UNIVERSITY OF SOUTHAMPTON

Semiotic Term Expansion as the Basis for Thematic Models in Narrative Systems.

by

Charlie Hargood

A thesis submitted in partial fulfillment for the degree of Doctor of Philosophy

in the
Faculty of Engineering and Applied Science
Department of Electronics and Computer Science

November 2011

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE
DEPARTMENT OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Charlie Hargood

Narratives are a method of communicating information that comes naturally to people and is present in much of our digital and non-digital lives. While work has been undertaken investigating the nature of plot and content within narrative systems little has been done to model subtext or themes. In this thesis a machine understandable thematic model is presented for representing themes within narrative. Each instance of this model forms a definition of a theme and how it may be deconstructed into other thematic elements and their related features. The model is based on semiotic term expansion where terms may be shown to *denote* motifs which in turn *connote* themes. An authoring method has been developed to allow for instances of the model to be created. The effectiveness of this approach is demonstrated in four experiments presented within this thesis centred around the concept of creating thematic definitions and generating thematically relevant images. The first experiment explored a semiotic term expansion method for creating thematic definitions in terms of the model and a guide to support authors in doing so. This demonstrated that, though further support for authors is needed, creating valid definitions of themes was possible using the method. The following two experiments used a system called the Thematic Montage Builder; a prototype using definitions of the model to create themed photo montages. The first of these experiments compares the ability of this system to generate montages relevant to specific titles containing themes to Flickr keyword searches while the second compares this system to a term expansion system based on co-occurrence. In both cases the TMB generates montages that are judged by participants to better represent the theme in question. In the final experiment the effect of thematic emphasis on narrative cohesion is investigated. In this experiment a set of variables for measuring narrative cohesion are identified and the impact of using themed illustrations from the TMB on short stories is measured. The illustrations reduced the thematic noise of the short stories and further analysis shows a correlation between thematic cohesion and the perceived 'logical sense' and 'genre cohesion' of the narratives. This work shows that better machine understandable models of narrative can benefit from an understanding of themes, and that semiotic term expansion may be used to build successful thematic models.

Contents

List of Figures

List of Tables

Academic Thesis: Declaration Of Authorship

I, Charlie Hargood

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

Semiotic Term Expansion as the Basis for Thematic Models in Narrative Systems

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as:

C. Hargood, D. Millard, and M. Weal. A thematic approach to emerging narrative structure. In Web Science at Hypertext '08, 2008.

C. Hargood, D. Millard, and M. Weal. Investigating a thematic approach to narrative generation. In DAH at Hypertext 09, 2009.

C. Hargood, D. Millard, and M. Weal. Using a thematic model to enrich photo montages. In Proceedings of the 20th ACM conference on Hypertext and hypermedia, pages 135{140. ACM, 2009.

C. Hargood, D. Millard, and M.Weal. Capturing the semiotic relationship between terms. New Rev. Hypermedia Multimedia, 16(1-2):71{84, 2010.

C. Hargood, D. Millard, and M. Weal. A semiotic approach for the generation of themed photo narratives. In HT '10: Proceedings of the 21st ACM conference on Hypertext and hypermedia, pages 19-28, New York, NY, USA, 2010. ACM.

C. Hargood, D. Millard, and M. Weal. Measuring narrative cohesion: A five variables approach. In Narrative and Hypertext at Hypertext '11, 2011.

Signed: .....

Date: .....

Acknowledgements

I would like to thank and acknowledge the following people:

My supervisors Dr David Millard and Dr Mark Weal for their outstanding support and assistance, particularly in the early part of this PhD. I hope we will continue to work together and make significant ground in this area.

My Examiners Prof. Paul Lewis and Dr Frank Nack for taking the time to understand and appraise my work.

Patrick McSweeney for his assistance in refining the code for the Thematic Model Builder and Thematic Illustrator.

Nathan Matias and Steve Ersinghaus for their assistance and cooperation with the narrative cohesion experiment.

The participants of my experiments, particularly those who took the time to construct thematic definitions for me.

The Learning Societies Lab and the school of Electronics and Computer Science at the University of Southampton for providing a research environment like no other. Particularly Alex Frazer, Rikki Prince, Sina Samangooei, and Yvonne Howard for their advice and council during the PhD.

All my family and friends who have supported me through the challenges and achievements of this work.

And my partner Jade, for being very patient.

Glossary

associated in the context of my thematic model refers to a thematic element that while connected to another element is not completely, in all cases, relevant to it and as such should not be considered a component of that element. 44

connote means to signify or imply beyond literal meaning. E.g. a red light connotes 'Danger'. 3

definition in the context of this thesis refers to a thematic definition meaning a representation of a given theme as sub themes and motifs within the terms of the thematic model. An instance of the thematic model. 3

denote means a literal representation of sign of. E.g. a picture of a chair denotes 'Chair'. 3

discourse in structuralism refers to the selection of narrative elements to expose from the story along with their relationships to each other. Occasionally also used to refer to the composed presentation of a narrative through a given media.. 11

fabula the russian formalist equivilent of story. 12

feature is identifiable evidence of the presence of a particular piece of content within a narrative or piece of narrative. This might be a piece of metadata, a tag, or automatically extracted keywords/descriptions. 39

genre cohesion in the context of my cohesion variables refers to the consistency with which the narrative conforms to the cultural conventions of an identifiable genre. 86

justification in the context of my thematic model refers to the type/reasoning of the connotation relationship shared between two elements. 41

logical sense in the context of my cohesion variables refers to a composite of the lexical, causal, and chronological cohesion of the narrative. 86

motif is an atomic component of a theme. Motifs are generalisations or classifications of elements with the narrative that connote a theme, they are directly denoted by features within the narrative. E.g. Snow, Rose, Grave, Champagne. 3

narrative a structured, purposeful, communication of an experience to an audience. 1

narrative cohesion the extent to which a narrative coherently and consistently flows and delivers its content. 6

narrative system a system principally concerned with the generation, analysis, or presentation of narrative. 2

narrator cohesion in the context of my cohesion variables refers to the consistency with which an identifiable storyteller is presented. 86

natom is a 'narrative atom'. A singular irreducible component of storytelling, where further reduction is either impossible or would cause the natom to no longer make sense. E.g. a photograph, a sentence or short paragraph, a short video clip. 39

sjuzhet the russian formalist equivilent of discourse. 12

story sometimes informally used interchangably with narrative, in structuralism refers to the set of all potential narrative elements (such as events, characters, locations) from which a discourse is constructed. 11

style cohesion in the context of my cohesion variables refers to the consistency with which language or other media is used to present different narrative elements. 86

subtext meaningful information, opinion, or perspective communicated to audience implictly such as through metaphor or themes as supposed to explicitly such as being overtly stated within the plot.. 3

thematic cohesion in the context of my cohesion variables refers to the consistency with which identifiable core themes are presented in a narrative. 86

thematic element in the context of the thematic model; a theme or motif. 101

thematic illustrator a variation of the TMB which generates themed illustrations for short stories. 6

thematic model a structural definition of elements and relationships coupled with rules for building definitions of themes within a narrative. 2

theme is a high level concept that is part of the subtext of a narrative. A theme is not directly present within the narrative but is connoted through the presence of motifs. E.g. Winter, Love, Death, Celebration. 1

TMB Thematic Montage Builder. A prototype system that utilises thematic definitions to generated themed photo montages. 4

Chapter 1

Introduction

In this thesis I explore the role of themes in constructing better machine readable models of narrative. To begin to understand the motivation behind the work discussed in this document we can consider two quotes:

"[Narrative] is simply there, like life itself."
-Roland Barthes

"Anything written in meaningful language has a theme; only intentionally meaningless works [created experimentally] have no theme.
To be coherent, a verbal structure must have a unifying theme running through it. Consequently, both the selection and the development of the theme are important aesthetically."
-Boris Tomashevsky

Both Barthes and Tomashevsky were structuralists prominent during the last century and saw importance of both theme and narrative. Narrative as a concept should not be limited to the idea of a novel, or a film, or what we would conventionally describe as a story. A narrative can be considered an ordered, composed, expression of human experience [84], fictional or not, and as such could be attached to many structured presentations of media and is not limited purely to traditional forms such as fictional novels or movies. It is better perhaps to think of narrative as a way of presenting information, whether to inform, persuade, or entertain, as a discourse rather than a specific type or set of types of media. Narrative is a very prevalent method of communicating information either face to face, through traditional media, or the Web.

