicked King

At the moment, the Virtual Storyteller is unable to create fabula structures, so we are forced to create one by hand if we want to test the methods proposed in this thesis. I've chosen to create a fabula for the story 'The Wicked King', which was also used in our experiment about human paragraph separation (see chapter 8). On the one hand, it could be considered unwise to use the same story twice. The lessons learned when analyzing this story, should be tested on another. On the other hand, it is prudent to check if the lessons learned are implemented in such a way that they lead to the desired results. Testing these steps on a different story might not lead to the desired results, while the method would not be at fault. Eventually, we need to repeat all the steps taken in this research for a large set of stories, but for now 'The Wicked King' will have to suffice.

Note that I have chosen to simplify the tale in several places, in order to create a fabula that is not necessarily complete, but sufficient to test my method for constructing a discourse tree.

11.1.1 Creating a fabula

First of all, let us identify the objects present in this fairy-tale, differentiating between characters, locations and other objects. I will provide a short summary of the story, noting the different characters.

Summary

The Wicked King (**K**) wants to be all-powerful. His soldiers (**S**) conquer the neighbouring lands, and kill cattle and pets. The king's architects (**A**) design palaces to store the king's spoils of war. The Wicked King humiliates the kings of other countries (**OK**) in front of his own subjects (**Sub**). The king wants to put statues on all city squares and in each church. The priests (**P**) resist him and tell him that he'll never be as powerful as God (**G**). The king now plans to conquer heaven. His architects design an airship and his masters at arms (**MaA**) design powerful cannons. On the day the attack would commence, the generals (**Gen**) address the troops. A swarm of mosquitoes attack the King, and the King is forced to flee into the forest and is never heard from again.

Here is short overview of the characters that occur in this tale. Note that some of the characters are actually groups of characters.

K	the King
S	Soldiers
A	Architects
OK	Other kings
Sub	Subjects
P	Priests
G	God
MaA	Masters at arms
Gen	Generals

Locations

Next, an overview of the different locations the story employs.

Within the storyworld, there are two main locations: Heaven and Earth. Earth is divided into several countries, but the notable distinction we need are: the King's country and other countries. These countries contain cities, which in turn contain squares and churches. The King's country also contains a clearing where the army assembles, next to a forest.

Other objects

Other objects of importance are the cattle and pets in neighbouring countries, the spoils of war, statues, the airship, the cannons and the mosquitoes. Note that it is also possible to view the mosquitoes as characters (or perhaps to view God as an object).

Listing fabula elements

Below, I've listed the fabula elements, together with a short description. The descriptions in parentheses can not to be found in the tale, but are implied.

Nr.	Fabula Element	Description
001.	$IE_{K(1)}$	(De vorst is ongelukkig.)
002.	$G_{K(1)}$	(De vorst wil gelukkig zijn.)
003.	$G_{K(2)}$	(De vorst wil aanbeden worden.)

004. $G_{K(3)}$ De vorst wil heersen over zijn buurlanden.
005. $A_{K(1)}$ (De vorst beveelt zijn soldaten de buurlanden te overwinnen.)
006. $P_{S(1)}$ (De soldaten horen het bevel.)
007. $IE_{S(1)}$ (De soldaten geloven dat de vorst de buurlanden wil overwinnen.)
008. $G_{S(1)}$ (De soldaten willen de buurlanden overwinnen.)
009. $A_{S(1)}$ Zijn soldaten overwinnen de buurlanden.
010. $A_{S(2)}$ Zijn soldaten vernietigen vele steden.
011. $A_{S(3)}$ De soldaten slachten het vee.
012. $A_{S(4)}$ De soldaten slachten de huisdieren.
013. $A_{S(5)}$ De soldaten verzamelen oorlogsbuit.
014. $P_{K(1)}$ (De soldaten overwinnen de buurlanden.)
015. $IE_{K(2)}$ (De vorst gelooft dat de soldaten de buurlanden overwinnen.)
016. $O^{+}_{K(3)}$ (De vorst heerst over zijn buurlanden.)
017. $G_{K(4)}$ (De vorst wil paleizen bouwen.)
018. $G_{K(5)}$ (De vorst wil middelen om paleizen mee te bouwen.)
019. $G_{K(6)}$ (De vorst wil oorlogsbuit.)
020. $P_{K(2)}$ (De soldaten verzamelen oorlogsbuit.)
021. $IE_{K(3)}$ (De vorst gelooft dat de soldaten oorlogsbuit verzamelen.)
022. $O^{+}_{K(6)}$ (De vorst heeft oorlogsbuit.)
023. $IE_{K(4)}$ (De vorst gelooft dat hij middelen heeft.)
024. $O^{+}_{K(5)}$ (De vorst heeft middelen.)
025. $A_{K(2)}$ (De vorst beveelt zijn architecten paleizen te bouwen.)
026. $P_{A(1)}$ (De architecten horen het bevel.)
027. IE_{A} (De architecten geloven dat de vorst wil dat zij paleizen bouwen.)
028. G_{A} (De architecten willen paleizen bouwen.)
029. $A_{A(1)}$ De architecten bouwen paleizen met de oorlogsbuit.
030. $P_{K(3)}$ (De vorst ziet de architecten paleizen bouwen.)
031. $IE_{K(5)}$ (De vorst gelooft dat de architecten paleizen bouwen.)
032. $O^{+}_{K(4)}$ (De vorst heeft paleizen.)
033. $G_{K(7)}$ De vorst wil heersen over de hele wereld.
034. $A_{K(3)}$ (De vorst beveelt zijn soldaten de hele wereld te overwinnen.)
035. $P_{S(2)}$ (De soldaten horen het bevel.)
036. $IE_{S(2)}$ (De soldaten geloven dat de vorst de hele wereld wil overwinnen.)
037. $G_{S(2)}$ (De soldaten willen de hele wereld overwinnen.)
038. A_{S6} De soldaten veroveren alle omringende landen.
039. $P_{K(4)}$ (De soldaten overwinnen de wereld.)
040. $IE_{K(6)}$ (De vorst gelooft dat de soldaten de wereld overwinnen.)
041. $O^{+}_{K(7)}$ (De vorst heerst over de wereld.)
042. $G_{K(8)}$ (De vorst wil de andere koningen vernederen.)
043. $A_{K(4)}$ De vorst vernedert de koningen van de omringende landen.
044. $G_{K(9)}$ De vorst wil op ieder stadsplein een standbeeld.
045. $G_{K(10)}$ De vorst wil in alle kerken een standbeeld.
046. $A_{K(4)}$ (De vorst beveelt de priesters.)
047. $P_{P(1)}$ (De priesters horen het bevel.)
048. $IE_{P(1)}$ (De priesters geloven dat de koning hen beveelt.)