A powerful tool of narrative is subtext, of which themes are a part. Subtext allows the author to communicate meaning implicitly, rather than as part of the narratives content, and to bind and flow the component parts of a narrative together with consistent concepts. In order for us to construct or understand narratives effectively we must also understand the notion of theme, how it becomes apparent within a narrative, and what effect it has, if any, on the experience of the consumer of the narrative.

This thesis will present a machine understandable representation of themes for use in narrative systems. I present a thematic model based on semiotic term expansion and present the results of a series of experiments evaluating both the performance of a system using the model to communicate themes, and the effect of this on the experience of consuming a narrative.

1.1 Semiotics, Thematics, and Narratives

Narratives, are potentially the most prevalent form of information communication between people. Research has shown [23][69][82] that people use stories to make sense of their own information and experiences, as well as to communicate them to others. Work has also shown that narratives are an important part of communication within communities [45] as well as the communication of knowledge between individuals [75]. A narrative's effect on the understanding of the information it presents has been used towards analytical goals. In social ethnography narrative analysis has been used to explore and interpret identity, such as the work done in analysing organisational identity by Brown [22] and Coupland [37], and in work by Alvarez [5] narrative is used to assist users in making sense of system requirement specifications. As well as to make sense of the flow of information between two people, narratives have been used, in perhaps their most familiar form, to entertain, in literature, cinema, the theatre, and computer games, as well as to convince in advertisements and debate. With the rise of collections of media on the web and the increasing problem of information overload [14] how we present and communicate information becomes an increasingly important question. This information is often in the form of narratives, or could be delivered to users as a narrative. In all of these areas increasing the ability of machines to understand narratives may lead to collections or individual narratives that can be better navigated, classified, presented, or even adapted and generated.

A 'narrative system' could be defined as any system or part of a system that is primarily concerned with narratives. This covers a broad range of systems from adaptive hypermedia, to narrative generators, and aspects of a large number of Web applications and services. Adaptive hypermedia and hypertext systems are often principally concerned with the delivery of personalised adaptive content to users from either user generated collections or predefined sources as detailed in chapter 2. This content is often in the form of a narrative or displayed as a narrative to engage users, and the adaptivity of the systems are reliant on understanding the narratives being delivered or created. Narrative generators (also detailed in chapter 2), on the other hand, are systems concerned with the creation of narratives based on user directed requirements. There are a variety of approaches to this problem, ranging from those that focus on characters present in the narrative, to those that simulate the process of authoring, but for any such approach, an understanding of how narrative is built is essential. Existing narrative systems ex-

perience mixed success, sometimes producing narratives lacking complexity, depth, or interesting plot, and often with problems with cohesion [91]. But the idea holds great potential and, should effective narrative generation be achieved, it would be possible to create a host of powerful narrative systems. As examples we might suggest: an automatic news reader which understood both events that had occurred and also what the user already knew and constructed a narrative of new developments, or a computer game which adapted and regenerated its plot based on the actions of the player, or a search and navigation system for a collection of information that presented results as an engaging discourse.

Existing research in narrative systems is largely concerned with content and plot: what occurs in the story and how it is structured. However, very little attention has been given to the notion of subtext (of which themes are a part), despite its importance to a rich and cohesive narrative. Themes in particular allow the author to give further meaning to their narratives. They allow the author to communicate subtle information without changing the focus of the narrative, and bind the components of a narrative together in a cohesive way under unifying concepts. Most narrative systems have no concept of what theme is, how it is present within a narrative, and how to interpret its features.

Within this research I have explored the creation of a machine understandable thematic model to represent themes within a narrative. By increasing a systems understanding of themes adaptive systems might better understand the content being adapted for the user, and narrative generators might create more coherent engaging narratives.

The thematic model takes a structuralist approach of breaking down the idea of a theme into its component elements. It represents a theme as a collection of other themes and motifs, which are directly denoted from features and, in the context of each other, connote the given theme. This allows a system to connect the occurrence of identifiable features with the presence of motifs and the potential presence of a theme. Definitions of themes in the terms of such a model can be used either analytically to determine the present themes of a narrative or generatively to understand what features can be used to embed a theme within a narrative.

The definitions of themes, motifs and their relationships can be very subjective. Themes are conceptual entities, implied through a narratives subtext rather than explicitly through the plots content. This implicit nature encouraged me to investigate a semiotic[1] approach to the creation of definitions. As there is no documentation of exactly what motifs make up given themes the semiotic approach allows us to capture the thematic definitions from subjects' individual perspectives of themes. The semiotic principle that the signifier of a connotative sign is a denotative sign becomes the basis for my thematic authoring method and, by formalising this in a five stage process, authors are able to

[1] a discipline devoted to understanding how signs are interpreted

create valid definitions of themes in the terms of my model. This process was condensed into a guide that potentially any author could pick up and use to create definitions. Creating such a formal process out of a task so subjective raised many potential problems and the guide was subject to repeated review, and a series of small experiments analyses its performance. While using the process did lead to valid models, the experiments show there is more work that could be done to support definition authors.

These definitions were utilised by the TMB (Thematic Montage Builder). Using the model the TMB was able to take titles for photo montages, that contained desired themes as well as defined content, and construct a photo montage from images in Flickr that was both relevant to the desired content as well as the desired themes, an example screenshot of this can be seen in 1.1. The TMB built a corpus using the content keyword and then calculated the thematic quality of each item based on the presence of desired thematic features within their tags. In essence the montages produced by the TMB were search results presented as a narrative, and the titles a text based query. The semiotic term expansion performed at the authoring stage of the thematic definitions was being utilised in order to improve the relevance to queries with a thematic nature.

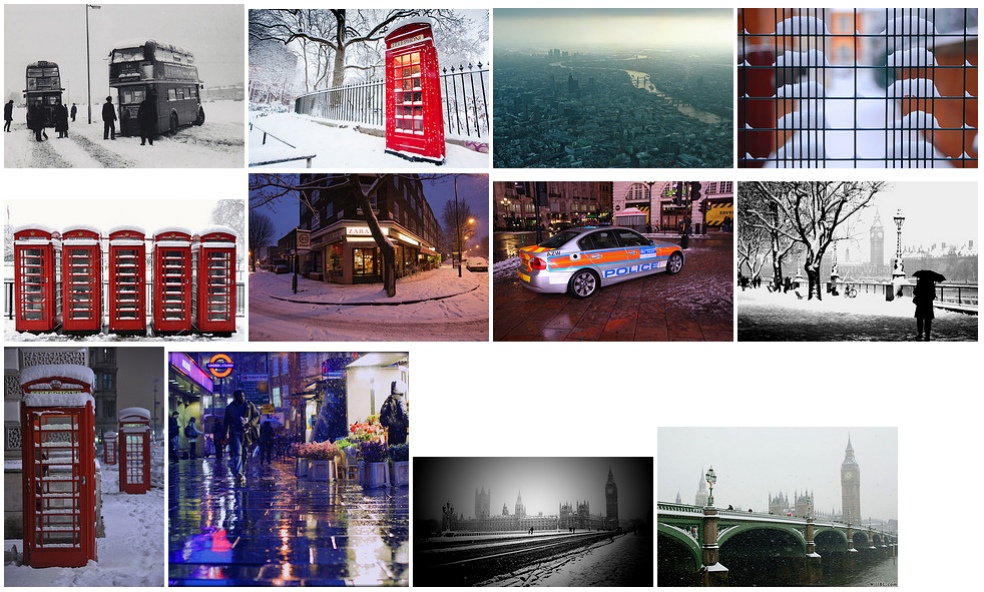


FIGURE 1.1: An example screen shot of the TMB constructing a montage on the theme of 'Winter' using images of London.

Term expansion has long been used in IR(Information Retrieval) research to improve the relevance of text search results. By expanding the terms used in a query to semantically similar ones the systems increase the chance of identifying relevant documents. This is done by identifying a relationship between the terms and, depending on the strength of this relationship, expanding or weighting the term in the query. Relationships commonly used are synonyms and meronyms from thesauri/lexical databases and detected co-occurrence within a corpus.