049. $A_{P(1)}$ De priesters verzetten zich.
050. $A_{P(2)}$ De priesters vertellen de vorst dat God machtiger is dan hij.
051. $P_{K(5)}$ (De vorst hoort de priesters.)
052. $IE_{K(7)}$ (De vorst gelooft dat de priesters denken dat God machtiger is.)
053. $G_{K(11)}$ De vorst wil de hemel veroveren.
054. $G_{K(12)}$ (De vorst wil een luchtschip.)
055. $A_{K(5)}$ (De vorst beveelt zijn architecten een luchtschip te maken.)
056. $P_{A(2)}$ (De architecten horen het bevel.)
057. $IE_{A(2)}$ (De architecten geloven het bevel.)
058. $G_{A(2)}$ (De architecten willen een luchtschip bouwen.)
059. $A_{A(2)}$ Zijn architecten maken een luchtschip.
060. $P_{K(6)}$ (De vorst ziet het luchtschip.)
061. $IE_{K(8)}$ (De vorst gelooft dat hij een luchtschip heeft.)
062. $O^{+}_{K(12)}$ (De vorst heeft een luchtschip.)
063. $G_{K(13)}$ (De vorst wil kanonnen.)
064. $A_{K(6)}$ (De vorst beveelt zijn wapenmeesters kanonnen te maken.)
065. $P_{MaA(1)}$ (De wapenmeesters horen het bevel.)
066. $IE_{MaA(1)}$ (De wapenmeesters geloven het bevel.)
067. $G_{MaA(1)}$ (De wapenmeesters willen kanonnen bouwen.)
068. $A_{MaA(1)}$ Zijn wapenmeesters maken kanonnen.
069. $P_{K(7)}$ (De vorst ziet de kanonnnen.)
070. $IE_{K(9)}$ (De vorst gelooft dat hij kanonnen heeft.)
071. $O^{+}_{K(13)}$ (De vorst heeft kanonnen.)
072. $G_{K(14)}$ (De vorst wil dat het leger zich voorbereidt.)
073. $A_{K(7)}$ (De vorst beveelt zijn leger zich voor te bereiden.)
074. $P_{Gen(3)}$ (De maarschalken horen het bevel.)
075. $IE_{Gen(3)}$ (De maarschalken geloven het bevel.)
076. $G_{Gen(3)}$ (De maarschalken wil de soldaten toespreken.)
077. $A_{Gen(7)}$ De maarschalken spreken de soldaten toe.
078. $P_{S(3)}$ (De soldaten horen het bevel.)
079. $IE_{S(3)}$ (De soldaten geloven het bevel.)
080. $G_{S(3)}$ (De soldaten wil zich opstellen.)
081. $A_{S(7)}$ De soldaten stellen zich op.
082. $E(1)$ Een wolk muggen komt aanzwermen.
083. $E(2)$ De muggen vallen de vorst aan.
084. $E(3)$ De muggen steken de vorst in zijn oor.
085. $P_{K(8)}$ (De vorst merkt de aanval.)
086. $IE_{K(10)}$ (De vorst heeft pijn.)
087. $G_{K(15)}$ (De vorst wil geen pijn meer hebben.)
088. $G_{K(16)}$ (De vorst wil de muggen doden.)
089. $A_{K(8)}$ De vorst valt de muggen aan.
090. $P_{K(9)}$ De vorst slaat alleen maar lucht.
091. $IE_{K(11)}$ (De vorst gelooft dat hij alleen lucht raakt.)
092. $O^{-}_{K(16)}$ (De vorst kan de muggen niet doden.)
093. $G_{K(17)}$ (De vorst wil weg bij de muggen.)
094. $A_{K(8)}$ De vorst vlucht het bos in.

095. $P_{S(4)}$ De soldaten zien de vorst vluchten.
096. $IE_{S(4)}$ (De soldaten zijn blij.)
097. $A_{S(8)}$ De soldaten lachen.
098. $P_{K(10)}$ (De vorst hoort de soldaten lachen.)
099. $IE_{K(12)}$ (De vorst gelooft dat de soldaten lachen.)
100. $O^{-}_{K(2)}$ (De soldaten aanbidden de vorst niet.)
101. $IE_{K(13)}$ (De vorst gelooft dat de soldaten hem niet aanbidden.)
102. $O^{-}_{K(1)}$ (De vorst is niet gelukkig.)

11.1.2 Adding rhetorical relations

Now that we've created the fabula structure, it is time to add the rhetorical relations as defined in chapter 9.

ObjectAttribute

The only real attribute that is being disclosed, can be found in element 001: " The king is unhappy". We can now add the rhetorical relation ObjectAttribute(element, 001) for each element that contains the King. The first four are shown below. See appendix E.1 for the complete list.

1. ObjectAttribute(002, 001)
2. ObjectAttribute(003, 001)
3. ObjectAttribute(004, 001)
4. ObjectAttribute(005, 001)
5. ...

ObjectFunction

Except for the actions that the characters undertake, the tale contains no information about functions they can perform.

Set-Member

Several objects could be considered set, namely the soldiers, the architects, the other kings, the subjects, the priests, the masters at arms, the generals and the mosquitoes. But for each of these sets, the simplified version of the tale represented by our fabula, does not explicitly refer to objects which are part of that set. Instead, each time the set as a whole is said to take an action or have a perception. Because of this, we can conclude that there are no Set-Member relations present.

Process-Step

The soldiers conquer the neighbouring countries. Subsequent actions are all motivated by the same goal and can therefore be considered steps in the process of conquering. We add:

1. Process-Step(009,010)
2. Process-Step(009,011)
3. Process-Step(009,012)
4. Process-Step(009,013)

Whole-Part

Of none of the objects present in the tale, information of parts of them is provided.

ParallelEvent

We can add several ParallelEvent relations. For these, we will have to determine the element which is more important to the story later on. Shown below are the first four; see appendix E.3 for the complete list.

1. ParallelEvent(009,010)
2. ParallelEvent(009,011)
3. ParallelEvent(009,012)
4. ParallelEvent(009,013)
5. ...

SeqTemporal

SeqTemporal connects fabula elements that contain actions or events. Only fabula elements 82, 83 and 84 contain events, but the list of actions is a lot longer. In chapter 9 I mentioned connecting only those combinations in which the satellite has a start time within 'threshold' of the start time or end time of the action or event presented in the nucleus. The problem with determining such a threshold is that these might exclude relations that are needed on a higher level. Therefore, I'll add every possible SeqTemporal relation to the fabula, disregarding the threshold. Once again, the first four are shown here as an example and the complete list can be found in appendix E.4.

1. SeqTemporal(005,009)
2. SeqTemporal(005,010)
3. SeqTemporal(005,011)
4. SeqTemporal(005,012)
5. ...

C/RVol

Once again we must examine each fabula element containing an action, and add VolCause and VolResult to the fabula. The first four are presented here, the complete list can be found in appendix E.5.

1. VolResult(004,005)
2. VolResult(008,009)
3. ...
4. VolCause(005,004)
5. VolCause(009,008)
6. ...

C/RNonVol

For each of the three events, I will add both NonVolCause and NonVolResult. Unfortunately, the motivation for the event listed in fabula element 082, is not to be found in the fabula.

1. NonVolCause(083,082)
2. NonVolCause(084,083)
3. NonVolResult(082,083)
4. NonVolResult(083,084)

Purpose

In this story, each action is motivated by a goal. Therefore the Purpose relations are identical to the VolCause relations. For completeness, the list has been added to appendix E.6.

1. Purpose(005,004)
2. Purpose(009,008)
3. ...

Condition

Condition adds a list of goals that are motivated by goals. The first four are shown as an example, the complete list is added to appendix E.7.

1. Condition(003,002)
2. Condition(004,003)
3. Condition(017,003)
4. Condition(018,017)
5. ...

Exception

No goal contains a clear opposite of another, so no Exception relations can be added.

Equative

The only attribute mentioned is "The king is unhappy." and this attribute isn't mentioned for anybody else. As mentioned at ObjectFunction no functions are presented either. This leaves different objects performing the same action. Both the architects and the masters at arms build, albeit different things (029, 058 and 068). Equative(029,068) Equative(058,068) Equative(068,029) Equative(068,058)

Contrast

Because little is known about the set of available attributes, it isn't useful to add that kind of Contrast relations. This leaves the Contrast relations where the nucleus contains an object performing an action, and the satellite containing a different object not performing that action (see appendix E.9 for the complete list).

1. Contrast(005,009)
2. Contrast(005,010)
3. Contrast(005,011)
4. Contrast(005,012)
5. ...

Comparison

Adding a simple comparison to each combination might prove very usefull in an automated version to connect 'stray' fabula elements, but when performing this algorithm by hand it just adds a lot of unnecessary relations. Mostly out of efficiency, I will leave Comparison for now.

11.1.3 Determining the importance of elements

The results of tallying all the times a fabula element was assigned as the nucleus of a rhetorical relation, it would seem that fabula element number 5 is our top element. Unfortunately this is not very likely, but displays a mistake in the algorithm. I will expand on this in the next section.

11.2 Conclusions

I attempted to provide a simple method to mathematically determine the importance of the different fabula elements, in order to have some way of prioritizing them. I choose to link their importance to the number of relations in which they occur as the nucleus, which seems reasonable. After all, the nucleus of rhetorical relations is the main part of the combination and should contain the more important element.

The problem is that I've defined importance in terms of relations, and then try to specify the nucleus in these relations by using the importance. This could work if we had a sufficient number of relations in which the nucleus could be determined without knowing the element's importance. The information that those relations provide could in turn specify the other relations which are added.

In chapter 12 I will examine my assumptions more closely.

Chapter 12

Conclusion

> 'But seriously now, Knipe. D'you really think they'd buy them?' 'Listen, Mr Bohlen. Who on earth is going to want custom-made stories when they can get the other kind at half the price? It stands to reason, doesn't it?'
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 599

In this chapter I will discuss the results, as presented through out this thesis. Unfortunately, the evaluation did not yield any results, but I'll come to that shortly.

12.1 Fields of research

I put forth an overview of different terms and concepts in order to clarify the setting for this research. I've found it difficult to combine the work done in the different fields of narratology, natural language generation and language analysis, because the different fields use similar terms for different concepts.

12.2 Integrating Hovy's work

The work done by Hovy seemed to be ideal for the Virtual Storyteller; he combined diverse definitions from the fields of natural language generation and text analysis to compile a list of rhetorical relations that would be sufficient endeavors in both fields. Unfortunately, when reviewing his list and collecting the sources Hovy used, I happened upon some weird references or other differences. Definitions didn't quite seem to match, and this made it even more unfortunate that Hovy himself did not provide definitions for these relations. I set forth to provide these definitions, in order to clarify the taxonomy of rhetorical relations.

At this point I should have realized that it might be necessary to focus all my attention on this subject, and not to maintain one of my original goals; automatic paragraph separation.

12.3 Separating paragraphs

I researched the theory behind the use of paragraphs, but found mostly information on the function of paragraphs in formal texts, while I was interested in fairy-tales. Outside of formal writing, most literature I found seems to assume you already know what to do with paragraphs. I've tried to transfer the wisdom on paragraphs in formal text to paragraphs in informal text, and executed a small experiment on this subject.

12.4 Implementing my ideas

When I had defined my two approaches to my research - my tool was a thorough understanding of rhetorical relations and my goal was to create paragraph separation - I started exploring the current implementation of the Virtual Storyteller. I was soon confronted with the difficulties of joining a project of considerable magnitude which was doing something relatively new and unknown. The earlier versions had given credence to the underlying ideas, but the current version tried to make a big leap forward on several fronts, simultaneously. After a considerable amount of effort, I gave up trying to integrate my work into the Virtual Storyteller. I instead focused on the ideas behind my methods, so that they might one day yet be implemented.

12.5 Algorithm on paper

Which leads us to the disappointing conclusion of this thesis: during evaluation, it quickly becomes clear that either the underlying theory is wrong, or it is simply untestable on paper. I would have loved to tinker about within a working environment, analyzing my methods while applying them on different fabulas, but when the creation of a fabula takes days this becomes much more cumbersome.

12.6 Conclusion

In conclusion, I can only say I've learned a great deal from this pursuit. In the following chapter, I will discuss several ideas for work that can be done in the future, on subjects that should have completed here and others.

Chapter 13

Future Work

> But that wasn't all. Control had also to be exercised during the actual writing process (which took about fifteen minutes per novel), and to do this the author had to sit, as it were, in the driver's seat, and pull (or push) a battery of labelled stops, as on an organ. By so doing, she was able continually to modulate or merge fifty different and variable qualities such as tension, surprise, humour, pathos, and mystery. Numerous dials and gauges on the dashboard itself told him throughout exactly how far along he was with his work.
>
> From *The Great Automatic Grammatizator* by Roahl Dahl [6], page 604

In this chapter I will suggest several improvements that can be made to the Virtual Storyteller, and explain some ideas I've had about what such a project as this could deliver in the future.

13.1 Character agents

Starting with the character agents, I must note that there is already work being done on improving the creativity of these agents. We want them to be able to solve problems in unpredictable ways, in order to provide the characters with an actual personality. I would like to add that these kinds of endeavors are truly future work, and could have perhaps waited until after a basic version of a planning algorithm was complete?

I am also curious to see what kind of regularities can be found in the cast of different kinds of narratives. What distribution of attributes do you find among the characters? And could these sets of characters be generated automatically?

13.2 The storyworld

The storyworld is defined rather rigorously, and by that I mostly point to the available actions. Though they are defined within an ontology, the lowest level is set to action as we would use them in a story. The characters can sing, fight and dance. But, as I mentioned earlier in this thesis, do we not also wish to describe the movement of the feet when the princess dances? Or describe the powerful moves of the fighting prince? This would of

course impose an immense complexity on the entire system, but it would be a marvelous improvement.

13.3 The Plot Agent

At the moment, the plot agent aspires to influence the plot using suggested actions and manipulating the world through events. I would like to expand these options in order to create a *goal based* storyteller. Configuring the storyteller to become goal based, so that it can generate stories with a certain end situation, would unlock several possibilities. For instance, it would enable characters to invoke this part of the program to reconstruct what happened in their absence. It would would also enable them to *lie*, by constructing a new piece of fabula that ends in a similar situation, but in which certain constraints aren't broken.

13.4 The Narrator Agent

It is this agent that still needs a lot of work. I would like to see all of the rhetorical relations listed by Hovy researched and implemented. I feel that then, despite the incredible amount of information you have to deal with, it becomes easier to create a certain type of discourse structure. It might be necessary to assign weights to the different relations, so that a different discourse structure tree would emerge when these weights are adjusted.

Because that is what this technology would also enable: if a technical manual would be created as a fabula, and the reader would be able to determine the amount of explanation that would be incorporated in the presented text, would that not be desirable? If the fabula contains so much more information than the reader would like to see, why not give the reader a choice in which bits he or she reads?

Returning to a topic closer to this research, I wonder if it would be possible to base a paragraph structure on a story grammar. I am not suggesting to create the story according to a story grammar, but try to match a story grammar to the discourse tree that is produced.

13.5 The Presentation Agent

But who says we have to present the reader with text? Merlin might still be willing to read the tales, but I am also thinking of output in the form of comics or movies. Once again, I think presenting these options to the user shows the true power of a system like the Virtual Storyteller.

13.6 The other way around

A truly remarkable, and perhaps impossible feat, would be to reverse the functionality of the Virtual Storyteller. Have the program listen to natural language, apply text analysis to create a discourse structure tree, and then store this information in a fabula structure.

Perhaps we could even try to run a simulation with character agents that try to determine what kind of characteristics they need to create such a fabula. The program would act like a human listener, getting to know the characters. And then the program could take this fabula and run with it.

13.7 Free to imagine

Despite this research coming to a less than satisfactory end, I am still glad to be a part of such a project. As is probably clear from the ideas presented above, these kinds of projects appeal to the imagination, as only storytellers can.

Bibliography

[1] H. Porter Abbott. *The Cambridge Introduction to Narrative*. Cambridge University Press, Cambridge, 2002.

[2] R. Aylett. Narrative in virtual environments - towards emergent narrative. In *AAAI Fall Symposium on Narrative Intelligence*, pages 83–86, 1999.

[3] J. Bates. The nature of characters in interactive worlds and the Oz project, 1992.

[4] Peter Burger and Jaap de Jong. *Handboek Stijl. Adviezen voor aantrekkelijk schrijven*. Martinus Nijhoff Uitgevers, 1997.

[5] Charles Callaway and James Lester. Narrative prose generation. In *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence*, pages 1241–1248, Seattle, WA, 2001.

[6] Roald Dahl. *Collected Short Stories*. Penguin, London, England, 1992.

[7] G. Davenport and M. Murtaugh. Automatist storyteller systems and the shifting sands of story. *IBM Systems Journal*, 36(3):446 – 456, 1997.

[8] S. Faas. Virtual storyteller: An approach to computational story telling. Master's thesis, University of Twente, 2002.

[9] F. Hielkema. Performing syntactic aggregation using discourse structure. Master's thesis, University of Groningen, 2005.

[10] E. Hovy and E. Maier. Parsimonious or profligate: How many and which discourse structure relations. unpublished, 1997.

[11] Sheldon Klein. Automatic novel writer: A status report. In János S. Petöfi, editor, *Research in Text Theory*, volume 3. 1973.

[12] LLC Lexico Publishing Group. Dictionary.com.