Our thematic definitions are in essence a term expansion based on semiotic relationships.

Themes and connotations are important to the way people use language including their use of language on the Web. When a term is selected, even outside the context of a narrative, such as for a query or a tag it could be selected with thematic connotations in mind. Take for example the image below in figure 1.2 [2].

FIGURE 1.2: An example of how themes may be implicit in selected tags. Tags are connected to motifs which are in turn connected to themes.

As we explore in figure 1.2 from a set of tags motifs may be drawn that in turn connote themes for a more complex understanding than what the picture literally presents. It is possible that the tags used for this image were selected with a thematic purpose towards these themes, that they were deliberately implying them. This thematic purpose (the intended implication) is currently something that could be used to improve the relevance of systems interacting with the language used to express this purpose. By expanding terms semiotically a system may have some understanding of the thematic content of the language used, this might allow search results to find thematically relevant items or have systems categorise items thematically.

The idea of thematic queries, as well as the effectiveness of thematic definitions for creating themed photo montages, led to my first couple of experiments, where I explored the ability of the TMB to react to different montage titles with thematic content and compared its performance to keyword search systems and co-occurrence term expansion. These experiments were designed to evaluate the performance of a system using the thematic definitions (in this case the TMB), explore the effect of presenting themes in individual items as well as narratives (in this case montages), the effect on the thematic system of using multiple themes and pairing themes with a difficult corpus, and investigate whether co-occurrence could itself simulate semiotics. If co-occurrence was just as effective at presenting themes then there would be no need for the lengthy process of constructing thematic definitions and a faster automatic approach could be used to expand the theme in question.

[2]Image taken from www.flickr.com, user findfando.

As well as understanding the implied meaning of an author in the language they use, themes serve another important purpose. One of the reasons themes are considered integral in literary theory is that they bring the components of a narrative together into a single flowing form. In his work Thematics [113] Tomashevsky claims it is the theme that "unites the separate elements". To explore this effect requires an understanding of narrative cohesion: how the message and content of a narrative are delivered to an audience in a way that makes coherent sense. If themes are a part of cohesion then how does their presence impact on the audience's perception of the narrative's cohesion as a whole? Does thematic cohesion itself genuinely have a positive effect on the narrative experience? To answer these questions an experiment was designed that used the TMB as a basis to create a thematic illustrator that would generate themed illustrations for short stories, based on their content, that emphasised an established theme within the story. This was to be compared to a similar system that was based on what was occurring within the narrative but not on the themes. To perform the comparison a number of variables by which cohesion could be measured were identified and used as the basis for a user evaluation.

Machine understandable narrative is a potentially powerful tool in how we present and manage information within systems, including the Web. Narrative has been a long prevalent form of presenting information between people and, as we increasingly digitise our information and the need for digital narratives increases, so does the importance of machine understandable models for describing them. Themes are an important part of narrative and, while their composition and effect has been explored in literary theory, there is currently little work exploring how to model them in a computational sense. Themes are complex entities, subjective in their definitions and subtle in their effect, but through semiotics there may be a way to capture the structure in which they exist in narrative. In this thesis I present a thematic model for representing themes in narratives. This includes a semiotic term expansion basis for authoring themes, the structure of the model, and evaluatory experiments for how systems using these definitions perform at creating thematic photo montages and illustrating short stories for perceived narrative cohesion.

1.2 Research Hypothesis

The research within this thesis examines the following hypothesis:

Thematic Models generated through Semiotic Term Expansion can be used to improve the relevance of search results in the context of a defined theme, and to improve the perceived cohesion of narratives through thematic illustration.

The following three key questions have been identified:

1. Can semiotic term expansion be used to generate thematic models?
2. Can thematic models be used to improve the relevance of search results in the context of a defined theme?
3. Can thematic models be used to improve the perceived cohesion of narratives through thematic illustration?

Question one explores how definitions of themes are created in the terms of the model, by "thematic models" here I refer to instances of a structure defining the presence of themes within a narrative, and by "semiotic term expansion" I mean term expansion using denotation and connotation as the semantic relationship between terms used for expansion. In order to answer this question I will need to first investigate how a machine understandable structure defining themes could be created based on existing work on thematic structures. To generate models in the terms of this structure semiotically I would also have to investigate how connotative and denotative relationships could be captured. As something very subjective that can be influenced on both a personal and cultural level an adequately formal method for capturing these relationships to form valid models that are semantically rich will need to be created. Finally it would also remain to be investigated whether semiotic term expansion is actually the best method for generating models within this structure and if another form of term expansion could be used to build thematic definitions in a way that performs better or at a lower cost.

Question two focuses on how some queries may be improved by machine understanding of thematics and semiotics. By performing semiotic expansion on terms used in queries or tags a search system might capture any implication intended from the user when those terms were used for finding relevant content that could otherwise be missed. By "in the context of a defined theme" I mean a search that is seeking content relevant to a theme that has been defined in the terms of the model. Searches with such a thematic content may have the relevance of their results improved by such an expansion. An experiment was conducted comparing the results of a thematic system to those of existing search and term expansion systems with users rating the relevance of the results presented by each. This user rating of relevance to given queries would be the metric I used to ascertain whether the system had "improved the relevance".

Question three concerns the tangible effect thematic understanding might have on narrative generation or narratives that are re-presented. Identifiable consistent thematic content has been attributed with binding a narrative together coherently. However it remains to be seen if the thematic models created by semiotic expansion are capable of this or what the tangible effect of enhanced thematic cohesion is on the cohesion of a story as a whole. By "thematic illustration" I mean images identified by a system as containing features that connote a given theme. By adding these illustrations to a story to emphasise a given theme present within the story a system may be able to strengthen

the thematic cohesion of the story and make the narrative as a whole more cohesive. This requires the identification of recordable variables pertaining to the cohesion of a story and an experiment that can record user perceptions of these variables. I can then compare the users perceptions of these variables for stories with illustrations selected based only of the stories content and for illustrations selected based on both content and theme. This will allow us to ascertain whether perceived thematic cohesion has been improved by the system and also what the effect of this is on the cohesion of the narrative as a whole.

1.3 Document Overview

This thesis is divided into seven chapters and within this section I will provide a summary of the content of each and a map of the document.

In this chapter I have established context and motivation, a dissemination of the hypothesis for this research including the research questions to be answered and how I intend to approach these questions, and this document overview.

Chapter 2 describes the background research and literature relevant to this research including structuralist and narratology theory, an exploration of existing narrative systems including adaptive hypermedia and narrative generation, and research undertaken in term expansion.

Chapter 3 describes the structure of the novel thematic model. This includes an explanation of the elements and relationships that make up my model of themes and the rules that govern valid definitions made within the terms of the model. This chapter also includes an exploration of how semiotic term expansion is captured and how authoring thematic definitions might occur. This process is captured as a series of steps and is then evaluated with an experiment that tests the ability of untrained individuals at creating thematic definitions using the defined process and the validity of the definitions they create. Finally this chapter also includes an exploration of how the thematic model might be integrated with existing methods of narrative generation.

Chapter 4 describes the first of two experiments on the generation of themed photo montages. The experiment compares the performance of the TMB, a prototype using definitions made in terms of the model, to Flickr's own keyword search in generating themed photo montages.

Chapter 5 describes a similar experiment where the TMB is compared to a system based on co-occurrence. My implementation of co-occurrence is explained as well as how the experiment was restructured to allow this different comparison. The experiment investigates whether co-occurrence could be used to simulate semiotics.

Chapter 6 documents a final experiment where the effect of using a thematic system on the narrative cohesion of short stories is presented. The chapter outlines the extraction of key variables related to narrative cohesion based on literature and existing systems and the development of a user based evaluation for capturing audience perception of cohesion based on these variables. The experiment compares the experience of reading short stories where key themes have been emphasised through thematic illustrations with those with illustrations based just on content.

Chapter 7 concludes this thesis, summarises the work, and outlines the key contributions. A description of possible future work explaining how this research can be taken forward is also undertaken and broken down into potential projects. Finally the work is concluded with a discussion of the key discoveries of this PhD.