[13] William C. Mann and Sandra A. Thompson. Rhetorical structure theory: A theory of text organization. Technical report, USC Information Sciences Institute, Marina del Rey, California, 1987.

[14] James R. Meehan. Tale-spin, an interactive program that writes stories. In *Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, volume one, pages 91–98, Cambridge, Massachusetts, USA, august 1977. IJCAI, Morgan Kaufman Publishers, Inc.

[15] Susan Onega and Jose Angel Garcia Landa. *Narratology.* Longman Publishing, London, New York, 1996.

[16] Rui Prada, Isabel Machado, and Ana Paiva. TEATRIX: Virtual environment for story creation. In *Intelligent Tutoring Systems*, pages 464–473, 2000.

[17] Vladimir Propp. *De morfologie van het toversprookje: vormleer van een genre (S. Butcher, Translation).* Spectrum, Utrecht, 1997 (Original published in 1968).

[18] S.H.J. Rensen. De virtuele verhalenverteller: agent-gebaseerde generatie van interessante plots. Master's thesis, University of Twente, 2004.

[19] N. Slabbers. Narration for virtual storytelling. Master's thesis, University of Twente, March 2006.

[20] I. Swartjes. The plot thickens: bringing meaning and structure into automated story generation. Master's thesis, University of Twente, March 2006.

[21] M. Theune, S. Faas, A. Nijholt, and D. Heylen. The virtual storyteller: Story creation by intelligent agents. In *Proceedings of the Technologies for Interactive Digital Storytelling and Entertainment (TIDSE) Conference*, pages 205–214, 2003.

[22] Eric Tiggeler. *Vraagbaak Nederlands. Van spelling tot stijl: snel een helder antwoord op praktijkvragen over taal.* Sdu Uitgevers, Den Haag, 2001.

[23] J.R.R. Uijlings. Designing a virtual environment for story generation. Master's thesis, University of Amsterdam, June 2006.

[24] Jan van de Pol. *In goed Nederlands.* Sdu Uitgevers, Den Haag, 2002.

[25] Han van Gessel; Jan Kees Hulsbosch; Bas van Kleef; Theo Stielstra, editor. *De Volkskrant Stijlboek.* Sdu Uitgevers, Den Haag, 2002.

Appendix A

Output Examples

A.1 Automated Novel Writer

```
JAMES WAS VERY RICH.
CLIVE WAS IMPOVERISHED.
CLIVE WANTED MONEY.
THE BUTLER WAS RELATED TO JAMES.
CLIVE THOUGHT THAT CLIVE INHERITED THE MONEY.
CLIVE KNEW THAT JAMES DRANK A MILK.
CLIVE POISONED THE MILK.
JAMES DRANK THE MILK.
JAMES WENT TO BED.
JAMES DIED.
THE OTHERS THOUGHT THAT JAMES WAS ASLEEP.
THE BUTLER RETURNED THE BOTTLE.
RONALD AWAKENED.
RONALD GOT UP.
RONALD THOUGHT THAT THE DAY WAS BEAUTIFUL.
RONALD FOUND JAMES.
RONALD SAW THAT JAMES WAS DEAD.
RONALD YELLED.
THE OTHERS AWAKENED.
THE OTHERS RAND TO RONALD.
THE OTHERS SAW JAMES.
EVERYONE TALKED.
HEATHER CALLED THE POLICEMEN.
HUME EXAMINED THE BODY.
DR. BARTHOLOMEW HUME SAID THAT JAMES WAS KILLED BY POISON.
```

A.2 TALE-SPIN

ONCE UPON A TIME GEORGE ANT LIVED NEAR A PATCH OF GROUND. THERE WAS A NEST IN AN ASH TREE. WILMA BIRD LIVED IN THE NEST. THERE WAS SOME WATER IN A RIVER. WILMA KNEW THAT THE WATER WAS IN

THE RIVER. GEORGE KNEW THAT THE WATER WAS IN THE RIVER. ONE DAY WILMA WAS VERY THIRSTY. WILMA WANTED TO GET NEAR SOME WATER. WILMA FLEW FROM HER NEST ACROSS A MEADOW THROUGH A VALLEY TO THE RIVER. WILMA DRANK THE WATER. WILMA WASN'T THIRSTY ANY MORE.

GEORGE WAS VERY THIRSTY. GEORGE WANTED TO GET NEAR SOME WATER. GEORGE WALKED FROM HIS PATCH OF GROUND ACROSS THE MEADOW THROUGH THE VALLEY TO A RIVER BANK. GEORGE FELL INTO THE WATER. GEORGE WANTED TO GET NEAR THE VALLEY. GEORGE WANTED TO GET NEAR THE MEADOW. GEORGE COULDN'T GET NEAR THE MEADOW. WILMA WANTED GEORGE TO GET NEAR THE MEADOW. WILMA WANTED TO GET NEAR GEORGE. WILMA GRABBED GEORGE WITH HER CLAW. WILMA TOOK GEORGE FROM THE RIVER THROUGH THE VALLEY TO THE MEADOW. GEORGE WAS DEVOTED TO WILMA. GEORGE OWED EVERYTHING TO WILMA. WILMA LET GO OF GEORGE. GEORGE FELL TO THE MEADOW. THE END.

A.3 StoryBook

Once upon a time, a woodcutter and his wife lived in a small cottage. The woodcutter and his wife had a young daughter, whom everyone called Little Red Riding Hood. She was a merry little maid, and all day long she went singing aoubt the house. Her mother loved her very much.

One day her mother said, "My child, go to grandmother's house. We have not heard from her for some time. Take these cakes, but do not stay too long. And, beware the dangers in the forest."

Little Red Riding Hood was delighted because she was very fond of her grandmother. Her mother gave her a well-filled basket and kisser her goodbye.

The road to grandmother's house led through the dark forest, but Little Red Riding Hood was not afraid and she went on as happy as a lark. The birds sang her their sweetest song while the squirrels ran up and down the tall trees. Now and then, a rabbit would cross her path.

Little Red Riding Hood had not gone far when she met a wolf.

"Hello," greeted the wolf, who was a cunning-looking creature. "Where are you going?"

"I am going to my grandmother's house," Little Red Riding Hood replied.

"Ah, well then, take care in the forest, for there are many dangers." And then the wolf left.

Little Red Riding Hood was not in a hurry. Indeed, she gathered wild flowers and chased the pretty butterflies.

Meanwhile the wolf ran ahead very quickly and soon arrived at grandmother's house. He knocked on the door gently. The old lady asked, "Who is there?"

The wolf replied, "It is Little Red Riding Hood, your granddaughter."

And so the old lady opened the cottage door. The wolf rushed in immediately and devoured the lady in one bite. Then he shut the door and climbed into the old lady's bed.

Much later Little Red Riding Hood arrived at grandmother's house. She knockd on the door and shouted, "Grandmother, it is Little Red Riding Hood."

"Pull the string. The door will open."

And so Little Red Riding Hood opened the door and walked in. "Grandmother, what big eyes you have."

"All the better to see with, dear."

"Grandmother, what big ears you have."

"All the better to hear with, dear."

"And, grandmother, what big teeth you have!"

"All the better to eat up with!" yelled the wolf.

And then the wolf jumped up and devoured Little Red Riding Hood in one bite.

Appendix B

Summaries of stories

Sander Faas [8] wrote summaries of four different well-known fairytales in order to analyse the storyline of such stories.

B.1 Hansel and Gretel

Once upon a time there lived two children, Hansel and Gretel. One day Hansel and Gretel were lost in the forest. After some time they arrived at a gingerbread house. Inside the house lived a witch. The witch put Hansel into a cage and wanted to eat him. Gretel pushed the witch into the oven. She freed Hansel and together they lived happily ever after.

B.2 Little Red Riding Hood

Once upon a time there was a girl called Little Red Riding Hood. One day Little Red Riding Hood walked through the woods to her grandmother's house. Instead of her grandmother there was a big bad wolf inside the house! The wolf ate Little Red Riding Hood and fell asleep. Then a hunter passed by. He cut open the wolf's belly and Little Red Riding Hood came out. She lived happily ever after.

B.3 Snowwhite

Once upon a time there lived a beautiful princess, called Snowwhite. She lived with seven dwarfs. One day her evil stepmother came by. She gave Snowwhite a poisonous apple. Showwhite died. The dwarves carried Snowwhite away, but suddenly one stumbled. The chunk of poisonous apple fell out of Snowwhite's mouth and she woke up. She married a prince and lived happily ever after.

B.4 Tom Thumb

Once upon a time there lived a boy so small that he was called Tom Thumb. One day Tom was lost in the woods. He came at the house of a giant. The giant wanted to eat Tome Thumb, but he ran away. The giant put on his seven-league boots and ran after

Tom. Soon the giant got tired and fell asleep. Tom snatched the seven-league boots which made him run faster than the giant. The giant gave up and Tom Thumb lived happily ever after.

Appendix C

De Slechte Vorst - Andersen

C.1 the tale

De Slechte Vorst

Er was eens een vorst die de hele wereld onder zijn heerschappij wilde brengen. En niet door wijsheid, goedheid, eerlijkheid, rechtvaardigheid en vrede, maar door oorlog te voeren. Hij was zo slecht dat hij eerder een schepsel van de duivel dan van God leek te zijn.

Te vuur en te zwaard overwonnen zijn soldaten de buurlanden; er stroomden rivieren van bloed en tranen, en van vele trotse steden bleven slechts smeulende puinhopen over. De wrede huurlingen van de vorst slachten zelfs hele kuddes vee af; honden en katten die zich niet snel genoeg uit de voeten maakten, werden door hen aan het zwaard geregen.

De slechte vorst was meedogenloos. Hij was uitsluitend genteresseerd in een zo groot mogelijke oorlogsbuit uit de paleizen, kerken en kloosters die hij overwonnen had. Zijn schatkamer stroomde over met goud, edelstenen, parels en juwelen: zelfs gouden deurknoppen liet hij loswrikken! Met al deze schatten bouwden zijn architecten schitterende paleizen, burchten en kastelen, zodat iedereen die ze zag van angst voor hem vervuld zou raken en zou buigen voor zijn almacht. Maar daar was de vorst nog niet tevreden mee. 'Ik wil de machtigste vorst op aarde worden!' zei hij iedere dag.

Toen hij alle omringende landen veroverd had, liet hij hun koningen vastbinden achter zijn gouden triomfwagen, en hij liet ze vertonen in zijn hoofdstad alsof het wilde dieren waren. Hij genoot van de bewondering van zijn onderdanen.

Na vele veldtochten stonden zijn legers aan de oever van de zee: er was geen enkel land op aarde dat nog niet veroverd was. De vorst besloot nu dat op ieder stadsplein een standbeeld van hem moest worden opgericht. Maar ook dat was hem nog niet voldoende eer. Hij besloot dat zijn standbeeld ook moest worden opgericht in alle kathedralen, kerken, moskeen en synagogen van de overwonnen landen. De priesters verzetten zich als enige tegen dit voorstel. 'U heerst over de mensen in vele landen, machtige vorst, maar u bent niet de heerser over de hemel. God is nog altijd machtiger dan u. We kunnen

uw bevel niet opvolgen: er kunnen geen standbeelden van u worden opgericht in onze godshuizen!'

De vorst bedreigde hen met de marteldood, maar de priesters lieten zich hierdoor niet afschrikken.

Nu had de vorst nog maar één doel voor ogen: hij zou ook de hemel veroveren. Hij begon onmiddellijk met de voorbereidingen. Zijn architecten maakten bouwtekeningen voor een luchtschip, dat voortgetrokken zou worden door duizend adelaars: daarmee zou hij de hemel bestormen. Zijn wapenmeesters maakten kanonnen die zo groot waren dat ze het hemelse paleis in puin konden schieten.

Zeven jaar lang was het leger in de weer met de voorbereidingen voor de aanval. Toen was alles klaar: een groot luchtschip voor de legertroepen, en een kleiner luchtschip voor de vorst, die zelf de aanval wilde inzetten.

De militaire kapel speelde. De maarschalken staken lange redevoeringen af voor de aanval werd geopend. De adelaars die het merkwaardige luchtschip zouden voorttrekken zaten nog in hun kooien.

De soldaten stelden zich in een mijlenlange rij op om plaats te nemen in hun luchtschip. Ook de vorst zelf, gehuld in zijn vergulde harnas, steeg aan boord. Hij stond op de brug, klaar om het eerste bevel te geven.

Plotseling kwam er een wolk muggen aanzwermen: alle muggen stortten zich op de vorst en zoemden oorverdovend. De machtige vorst trok zijn zwaard en sloeg om zich heen, maar hij sloeg alleen maar in de lucht.

Een van de muggen wist zijn helm binnen te dringen en stak de meedogenloze vorst in zijn oor. De vorst begon om hulp te roepen, sprong uit zijn luchtschip en zette het op een lopen in de richting van het bos. Ondertussen rukte hij alle onderdelen van zijn wapenuitrusting los: poedelnaakt rende hij weg, voor het oog van al zijn legertroepen; en aldoor werd hij omstuwd door een horde boze muggen.

De vorst verdween in het bos, en werd nooit meer teruggezien. Het hele leger lachte en slaakte een zucht van opluchting. De machtige vorst die de hemel wilde bestormen ging roemloos ten onder!

Appendix D

Hovy's discourse structure relations

- Ideational
  - Elaboration
    - Elab-Object
      - Object-Attribute
      - Object-Function
    - Elab-Part
      - Set-Member
      - Process-Step
      - Whole-Part
    - Elab-Generality
      - General-Specific
      - Abstract-Instance
    - Identification
    - Restatement
      - Conclusion (interp at end)
      - Summary (short restatement)
  - Circumstance
    - Location
    - Time
    - Means
    - Manner
    - Instrument
    - Parallel-Event
  - Sequence
    - Seq-Temporal
    - Seq-Spatial
    - Seq-Ordinal

  - Cause/Result
    - C/RVol (conditional)
      - Vol-Cause
      - Vol-Result
    - C/RNonvol (nonvolitional)
      - NonVol-Cause
      - NonVol-Result
    - Purpose
  - General-Condition
    - Condition
    - Exception
  - Comparative
    - Equative (like, while)
    - Contrast
    - Otherwise (if then else)
    - Comparison
    - Analogy

- Interpersonal
  - Interpretation
    - Evalulation (A opinion)
  - Enablement
    - Background
  - Antithesis
  - Exhortation
    - Support
      - Solutionhood (general prob)
      - Answer (numeric prob)
      - Evidence (support claim)
      - Proof
      - Justification (for A act)
      - Motivation (for R act)
    - Concession
    - Qualification

- Textual
  - Logical-Relation
    - Conjunction
    - Disjunction
  - Pres-Sequence
  - Joint

Appendix E

Relations added to fabula example

E.1 ObjectAttribute

1. ObjectAttribute(002, 001)
2. ObjectAttribute(003, 001)
3. ObjectAttribute(004, 001)
4. ObjectAttribute(005, 001)
5. ObjectAttribute(015, 001)
6. ObjectAttribute(016, 001)
7. ObjectAttribute(017, 001)
8. ObjectAttribute(018, 001)
9. ObjectAttribute(019, 001)
10. ObjectAttribute(021, 001)
11. ObjectAttribute(022, 001)
12. ObjectAttribute(023, 001)
13. ObjectAttribute(024, 001)
14. ObjectAttribute(025, 001)
15. ObjectAttribute(030, 001)
16. ObjectAttribute(031, 001)
17. ObjectAttribute(032, 001)
18. ObjectAttribute(033, 001)
19. ObjectAttribute(034, 001)