Chapter 2

Background Literature

This chapter presents the theoretical basis of this work, and looks at how others have covered the challenge of creating narrative systems. The chapter covers the narratological theories that act as a basis for this contribution in the form of the thematic model, as well as research into narrative systems more generally.

2.1 Narratology

By "Narratology" I mean the theory of narrative that arises from literary theory, criticism, and philosophy. Narratology has much to contribute in forming better machine-understandable models of narrative.

2.1.1 Structuralism

Structuralism is a philosophy concerned with identifying structures emergent through language. As a philosophy it has been applied to anthropology, such as the identification of phonemes in cultural behaviour by Levi-Strauss [74] and to literary theory and semiotics in the works of Barthes [12] and Saussure [102]. Structuralism in literary theory was heavily influenced by earlier work in formalism, investigating patterns and re-occurring forms in narrative. This included the work of Propp on folk tales [97], where he identified a sequence of 31 functions that comprise a folk tale (ranging from 'absentation' through 'struggle' and to 'wedding') and 8 core character archetypes (such as 'villain' or 'hero'). His analysis and deconstruction of the component elements within narrative and the structures that control them demonstrates what could be called early structuralist concepts.

Structuralist analysis of narrative is based on structures comprised of a series of abstract elements and their relationships with each other, while the narrative itself is a represen-

tation of this structure modelled with language. Often the method of deconstruction of these structures is semiotic (described later in this chapter) representing the narrative as a series of signs. Structuralism has been criticised for its rigidity [107], in that narratives do not always conform to a given explicit structure. It was philosophically followed by post-structuralism (such as the works of Derrida [38]) which argued that narrative was inherently more complex than the models offered by structuralism and favoured a less determinate theory of language. However from the perspective of this research (which requires machine readable structures, that we accept are necessary simplifications of something much richer) the discrete rules, elements, and relationships that structuralism offers are useful when beginning to build better machine understandable models of narrative.

It has been asserted in structuralist works that narrative is a composed series of human experiences [84], this allows not just written novels to be perceived as narratives but all forms of composed communication from a human perspective. However the perspectives of what comprises the elements within a narrative differs in a variety of works. A classic distinction is between what is told in the narrative and how it is told, these were identified respectively by Russian Formalists as the '*Fabula*' and the '*Sjuzhet*'. This was adapted by French structuralists, particularly Roland Barthes [13], as '*Histoire*' and '*Discours*', which in turn is interpreted in English structuralism as '*Story*' and '*Discourse*'. Terminology aside, they all allude to the same basic structure of the story which represents a collection of all the information that could be covered by the narrative and the discourse which represents the selections from the story of what is exposed and how it is exposed within the narrative, as demonstrated in figure 2.1.

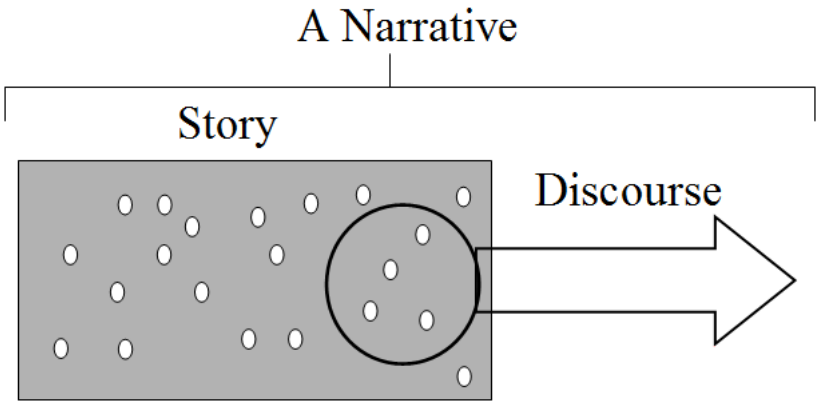


FIGURE 2.1: A narrative can be deconstructed into story and discourse

Discourse is a complex entity; a process by which elements of the story are selected for exposure, form relationships with each other that could be causal, chronological, or interactive, and are then presented to the audience. The discourse is what represents

the telling of the story. As well as the selection and ordering of content, what could be called the narratives plot, the discourse includes other subtler elements of the narrative that are compiled in the telling, such as any subtext and the themes of the narrative. Structuralist studies of discourse will also discuss how the author, as an important structure within the narrative itself, is a part of the narrative's discourse. In his work on Rhetoric [18] Booth talked of the importance of the authorial voice to narratives. The authorial voice was the storyteller that manifested either dramatically as a storyteller character, explicitly as a narrator, or implicitly in the style and subtle bias of the telling of the story. Each different authorial voice could be shown to have marked effects on the narrative, building a relationship between the author and the audience.

As well as the Story - Discourse model of narrative there are other structuralist theories that analyse narrative differently. Most notable perhaps is Bal's [11] who describes narrative as comprised of three layers; '*Fabula*', '*Story*', and '*Narrative*' which represent the sum of all elements in the narrative, the elements selected for exposition, and the way the selected elements are exposed respectively. Ultimately though, although Bal's model differs slightly from Barthes they are identifying the same overall narrative process, and what Bal refers to as fabula could easily be story in more classical structuralism, and what he refers to as story and narrative is broadly equivalent to discourse. Other structuralist models of narrative are more concerned specifically with the plot; the sequencing of events within the discourse and the formation of relationships between story elements, such as in Todorov's work on prose [112]. Todorov's laws of equilibrium asserted that plots were often centred around the notion of a stable state (an equilibrium) that becomes destabilised by an agent of change, before culminating in a new equilibrium. Todorov concluded in his structuralist analysis centred on stories by Boccaccio that there were two master plots: 'conversion' (disequilibrium becoming an equilibrium) or 'punishment evaded' (equilibrium to disequilibrium to equilibrium). This work was further commented on by Bruner [23], in his work on narrative constructions of reality, where amongst 9 other features, he identified 'Canonicity and breech' as a defining narrative feature. Like Todorov Bruner felt a narrative was not just a series of experiences, but needed a motivation, a point to the telling of the story, and this often took the form of 'the unusual' by breaking equilibrium.

How fabula becomes sjuzhet (or story becomes discourse) through the process of both authorship and consumption has been explored in literary theory through the notion of plot selection. As demonstrated in the Barthesian model of narrative, the conventional view is that the author selects story elements from the fabula to be a part of the plot. This concept was further explored by Musarra-Schroeder, based on Calvino's writings [92] as 'The Garbage Axiom' representing the process of the author deliberately omitting potential story elements in their construction of a plot grounded within the reality of the narrative. Calvi's work 'Lector in Rebus' [29] seeks to explore the role of the reader, rather than the author, in the formation of discourse in hypertext and follows on from

less specific work by Eco [39] on the role of the reader in narrative. Calvi explores how as well as the selections made by the author through 'The Garbage Axiom' the choices of the reader are also responsible for the discourse, formed through their decisions of how to consume the narrative driven by their own motivations. Calvi reports [29] how originally Calvino [30] refers to this choice as the 'Labyrinth Challenge', where the reader either consumes a narrative to get to the end (Calvino's first-order model reader) or to discover something in particular (Calvino's second-order model reader). Either choice leads to a different discourse consumed as the reader omits or gives less attention to plot elements less key to their reading strategy, leading to the reader affecting the eventual discourse created through their own selection and omission of consumed plot elements. Eco [40] (as reported by Calvi [29]) further explores this, explaining how the 'Labyrinth Challenge' is also in part the reader attempting to discover the discourse the author intended them to consume through the selections made by the author. However, Calvi explains that such selections and omissions made by both author and audience are often precluded by analepsis (explanation of past plot elements) and prolepsis (anticipation of elements to occur). Thus the inclusion (authorially or through attention) of one plot element leads to the inevitable inclusion of other related elements, through explanation or climax.

2.1.2 Semiotics

Semiotics, or semiology, is the study of signs and how we extract meaning from them. Sassaure [102] wrote that all signs are made up of a signifier and a signified; something we are observing and our understanding of it. This literal interpretation is that of denotation; we see a specific football and to us this denotes the concept of a ball. Barthes [12] expanded on this by describing the idea of connotation, that signs have a meaning beyond their literal expression, he wrote that the entire denotative sign becomes a signifier for a further signified; we may connote from the ball the concept of competition. Semiotics goes some way to explain a connection between language and more than just what it might be literally describing, and to explain that when a term is used, such as in a tag, the user is trying to express more than just its denoted meaning. The example above of how a denotation can lead to something other than its literal meaning is displayed diagrammatically in figure 2.2.

Structurally speaking this defines two key relationships between entities within a narrative; that of *connotation* and *denotation*. Conceptually this also divides what might originally have been thought of as a single part of the narrative into two things what the audience sees (the literal denotation) and what the audience understands (the connotation inferred from what they are presented).

In his work 'Cybertext' [1] Aarseth offers a more contemporary structuralist viewpoint building on the semiotic division of narrative elements. He describes that literary works

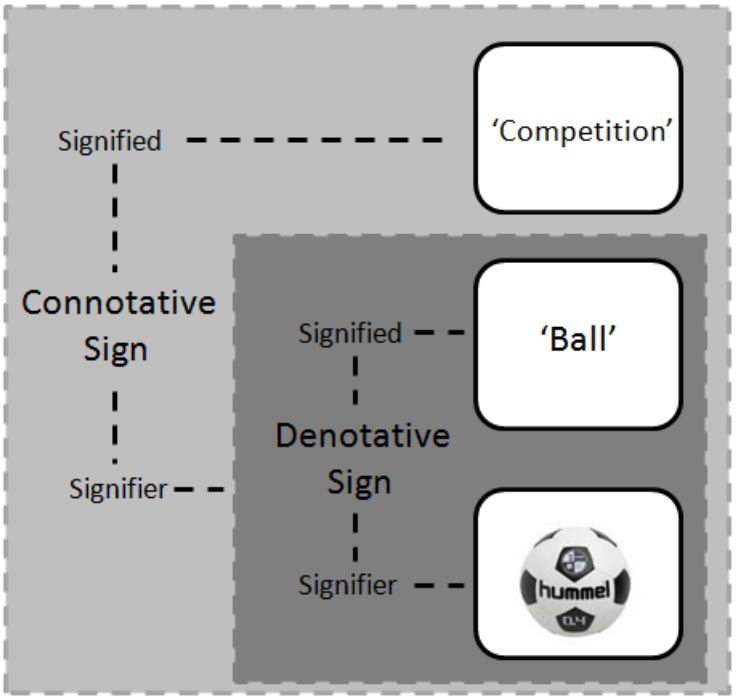


FIGURE 2.2: Barthes theory of how a denotative sign forms a connotative one

are built of *textons* and *scriptons*; the textons representing the textual elements themselves (a 'string of signs' as Aarseth describes it) and the scriptons the readers understanding and interpretation of them, leading to a model of a series of signifiers connected in turn to a series of signifieds. This would mean that a text of ambiguity may well have a much larger number of scriptons then it does textons as the list of possible interpretations lengthens. Aarseth describes how with the rise of modern forms of literature the connection between textons and scriptons has become less fixed. For example in a game while the number of textons is fixed the number of scriptons changes depending on the actions of the player causing which textons they are exposed to, and non-deterministic or generative texts may have a varying number of textons as well as scriptons. Based on his theory of textons and scriptons Aarseth explains how any narrative may be expressed through the values of seven identified variables; Dynamics, Determinability, Transiency, Perspective, Access, Linking, and User Function, each based on how a text is designed for an audience to traverse its textons and scriptons. While his work is not concerned explicitly with the plot and content of the narrative Aarseth's model seeks to express the structure of the audiences understanding of the narrative from the way it is presented. This is done in a way that can be applied to more modern and less conventional ways of presenting narratives and highlighting the importance of what is connoted from the elements presented by the author to the structure of narrative.

2.1.3 Thematics

Thematics can be described as a formalist approach to the concept of themes within narratives, it is also the title of the prominent work in the area by Tomashevsky [113]. Tomashevsky deconstructs thematic elements into themes (broad ideas such as 'politics' or 'drama') and motifs (more atomic elements directly related to the narrative such as 'the helpful beast' or 'the thespian'). A motif is the smallest atomic thematic element and refers to an individual element within the narrative which connotes in some way the theme. Themes may always be deconstructed into other themes or motifs whereas a motif may not be deconstructed.

Tomashevsky built up a semiotic model of themes in narrative where the denoted literal meaning was connected to the connoted implied understanding. Though Tomashevskys motifs were not as simple as a conventional denotative sign, being more narrative constructs which he explained were drawn from a text's genre, they were still deliberate signs from which the audience inferred further meaning in the form of themes. Crucially themes could be sub themes and as such lead to further themes. This builds to a hierarchy with the specific motif at the bottom denoted by the text and a tree of connotations of further themes above.

Tomashevsky further believed that there was an intrinsic connection between theme and genre. He explains that the motifs that comprised each theme were all of a list of repeating motifs available to a given genre, and that within this genre these motifs were often reused in a range of narratives and often for different themes. This might explain why some themes and genres are often intrinsicly connected, such as the theme of *supernatural* and the genre of *horror*, or the theme of *love* and the genre of *drama* or *romance*. It is possible, given Tomashevsky's model, that the set of motifs for these genres are predisposed to connoting particular themes. Further more, it is possible that this assits in a theme's role in narrative cohesion, where a strongly defined coherent theme will include motifs important to a genres conventions, and as such improve the genre cohesion of the narrative. The relationship between theme and genre, and how they impact cohesion, is something we explore in more detail in chapter 6.

Tomashevsky believed that themes were at the root of giving a narrative meaning and cohesion. Through themes an author could give a story purpose allowing for the author to entertain, convince, or otherwise communicate by presenting a perspective rather than a report of events.

2.2 Narrative Systems

While many systems are concerned with the processing and/or presentation of information in some way we might consider those that are primarily concerned with information

as a narrative to be narrative systems. As discussed earlier in this chapter narrative could be considered to be a composed series of human experiences [84], that has a purpose or out of the ordinary quality to give its telling purpose [23]. This means that narrative systems are not limited to written prose but a variety of representations of human experience. This is something demonstrated in the StoryBank project [45] where mobile digital storytelling is explored as a means to aid communication in rural communities in India. There are important affordances of multimedia narratives as opposed to textual ones, particularly in the case of this project when dealing with accessibility to a semi-literate audience, and these are instrumental to the success of the project.

This can also be seen in other projects such as PhotoCopain [115] which uses narrative as a lens through which to view large collections of images by offering automatic contextual annotations of photo collections. The information becomes a narrative through its purposeful presentation, this narration of the information by a system or a system concerned with assisting or presenting the process is what we would call a narrative system. Similar work can be seen work by Kuchinsky et al on the FotoFile project [68] where narrative structure is used as an aid to media annotation. Kuchinsky explained that while annotation is tedious humans enjoy storytelling and make sense of their experiences using narrative. Their users attach images to 'narrative episodes' rather then tagging them as a method of gathering image metadata.

Functionally speaking a narrative system may be focussed on dissemination, analysis, or creation. Systems such as StoryBank [45] play no part in the creation of narrative but assist in the sharing and distribution of narratives. Work by Chau [35] is an example of a system focussed on narrative analysis using established methods of neural network text analysis to make semantic sense of narratives generated by police reports. The OntoMedia project [65] also is not concerned with creation but the building of a shared vocabulary for describing narrative both to aid search and navigation and to provide semantic machine readable annotations. This vocabulary takes the form of an ontology modelling the components of a narrative that might be annotated with 'events' (actions and scenes within the plot) or 'entities' (characters, props, locations, and other key elements within the story). Much of the formalisation of OntoMedias ontology can be found in Lawrence's work [70]. In other cases the presentation of narratives can be a part of learning and assessment, either where narrative can be used to assist a student to make sense of a scenario or context [106], or to provide a framework for disseminating personalised materials in a centre of learning (such as a museum) in a way that engages the learner [62].

Narrative systems may however also be directly involved in narrative creation. These may be tools for the support of human created narratives such as the NICE project [100] where children are encouraged to interact in virtual worlds constructing their own narratives as a support for the learning process. However some systems take a more active role in narrative creation such as narrative generators, putting the system in

the role of the author in actually creating the narrative. But there are also narrative systems elsewhere concerned with the adaptation or presentation of narratives. Adaptive Hypermedia and Hypertext systems in particular are often concerned with the purposeful presentation of a series of connected information to users.