20. ObjectAttribute(040, 001)
21. ObjectAttribute(041, 001)
22. ObjectAttribute(042, 001)
23. ObjectAttribute(043, 001)
24. ObjectAttribute(044, 001)
25. ObjectAttribute(045, 001)
26. ObjectAttribute(046, 001)
27. ObjectAttribute(051, 001)
28. ObjectAttribute(052, 001)
29. ObjectAttribute(053, 001)
30. ObjectAttribute(054, 001)
31. ObjectAttribute(055, 001)
32. ObjectAttribute(060, 001)
33. ObjectAttribute(061, 001)
34. ObjectAttribute(062, 001)
35. ObjectAttribute(063, 001)
36. ObjectAttribute(064, 001)
37. ObjectAttribute(069, 001)
38. ObjectAttribute(070, 001)
39. ObjectAttribute(071, 001)
40. ObjectAttribute(072, 001)
41. ObjectAttribute(073, 001)
42. ObjectAttribute(083, 001)
43. ObjectAttribute(084, 001)
44. ObjectAttribute(085, 001)
45. ObjectAttribute(086, 001)
46. ObjectAttribute(087, 001)
47. ObjectAttribute(088, 001)
48. ObjectAttribute(089, 001)

49. ObjectAttribute(090, 001)
50. ObjectAttribute(091, 001)
51. ObjectAttribute(092, 001)
52. ObjectAttribute(093, 001)
53. ObjectAttribute(094, 001)
54. ObjectAttribute(098, 001)
55. ObjectAttribute(099, 001)
56. ObjectAttribute(100, 001)
57. ObjectAttribute(101, 001)
58. ObjectAttribute(102, 001)

E.2 Process-Step

1. Process-Step(009,010)
2. Process-Step(009,011)
3. Process-Step(009,012)
4. Process-Step(009,013)

E.3 Parallel-Event

1. ParallelEvent(009,010)
2. ParallelEvent(009,011)
3. ParallelEvent(009,012)
4. ParallelEvent(009,013)
5. ParallelEvent(010,009)
6. ParallelEvent(010,011)
7. ParallelEvent(010,012)
8. ParallelEvent(010,013)
9. ParallelEvent(011,009)
10. ParallelEvent(011,010)
11. ParallelEvent(011,012)

12. ParallelEvent(011,013)
13. ParallelEvent(012,009)
14. ParallelEvent(012,010)
15. ParallelEvent(012,011)
16. ParallelEvent(012,013)
17. ParallelEvent(013,009)
18. ParallelEvent(013,010)
19. ParallelEvent(013,011)
20. ParallelEvent(013,012)
21. ParallelEvent(055,064)
22. ParallelEvent(064,055)
23. ParallelEvent(059,068)
24. ParallelEvent(068,059)
25. ParallelEvent(077,081)
26. ParallelEvent(081,077)

E.4 SeqTemporal

1. SeqTemporal(005,009)
2. SeqTemporal(005,010)
3. SeqTemporal(005,011)
4. SeqTemporal(005,012)
5. SeqTemporal(005,013)
6. SeqTemporal(005,025)
7. SeqTemporal(005,029)
8. SeqTemporal(005,034)
9. SeqTemporal(005,038)
10. SeqTemporal(005,043)
11. SeqTemporal(005,046)
12. SeqTemporal(005,049)

13. SeqTemporal(005,050)
14. SeqTemporal(005,055)
15. SeqTemporal(005,059)
16. SeqTemporal(005,064)
17. SeqTemporal(005,068)
18. SeqTemporal(005,073)
19. SeqTemporal(005,077)
20. SeqTemporal(005,081)
21. SeqTemporal(005,082)
22. SeqTemporal(005,083)
23. SeqTemporal(005,084)
24. SeqTemporal(005,089)
25. SeqTemporal(005,094)
26. SeqTemporal(005,097)
27. SeqTemporal(009,010)
28. SeqTemporal(009,011)
29. SeqTemporal(009,012)
30. SeqTemporal(009,013)
31. SeqTemporal(009,025)
32. SeqTemporal(009,029)
33. SeqTemporal(009,034)
34. SeqTemporal(009,038)
35. SeqTemporal(009,043)
36. SeqTemporal(009,046)
37. SeqTemporal(009,049)
38. SeqTemporal(009,050)
39. SeqTemporal(009,055)
40. SeqTemporal(009,059)
41. SeqTemporal(009,064)

42. SeqTemporal(009,068)
43. SeqTemporal(009,073)
44. SeqTemporal(009,077)
45. SeqTemporal(009,081)
46. SeqTemporal(009,082)
47. SeqTemporal(009,083)
48. SeqTemporal(009,084)
49. SeqTemporal(009,089)
50. SeqTemporal(009,094)
51. SeqTemporal(009,097)
52. SeqTemporal(010,011)
53. SeqTemporal(010,012)
54. SeqTemporal(010,013)
55. SeqTemporal(010,025)
56. SeqTemporal(010,029)
57. SeqTemporal(010,034)
58. SeqTemporal(010,038)
59. SeqTemporal(010,043)
60. SeqTemporal(010,046)
61. SeqTemporal(010,049)
62. SeqTemporal(010,050)
63. SeqTemporal(010,055)
64. SeqTemporal(010,059)
65. SeqTemporal(010,064)
66. SeqTemporal(010,068)
67. SeqTemporal(010,073)
68. SeqTemporal(010,077)
69. SeqTemporal(010,081)
70. SeqTemporal(010,082)

71. SeqTemporal(010,083)
72. SeqTemporal(010,084)
73. SeqTemporal(010,089)
74. SeqTemporal(010,094)
75. SeqTemporal(010,097)
76. SeqTemporal(011,012)
77. SeqTemporal(011,013)
78. SeqTemporal(011,025)
79. SeqTemporal(011,029)
80. SeqTemporal(011,034)
81. SeqTemporal(011,038)
82. SeqTemporal(011,043)
83. SeqTemporal(011,046)
84. SeqTemporal(011,049)
85. SeqTemporal(011,050)
86. SeqTemporal(011,055)
87. SeqTemporal(011,059)
88. SeqTemporal(011,064)
89. SeqTemporal(011,068)
90. SeqTemporal(011,073)
91. SeqTemporal(011,077)
92. SeqTemporal(011,081)
93. SeqTemporal(011,082)
94. SeqTemporal(011,083)
95. SeqTemporal(011,084)
96. SeqTemporal(011,089)
97. SeqTemporal(011,094)
98. SeqTemporal(011,097)
99. SeqTemporal(012,013)

100. SeqTemporal(012,025)
101. SeqTemporal(012,029)
102. SeqTemporal(012,034)
103. SeqTemporal(012,038)
104. SeqTemporal(012,043)
105. SeqTemporal(012,046)
106. SeqTemporal(012,049)
107. SeqTemporal(012,050)
108. SeqTemporal(012,055)
109. SeqTemporal(012,059)
110. SeqTemporal(012,064)
111. SeqTemporal(012,068)
112. SeqTemporal(012,073)
113. SeqTemporal(012,077)
114. SeqTemporal(012,081)
115. SeqTemporal(012,082)
116. SeqTemporal(012,083)
117. SeqTemporal(012,084)
118. SeqTemporal(012,089)
119. SeqTemporal(012,094)
120. SeqTemporal(012,097)
121. SeqTemporal(013,025)
122. SeqTemporal(013,029)
123. SeqTemporal(013,034)
124. SeqTemporal(013,038)
125. SeqTemporal(013,043)
126. SeqTemporal(013,046)
127. SeqTemporal(013,049)
128. SeqTemporal(013,050)

129. SeqTemporal(013,055)
130. SeqTemporal(013,059)
131. SeqTemporal(013,064)
132. SeqTemporal(013,068)
133. SeqTemporal(013,073)
134. SeqTemporal(013,077)
135. SeqTemporal(013,081)
136. SeqTemporal(013,082)
137. SeqTemporal(013,083)
138. SeqTemporal(013,084)
139. SeqTemporal(013,089)
140. SeqTemporal(013,094)
141. SeqTemporal(013,097)
142. SeqTemporal(025,029)
143. SeqTemporal(025,034)
144. SeqTemporal(025,038)
145. SeqTemporal(025,043)
146. SeqTemporal(025,046)
147. SeqTemporal(025,049)
148. SeqTemporal(025,050)
149. SeqTemporal(025,055)
150. SeqTemporal(025,059)
151. SeqTemporal(025,064)
152. SeqTemporal(025,068)
153. SeqTemporal(025,073)
154. SeqTemporal(025,077)
155. SeqTemporal(025,081)
156. SeqTemporal(025,082)
157. SeqTemporal(025,083)