2.2.1 Adaptive Hypermedia and Hypertext

Through linking and adaptive structure Hypermedia and Hypertext systems present collections of information and multimedia as narratives. The vast majority of these systems are dynamic and as the user often controls which links are followed they are in a sense partially controlling the flow of the narrative; co-authoring it with the system. Hypermedias such as HyperCafe [103] present a series of possible narratives for the users to follow which they do through a series of different links. In the case of HyperCafe the user is placed amidst a fictional cafe where they see a multimedia presentation of the conversations taking place between characters. Using different links the user can control which conversations they are exposed to and as such what narrative they ultimately consume. The individual narratives themselves are created by human authors but the system facilitates the creation of the structure based on the decisions of the audience.

Narrative can be used to make the delivery of information more engaging to users. By taking information a user needs and presenting it as a discourse a system can make the information more engaging by embedding further meaning through the way the narrative is presented. An example of this is the Topia project [4] where search results are presented as a narrative to users using a variety of techniques in order to create a discourse. These include sequencing the results into what could be considered a plot and then emphasising and/or omitting particular elements that are either important to the user or damaging to the cohesion of the discourse.

Adaptive hypermedia ultimately aims to make static content dynamic in a way that better suits the objectives of the system and work in the area explores a variety of different ways to make content adaptable [20]. The way in which this adaptivity occurs varies on a system by system basis depending on the objectives of the system or the desires of its users. A system such as the Hyperdoc [87] handles multimedia narratives and adapts them at different layers of narrative to effect a presentation suited to the users desires. The hyperdoc follows Bal's model of narrative [11], as mentioned earlier and modifies both the content and the style in which videos such as interviews are shown to an audience. Similarly the AUTEUR system by Nack [94] provides an automated video editing system for delivery of completed themed videos based on a precompiled and time annotated collection of clips. AUTEUR's themes are more akin to narrative styles or genres (such as 'humor' in the prototype demonstrated) rather than connotated concepts as later discussed in my own investigation of themes. AUTEUR utilises an expert defined knowledge base that dictates both the rules for its themes and a semantic network of how

different actions within the scenes may achieve different effects or lead into other actions. AUTEUR then uses a two stage process where a 'scene planner' builds a strategy or plan for a series of scenes based on the user defined desired theme and motive for the piece. This is followed by a 'story planner' that works at the more detailed level of selecting clips with listed features from the database that match semantically desired actions to build the planned sections. AUTEUR represents a sophisticated exploration of what is possible in automatic narrative presentation. However, the success of such an approach could be largely dependent both on the quality of the expert defined knowledge base, and the pre-prepared, purpose made, and purposefully annotated recordings, on which it relies to generate its material, both of which are time consuming and subjective. In subsequent work [93] Nack has called for more semiotic aware hypermedia systems that understands the denotative and connotative connections present within their narratives in this particular case to overcome the analytical problem of interpretation of complex visual scenes. Nack proposes an architecture in which an expert defined knowledge base supports a semantic network of semiotic connections so that both denotative and connotative annotations may be applied to multimedia to lend a greater understanding of its content. The system proposed first of all highlights the importance of interpretation when dealing with rich multimedia narratives (and as such, the importance of building workable semiotic models) but also, in its reliance again on expert defined knowledge bases, the importance of understanding how subjective information (such as semiotic connotation) can be captured.

A different way to provide adaptation based on the user is user modelling. Rather than providing results based on user requests these systems seek to understand the characteristics of the user operating the system based on user meta data either entered by the user or based on system usage history. AHA! [19] uses user modelling to adapt the presentation of materials to users, normally course materials to students. The modelling allows systems using AHA! to understand what material the student has already covered and as such present appropriate material. AHA! also adapts its material based on the links the user follows and utilises a range of adaptive techniques some of which are automatic and others driven by the users leading to the presentation of content that is partly dictated by the author, partly by the user, and partly by what the system has inferred about the author. There are other methods adaptive hypermedia systems employ in order to provide dynamic content to their users many of which are detailed in Brusilovskys taxonomy [25] as detailed in figure 2.3. There are also other similar examples to AHA! where educational material is delivered through adaptive narratives (sometimes referred to as hyper-books [91]), another significant example being Interbook [24], which offers a simpler set of tools for supporting the authoring of adaptive hypermedia and serving it. Murray et al. in their work on MetaLinks [91] explain that many adaptive works, while delivering personalised content, can damage the cohesion of the narratives they present. He identifies hyper-book narratives as suffering from problems with 'Disorientation' (users lost in links), 'Cognitive Overload' (confusion of

choice in multipath hypertext), 'Discontinuous Flow' (the coherence of the work which he further separates into 'Narrative Flow' and 'Conceptual Flow'), and 'Content Readiness'(the appropriateness of the material for the given user). Murray et al. identify a series of potential solutions to these issues including contents pages (for disorientation), glossaries (for cognitive overload), and narrative smoothing through constructed introductions (for cohesion) experiencing some success but also identifying the area as needing further attention.

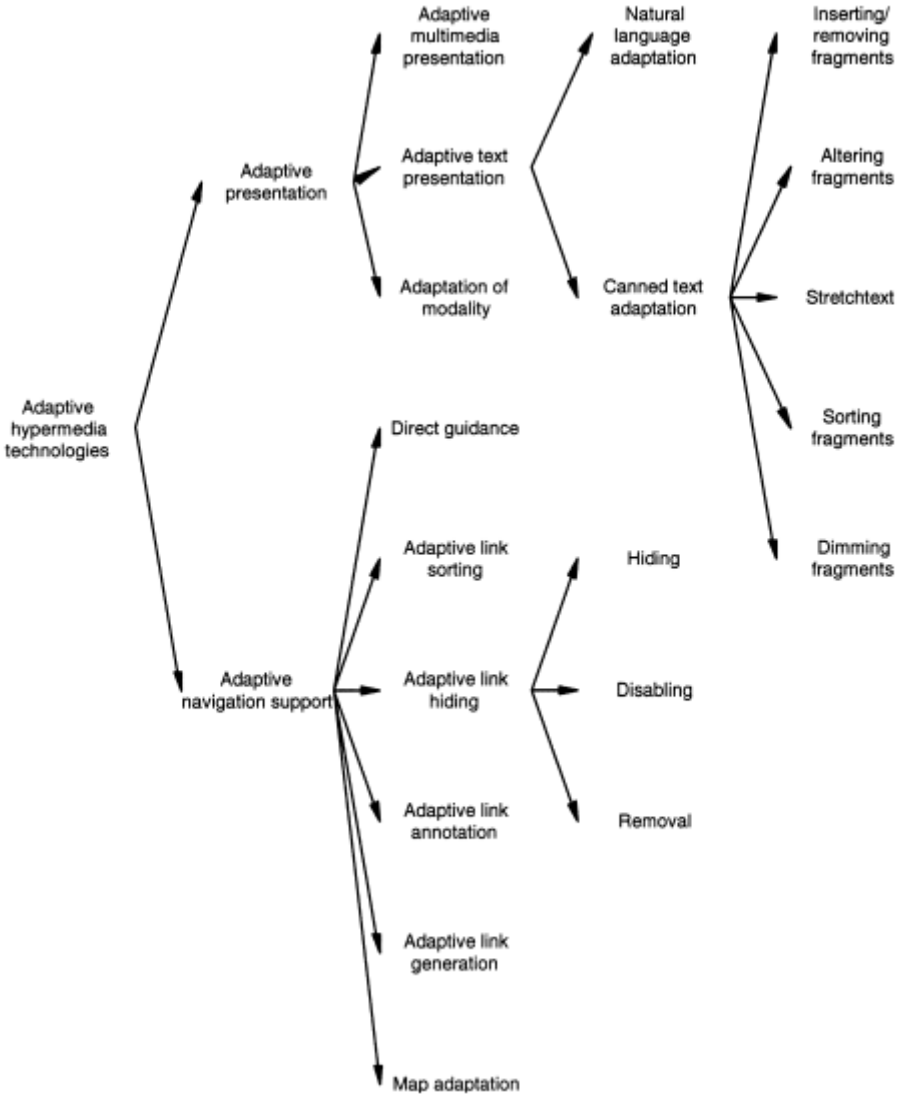


FIGURE 2.3: Brusilovskys taxonomy of adaptive hypermedia [25]

Through adaptation the system itself is making decisions based on its own rules or templates about the structure and presentation of content. This makes the system itself a co-author of the materials the user ultimately consumes as it becomes an implicit narrator through is adaptation. Adaptive hypermedia systems are narrative generators in the sense that they create narratives based on the desires of the user. While the content itself is not generated, the structure, sequencing, inclusion, and emphasis are

and the rules for adaptation collectively model the process of an author creating personalised material for their audience. Both AHA! and the Hyperdoc, and other adaptive hypermedia systems, could be considered in some ways to be *Author Centric* narrative generation systems which I describe in a later section.

2.2.2 Interactive Narrative

A common form of narrative system is the interactive narrative or drama. These are however normally explorations of writing rather than technology with simplistic systems delivery a story in which the user has control of the path of the story sometimes with the reader playing the role of the protagonist (much like a simple roleplaying game or a 'choose your own adventure' book). This can include work in Hypertext fiction (such as Afternoon by Joyce[1]) and Hypermedia such as the previously discussed HyperCafe [103]. These works come with a prewritten narrative in the form of a set of possible narratives represented by narrative fragments linked together in different sequences, the narrative is interactive in that the user decides his or her route through the links to form the eventual narrative.

As well as being applied for entertainment, interactive narratives have been used in other areas such as learning and assessment where the ability of narrative to engage readers and understand events has been used with virtual scenarios [77]. Interactive narratives have been used in this way in the Umpire and Management systems as discussed by Skov in [106] as part of his exploration into how the object orientated design process affects the creation of interactive narrative systems. His findings showed that while the approach facilitated a better design understanding of the purpose of such systems it as a formal process struggled to integrate with the creative and inherently unstructured process of writing stories for use in the system. This conflict between the structure of systems and the unstructured creativity of narrative is a common criticism of a structuralist approach and a typical and crucial problem facing narrative systems, particularly where the system itself (based on structure) becomes a co-author such as in narrative generators.

Most of these systems could not be considered narrative generation as no inferred decision made by the system itself affects the outcome of the narrative, they are just slaves to the decisions and instructions of the author and reader. In some interactive narratives this distinction blurs however, where as well as the user's interactions the systems models and simulates the personalities of characters other than the player to dictate their actions and the resulting effect on the narrative such as in Facade [80] or work undertaken by El-Nasr [42] both of which could be considered narrative generators and are discussed in more detail later. Another example of such a case is the system IDtension created by Szilas [109] which presents an interactive narrative approach that utilises narrative generation to solve the problem of interactivity being limited by predetermined

[1] http://www.eastgate.com/catalog/Afternoon.html as of 11/5/2011

plots and is again explored in more depth later when I discuss narrative generation. Similarly work by Charles [34] explores interactive generated narratives, in this case for safety, as part of the CALLAS project. Charles system uses voice recognition to allow users to intervene in a continually generated narrative highlighting kitchen dangers, this creates an interactive feedback loop allowing for interactivity in a completely generated narrative.

Interactive narratives sometimes encounter difficulty in both providing interactive creativity and cohesive narrative. This problem is one of essentially trying to merge purposefully created structured plot and unstructured chaotic user interactions. This has been referred to in the work of Aylett and Louchart on narrative in virtual environments [7] [76] as the 'narrative paradox', that interaction and narrative cohesion are at odds with each other. Aylett and Louchart created various virtual environment interactive narrative systems during their investigation into this topic including Virtual Teletubbies [7], FearNot! [8][9], and the double appraisal method [9]. Aylett and Louchart identify pencil and paper role playing games (RPGs) as an example of interaction and cohesive storytelling combined through the games referee's (often referred to as Games Master or GM) ability to plan a structure but then negotiate with players throughout to alter this structure based on their interactions. Aylett and Louchart identify that this is in part due to two key mechanisms of RPGs; the modular structure of the narrative as encounters, and the ability of the author (in this case the GM) to constantly negotiate the structure with the audience (players) based on their wishes. This led to work exploring how these could be harnessed in order to challenge the narrative paradox in VEs [76]. The modular nature of RPG storytelling led to FearNot! [9] which presented narrative as episodes with interaction opportunities in-between and agents playing characters resolving the content of the following episode which I explore further in the narrative generation section. Whereas the concept of the GM negotiating with players was expanded on in the double appraisal system [9] where the emergent narrative was delivered through the actions of character agents that appraised their options first based on their goals and situations and secondly based on the goals of the player and their 'emotional impact'. Both approaches demonstrate some success at the implementation level in the delivery of interactive narratives with increased cohesion though the authors note there is still significant work to be done on modelling the impact of a GM on narrative. The work covered by Aylett and Louchart could be said to be a new approach to the question of plot in narrative system, much like the narrative generators in the following section. However, as with most largely character centric approaches the emphasis is on the explicit content of the narrative and construction of subtleties such as the subtext or themes is largely not a part of the process.

Aylett and Loucharts identified problem of 'Narrative Paradox' has been explored by others under differing names, such as the 'boundary problem' in work by Magerko [77], as well as being explored by Harrell in his work on agency [58], and generative interactive

poetry (the GRIOT system discussed more fully later) [56]. Harrell explains there is a need for structured balance between user agency (interactive actions taken by user through the protagonist) and system agency (actions of protagonist forced by the system to ensure plot cohesion). Harrells 'agency play' [58] represents an initial attempt to model this structure in terms of relationships, scope, dynamics, input direction and shows similarities in the contrasting agencies it intends to consolidate to the work in compromise approaches of narrative generation explained below. It is possible that this represents a struggle between simulation (what the characters would likely do or what the player chooses to do) and storytelling (what structurally makes an engaging plot) that might be present in all narratives and manifests in narrative systems as difficult issues surrounding agency.

2.2.3 Narrative Generation

As a process narrative generation can be broken down into three stages; story, plot, and presentation generation. Depending on the project in question these stages can be consolidated together or separated, (for example, in the virtual storyteller, presentation generation is broken down into narration and presentation [110]). The majority of narrative generation projects deal with the creation of the narrative elements (story generation); resolution of the sequence of events that comprise the narrative and selection of narrative elements to be exposed and building of relationships between these elements (plot generation); and presentation of the narrative through a chosen medium (presentation generation). Figure 2.4 illustrates this process.

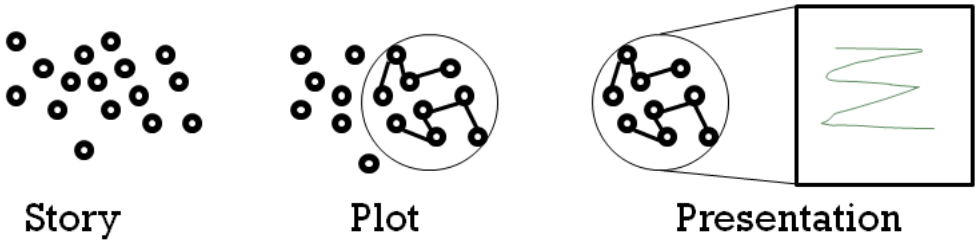


FIGURE 2.4: Narrative generation can be broken down into three stages

Riedl and Young in their work on narrative generation [99] describe the success of narrative generation systems through two variables; *Plot Coherence* and *Character Believability*, with the ideal system producing high values of both. Riedl and Young go on to say narrative systems can be classified in the space as either taking a *character centric* or *author centric* approach depending on whether the system seeks to model the characters within the story, the authorial process itself, or whether the system is a compromise of both approaches. Typically they say author centric approaches lead to a high plot coherence whereas character centric approaches have a high character believability.

Riedl and Young also identify a third approach in the form of story centric approaches

focussed more on the linguistic process of storytelling itself. An example of a story centric approach can be found in the work done by Bailey [10] on the readers perspective. Baileys approach involves modelling the audiences reaction to different parts of the story discourse and modifying the narrative to be generated at the next step of the plot based on this. In this sense Baileys approach could be considered to be a modified author centric approach as it has a modelling focus external to the content itself, but it is concerned primarily with the telling rather than the writing. The linguistic focus of story centric approaches makes them less concerned with modelling the narrative and more concerned with language of the discourse and as such I will concentrate on the other two more common approaches more relevant to my work.