158. SeqTemporal(025,084)
159. SeqTemporal(025,089)
160. SeqTemporal(025,094)
161. SeqTemporal(025,097)
162. SeqTemporal(029,029)
163. SeqTemporal(029,034)
164. SeqTemporal(029,038)
165. SeqTemporal(029,043)
166. SeqTemporal(029,046)
167. SeqTemporal(029,049)
168. SeqTemporal(029,050)
169. SeqTemporal(029,055)
170. SeqTemporal(029,059)
171. SeqTemporal(029,064)
172. SeqTemporal(029,068)
173. SeqTemporal(029,073)
174. SeqTemporal(029,077)
175. SeqTemporal(029,081)
176. SeqTemporal(029,082)
177. SeqTemporal(029,083)
178. SeqTemporal(029,084)
179. SeqTemporal(029,089)
180. SeqTemporal(029,094)
181. SeqTemporal(029,097)
182. SeqTemporal(034,038)
183. SeqTemporal(034,043)
184. SeqTemporal(034,046)
185. SeqTemporal(034,049)
186. SeqTemporal(034,050)

187. SeqTemporal(034,055)
188. SeqTemporal(034,059)
189. SeqTemporal(034,064)
190. SeqTemporal(034,068)
191. SeqTemporal(034,073)
192. SeqTemporal(034,077)
193. SeqTemporal(034,081)
194. SeqTemporal(034,082)
195. SeqTemporal(034,083)
196. SeqTemporal(034,084)
197. SeqTemporal(034,089)
198. SeqTemporal(034,094)
199. SeqTemporal(034,097)
200. SeqTemporal(038,038)
201. SeqTemporal(038,043)
202. SeqTemporal(038,046)
203. SeqTemporal(038,049)
204. SeqTemporal(038,050)
205. SeqTemporal(038,055)
206. SeqTemporal(038,059)
207. SeqTemporal(038,064)
208. SeqTemporal(038,068)
209. SeqTemporal(038,073)
210. SeqTemporal(038,077)
211. SeqTemporal(038,081)
212. SeqTemporal(038,082)
213. SeqTemporal(038,083)
214. SeqTemporal(038,084)
215. SeqTemporal(038,089)

216. SeqTemporal(038,094)
217. SeqTemporal(038,097)
218. SeqTemporal(043,046)
219. SeqTemporal(043,049)
220. SeqTemporal(043,050)
221. SeqTemporal(043,055)
222. SeqTemporal(043,059)
223. SeqTemporal(043,064)
224. SeqTemporal(043,068)
225. SeqTemporal(043,073)
226. SeqTemporal(043,077)
227. SeqTemporal(043,081)
228. SeqTemporal(043,082)
229. SeqTemporal(043,083)
230. SeqTemporal(043,084)
231. SeqTemporal(043,089)
232. SeqTemporal(043,094)
233. SeqTemporal(043,097)
234. SeqTemporal(046,049)
235. SeqTemporal(046,050)
236. SeqTemporal(046,055)
237. SeqTemporal(046,059)
238. SeqTemporal(046,064)
239. SeqTemporal(046,068)
240. SeqTemporal(046,073)
241. SeqTemporal(046,077)
242. SeqTemporal(046,081)
243. SeqTemporal(046,082)
244. SeqTemporal(046,083)

245. SeqTemporal(046,084)
246. SeqTemporal(046,089)
247. SeqTemporal(046,094)
248. SeqTemporal(046,097)
249. SeqTemporal(049,050)
250. SeqTemporal(049,055)
251. SeqTemporal(049,059)
252. SeqTemporal(049,064)
253. SeqTemporal(049,068)
254. SeqTemporal(049,073)
255. SeqTemporal(049,077)
256. SeqTemporal(049,081)
257. SeqTemporal(049,082)
258. SeqTemporal(049,083)
259. SeqTemporal(049,084)
260. SeqTemporal(049,089)
261. SeqTemporal(049,094)
262. SeqTemporal(049,097)
263. SeqTemporal(050,055)
264. SeqTemporal(050,059)
265. SeqTemporal(050,064)
266. SeqTemporal(050,068)
267. SeqTemporal(050,073)
268. SeqTemporal(050,077)
269. SeqTemporal(050,081)
270. SeqTemporal(050,082)
271. SeqTemporal(050,083)
272. SeqTemporal(050,084)
273. SeqTemporal(050,089)

274. SeqTemporal(050,094)
275. SeqTemporal(050,097)
276. SeqTemporal(055,059)
277. SeqTemporal(055,064)
278. SeqTemporal(055,068)
279. SeqTemporal(055,073)
280. SeqTemporal(055,077)
281. SeqTemporal(055,081)
282. SeqTemporal(055,082)
283. SeqTemporal(055,083)
284. SeqTemporal(055,084)
285. SeqTemporal(055,089)
286. SeqTemporal(055,094)
287. SeqTemporal(055,097)
288. SeqTemporal(059,064)
289. SeqTemporal(059,068)
290. SeqTemporal(059,073)
291. SeqTemporal(059,077)
292. SeqTemporal(059,081)
293. SeqTemporal(059,082)
294. SeqTemporal(059,083)
295. SeqTemporal(059,084)
296. SeqTemporal(059,089)
297. SeqTemporal(059,094)
298. SeqTemporal(059,097)
299. SeqTemporal(064,068)
300. SeqTemporal(064,073)
301. SeqTemporal(064,077)
302. SeqTemporal(064,081)

303. SeqTemporal(064,082)
304. SeqTemporal(064,083)
305. SeqTemporal(064,084)
306. SeqTemporal(064,089)
307. SeqTemporal(064,094)
308. SeqTemporal(064,097)
309. SeqTemporal(068,073)
310. SeqTemporal(068,077)
311. SeqTemporal(068,081)
312. SeqTemporal(068,082)
313. SeqTemporal(068,083)
314. SeqTemporal(068,084)
315. SeqTemporal(068,089)
316. SeqTemporal(068,094)
317. SeqTemporal(068,097)
318. SeqTemporal(073,077)
319. SeqTemporal(073,081)
320. SeqTemporal(073,082)
321. SeqTemporal(073,083)
322. SeqTemporal(073,084)
323. SeqTemporal(073,089)
324. SeqTemporal(073,094)
325. SeqTemporal(073,097)
326. SeqTemporal(077,081)
327. SeqTemporal(077,082)
328. SeqTemporal(077,083)
329. SeqTemporal(077,084)
330. SeqTemporal(077,089)
331. SeqTemporal(077,094)

332. SeqTemporal(077,097)
333. SeqTemporal(081,082)
334. SeqTemporal(081,083)
335. SeqTemporal(081,084)
336. SeqTemporal(081,089)
337. SeqTemporal(081,094)
338. SeqTemporal(081,097)
339. SeqTemporal(082,083)
340. SeqTemporal(082,084)
341. SeqTemporal(082,089)
342. SeqTemporal(082,094)
343. SeqTemporal(082,097)
344. SeqTemporal(083,084)
345. SeqTemporal(083,089)
346. SeqTemporal(083,094)
347. SeqTemporal(083,097)
348. SeqTemporal(084,089)
349. SeqTemporal(084,094)
350. SeqTemporal(084,097)
351. SeqTemporal(089,094)
352. SeqTemporal(089,097)
353. SeqTemporal(094,097)

E.5 C/RVol

1. VolResult(004,005)
2. VolResult(008,009)
3. VolResult(008,010)
4. VolResult(008,011)
5. VolResult(008,012)

6. VolResult(008,013)
7. VolResult(017,025)
8. VolResult(028,029)
9. VolResult(033,034)
10. VolResult(037,038)
11. VolResult(042,043)
12. VolResult(045,046)
13. VolResult(048,049)
14. VolResult(048,050)
15. VolResult(054,055)
16. VolResult(058,059)
17. VolResult(063,064)
18. VolResult(067,068)
19. VolResult(072,073)
20. VolResult(076,077)
21. VolResult(080,081)
22. VolResult(088,089)
23. VolResult(093,094)
24. VolResult(096,097)
25. VolCause(005,004)
26. VolCause(009,008)
27. VolCause(010,008)
28. VolCause(011,008)
29. VolCause(012,008)
30. VolCause(013,008)
31. VolCause(025,017)
32. VolCause(029,028)
33. VolCause(034,033)
34. VolCause(038,037)

35. VolCause(043,042)
36. VolCause(046,045)
37. VolCause(049,048)
38. VolCause(050,048)
39. VolCause(055,054)
40. VolCause(059,058)
41. VolCause(064,063)
42. VolCause(068,067)
43. VolCause(073,072)
44. VolCause(077,076)
45. VolCause(081,080)
46. VolCause(089,088)
47. VolCause(094,093)
48. VolCause(097,096)

E.6 Purpose

1. Purpose(005,004)
2. Purpose(009,008)
3. Purpose(010,008)
4. Purpose(011,008)
5. Purpose(012,008)
6. Purpose(013,008)
7. Purpose(025,017)
8. Purpose(029,028)
9. Purpose(034,033)
10. Purpose(038,037)
11. Purpose(043,042)
12. Purpose(046,045)
13. Purpose(049,048)

14. Purpose(050,048)
15. Purpose(055,054)
16. Purpose(059,058)
17. Purpose(064,063)
18. Purpose(068,067)
19. Purpose(073,072)
20. Purpose(077,076)
21. Purpose(081,080)
22. Purpose(089,088)
23. Purpose(094,093)
24. Purpose(097,096)

E.7 Condition

1. Condition(003,002)
2. Condition(004,003)
3. Condition(017,003)
4. Condition(018,017)
5. Condition(033,003)
6. Condition(042,003)
7. Condition(044,003)
8. Condition(054,053)
9. Condition(063,053)
10. Condition(072,053)
11. Condition(088,087)
12. Condition(093,087)

E.8 Equative

1. Equative(029,068)
2. Equative(058,068)
3. Equative(068,029)
4. Equative(068,058)

E.9 Contrast

1. Contrast(005,009)
2. Contrast(005,010)
3. Contrast(005,011)
4. Contrast(005,012)
5. Contrast(005,013)
6. Contrast(005,029)
7. Contrast(005,038)
8. Contrast(005,049)
9. Contrast(005,050)
10. Contrast(005,059)
11. Contrast(005,068)
12. Contrast(005,077)
13. Contrast(005,081)
14. Contrast(005,097)
15. Contrast(025,009)
16. Contrast(025,010)
17. Contrast(025,011)
18. Contrast(025,012)
19. Contrast(025,013)
20. Contrast(025,029)
21. Contrast(025,038)
22. Contrast(025,049)
23. Contrast(025,050)
24. Contrast(025,059)
25. Contrast(025,068)
26. Contrast(025,077)
27. Contrast(025,081)
28. Contrast(025,097)

29. Contrast(034,009)
30. Contrast(034,010)
31. Contrast(034,011)
32. Contrast(034,012)
33. Contrast(034,013)
34. Contrast(034,029)
35. Contrast(034,038)
36. Contrast(034,049)
37. Contrast(034,050)
38. Contrast(034,059)
39. Contrast(034,068)
40. Contrast(034,077)
41. Contrast(034,081)
42. Contrast(034,097)
43. Contrast(043,009)
44. Contrast(043,010)
45. Contrast(043,011)
46. Contrast(043,012)
47. Contrast(043,013)
48. Contrast(043,029)
49. Contrast(043,038)
50. Contrast(043,049)
51. Contrast(043,050)
52. Contrast(043,059)
53. Contrast(043,068)
54. Contrast(043,077)
55. Contrast(043,081)
56. Contrast(043,097)
57. Contrast(046,009)

58. Contrast(046,010)
59. Contrast(046,011)
60. Contrast(046,012)
61. Contrast(046,013)
62. Contrast(046,029)
63. Contrast(046,038)
64. Contrast(046,049)
65. Contrast(046,050)
66. Contrast(046,059)
67. Contrast(046,068)
68. Contrast(046,077)
69. Contrast(046,081)
70. Contrast(046,097)
71. Contrast(055,009)
72. Contrast(055,010)
73. Contrast(055,011)
74. Contrast(055,012)
75. Contrast(055,013)
76. Contrast(055,029)
77. Contrast(055,038)
78. Contrast(055,049)
79. Contrast(055,050)
80. Contrast(055,059)
81. Contrast(055,068)
82. Contrast(055,077)
83. Contrast(055,081)
84. Contrast(055,097)
85. Contrast(064,009)
86. Contrast(064,010)

87. Contrast(064,011)
88. Contrast(064,012)
89. Contrast(064,013)
90. Contrast(064,029)
91. Contrast(064,038)
92. Contrast(064,049)
93. Contrast(064,050)
94. Contrast(064,059)
95. Contrast(064,068)
96. Contrast(064,077)
97. Contrast(064,081)
98. Contrast(064,097)
99. Contrast(073,009)
100. Contrast(073,010)
101. Contrast(073,011)
102. Contrast(073,012)
103. Contrast(073,013)
104. Contrast(073,029)
105. Contrast(073,038)
106. Contrast(073,049)
107. Contrast(073,050)
108. Contrast(073,059)
109. Contrast(073,068)
110. Contrast(073,077)
111. Contrast(073,081)
112. Contrast(073,097)
113. Contrast(089,009)
114. Contrast(089,010)
115. Contrast(089,011)

116. Contrast(089,012)
117. Contrast(089,013)
118. Contrast(089,029)
119. Contrast(089,038)
120. Contrast(089,049)
121. Contrast(089,050)
122. Contrast(089,059)
123. Contrast(089,068)
124. Contrast(089,077)
125. Contrast(089,081)
126. Contrast(089,097)
127. Contrast(094,009)
128. Contrast(094,010)
129. Contrast(094,011)
130. Contrast(094,012)
131. Contrast(094,013)
132. Contrast(094,029)
133. Contrast(094,038)
134. Contrast(094,049)
135. Contrast(094,050)
136. Contrast(094,059)
137. Contrast(094,068)
138. Contrast(094,077)
139. Contrast(094,081)
140. Contrast(094,097)
141. Contrast(009,029)
142. Contrast(009,049)
143. Contrast(009,050)
144. Contrast(009,012)

145. Contrast(009,059)
146. Contrast(009,068)
147. Contrast(009,077)
148. Contrast(010,029)
149. Contrast(010,049)
150. Contrast(010,050)
151. Contrast(010,012)
152. Contrast(010,059)
153. Contrast(010,068)
154. Contrast(010,077)
155. Contrast(011,029)
156. Contrast(011,049)
157. Contrast(011,050)
158. Contrast(011,012)
159. Contrast(011,059)
160. Contrast(011,068)
161. Contrast(011,077)
162. Contrast(012,029)
163. Contrast(012,049)
164. Contrast(012,050)
165. Contrast(012,012)
166. Contrast(012,059)
167. Contrast(012,068)
168. Contrast(012,077)
169. Contrast(013,029)
170. Contrast(013,049)
171. Contrast(013,050)
172. Contrast(013,012)
173. Contrast(013,059)

174. Contrast(013,068)
175. Contrast(013,077)
176. Contrast(038,029)
177. Contrast(038,049)
178. Contrast(038,050)
179. Contrast(038,012)
180. Contrast(038,059)
181. Contrast(038,068)
182. Contrast(038,077)
183. Contrast(081,029)
184. Contrast(081,049)
185. Contrast(081,050)
186. Contrast(081,012)
187. Contrast(081,059)
188. Contrast(081,068)
189. Contrast(081,077)
190. Contrast(097,029)
191. Contrast(097,049)
192. Contrast(097,050)
193. Contrast(097,012)
194. Contrast(097,059)
195. Contrast(097,068)
196. Contrast(097,077)
197. Contrast(029,049)
198. Contrast(029,050)
199. Contrast(029,068)
200. Contrast(029,077)
201. Contrast(059,049)
202. Contrast(059,050)

203. Contrast(059,068)
204. Contrast(059,077)
205. Contrast(049,068)
206. Contrast(049,077)
207. Contrast(050,068)
208. Contrast(050,077)
209. Contrast(068,077)