In the following sections I explore a range of narrative generation systems structured around the *character centric/author centric* distinction made by Riedl and Young [99]. I begin with *character centric* systems which are largely based on simulating the content of a story (such as through intelligent agents playing characters), and then follow it with an exploration of *author centric* systems based on simulating the act of authoring itself (such as with plot structures and templating). Finally I conclude this section with an exploration of recent approaches that have sought to include elements of both as a *compromise* approach.

2.2.3.1 Character Centric Systems

Character centric narrative generation involves modelling the behaviour and goals of the characters of a story. With the characters successfully simulated they are released to pursue their goals and their actions are exposed, the idea being that stories are everywhere and an engaging narrative will naturally emerge from the actions of a set of well-motivated characters.

Character centric narrative generation systems often use agent technology to suitably simulate the characters and their behaviours with a purpose built agent taking the part of each character. An example of a character centric narrative generation system can be seen in the work done by Cavazza [33]. Cavazzas system uses intelligent agents to simulate the actions taken by characters within a story. The agents have predefined rules that represent the characters personality and goals and then in simulation the agent follows these to form plans as hierarchical task networks (formalised AND/OR networks) and then perform actions. A screen shot of the system shown next to the plans of action formed by two agents can be found in figure 2.5. The system uses the unreal engine [2] to then present the actions of the characters through virtual models, the camera following the protagonist or where the users wishes it to. The systems Cavazza built are also partially interactive allowing the user to alter the course of the plot through the characters and items. The importance of believable characters is evident in this

[2] A 3D game engine: http://www.unreal.com/ as of 31/7/2011

work and further work was completed investigating the generation of dialogue within the system [32], however at this point the focus of the work became more linguistic then narrative.

FIGURE 2.5: Cavazzas work on narrative generation [33]. A screenshot of the system presenting a scene and the plans formed by each character.

Facade was an interactive storytelling system by Mateas and Stern [80]. The story put the user in the role of a character along with two others in fairly typical drama (that of a quarrelling couple) in which the user is able to interact with the other characters and affect the course of the story. The story is a character centric narrative generator in the sense that it is constantly regenerating the remainder of the narrative based on the user's actions and this generation is based on the simulation of the other two characters in the story. Much like the work by Cavazza the characters are simulated using intelligent agents that are pre-authored with a set of rules that define their personalities and actions. The generation in Facade is however effected by a plot model, potentially making it a compromise approach, not character centric, as I will explore later.

Another example of interactive character centric narrative generation can be found in the IDtension system by Szilas [109]. IDtension describes itself explicitly as not character centric but property centric, however its approach matching a number of accepted definitions of character centric narrative so warrants its inclusion in this section. Key to the method taken by Szilas is the modelling of 'narrative properties' in the world in which the story is set. These properties are variables based upon what he describes as core emotions within the narratives plot such as 'violence' and 'law'. However the story itself is emergent from agents playing the characters (thus its inclusion as a character centric narrative) and it is the content of the world that is modelled rather then the authorial process. The agents have predefined positions along different variable axes, representing the narratives key properties defining whether they are promoting them or hindering them and present narrative possibilities to the player appropriately through a finite list of possible actions. This is where the approach taken by Szilas differs from traditional character centric narratives in that the actions of the characters are based on their disposition to key concepts underlying the story rather than a complete model of their personality. The focus on underlying concepts (his properties) and how they affect the story generated rather than the literal plot itself is unique in narrative generation

and hold similarities with my own thematic approach except that what comprises the concepts is only loosely defined and tied strongly to the characters themselves rather than the stories subtext. Szilas finds his approach effective in his evaluation, increasing the 'interactivity' of the stories through his generative approach. However the stories themselves are not presented through rich language suffering from the same problem as many character centric narrative generators of creating a series of reported actions rather than a rich discourse and through this the subtleties of his unifying concepts are somewhat lost. He identifies that this presentation of the narrative (the 'theatre') is a crucial area of further work.

Not all character centric narrative generation systems are based on agents simulating the actions of characters. Earlier approaches often had more simplistic reasoning systems that handled the actions and behaviour of all characters based on the goals of the protagonist. An example of such a system is the early TaleSpin [85], which generated narratives by simply reporting in text the action and position of every character sequence. However the narratives generated by TaleSpin and similar systems lack any notion of plot or narrative design and the simplicity of the characters and lack of personality can make the results feel rather bland.

Automatic generation of story elements is rare in character centric narrative generation. This is because elegantly written characters with sophisticated behaviour are key to narratives being successfully emergent from the generated result and at present the only way to ensure this is to build the characters by hand. Some story elements are generated by using character archetypes with cliched behaviour such as with the supporting characters in work by Cavazza [33], which are based on archetype templates rather than specifically written characters, but it is rare to find this for key characters. More completely automatic systems do exist but without sophisticated main characters the stories quickly suffer from the same problem as systems such as TaleSpin [85] with highly simplistic and bland results.

Plot generation in character centric systems is therefore a direct result of the characters behaviour as dictated by the agents playing them or the intelligence modelling all of the characters. The actions they take to achieve their goals builds the relationships between story elements and the sequence of events that makes a plot. Presentation generation is not specifically tied to the character centric approach but the focus on entities and modelling their actions make character centric approaches ideal for presentation in game engines (such as demonstrated by Cavazza). Although the presentation of character centric systems still sometimes uses text as a medium of choice either using sentence templates such as in Talespin [85] or generated text using natural language processing.

The main weakness of character centric narrative generation is its reliance on an engaging narrative successfully emerging from the exposition of the characters actions. Often these systems generate bland stories that merely report on a series of uninterest-

ing actions. These stories are thus often sensible and varied but lack narrative richness or interesting plot. Ultimately this can be traced to the agents or systems behind the narrative generation which are attempting to simulate what a character would do, not necessarily what is interesting to read. It is possible that narrative on some level is emergent from all human experience but without a structured plot or engaging presentation these narratives, presented alone as a report of events, can seem bland and directionless.

2.2.3.2 Author Centric Systems

Author Centric narrative generation seeks to model the authorial process itself rather than the content of the narrative. The systems seek to model the process by creating rule based systems or narrative grammars that use well defined structures that are typical of the desired genre of narrative in order to generate stories, it is typical of earlier work in the area such as that by Lebowitz [72] or Hovy [63] to seek the relationships between elements that make up the discourse of a story in order to build up structural grammars.

Author centric narrative generation can lend itself to the representation of existing knowledge as narrative as it is primarily concerned with simulating the plot elements and story structure created by the author then the narrative content itself. This is evidenced in some author centric narratives systems generating narratives using collections of material from elsewhere; normally the Web. ArtEquAKT [117], a system created at The University of Southampton, is an example of author centric narrative generation that automatically generated biographies for artists. ArtEquAKT creates narratives out of a variety of resources and media from the web about the relevant desired artist, for ArtEquAKT the story generation is the compilation of these resources together. ArtEquaAKT retrieves story elements off the Web [3] relevant to the artist in question from a mixture of search engines and predefined sources using its knowledge extraction tools. The relevant content retrieved is then processed into an ontology on the artist so that the content collected may be used in the correct context. Then, using structures and grammars designed specifically for the biography genre, a template for the biography is constructed and the link server AuldLinky [86] is used to manage a set of queries to the ontology that populate the template with content and generate the final narrative. A diagram representing ArtEquAKTs architecture can be found in figure 2.6.

Another example of author centric generation for a specific genre of narrative can be found in the work by Harrell on GRIOT [56]. GRIOT allows authors to prepare interactive poetry that generates fitting lines based on keywords the reader feeds to it. The author prepares templates for phrases to occur in the poem, theme domains for desired themes (a collection of associated keywords), and a collection of structural rules representing desired narrative structure. As the user enters keywords GRIOT returns phrases based on the desired term, themes, and structure. To do this GRIOT employs a technique of conceptual blending [57] where ontological representations of themes and

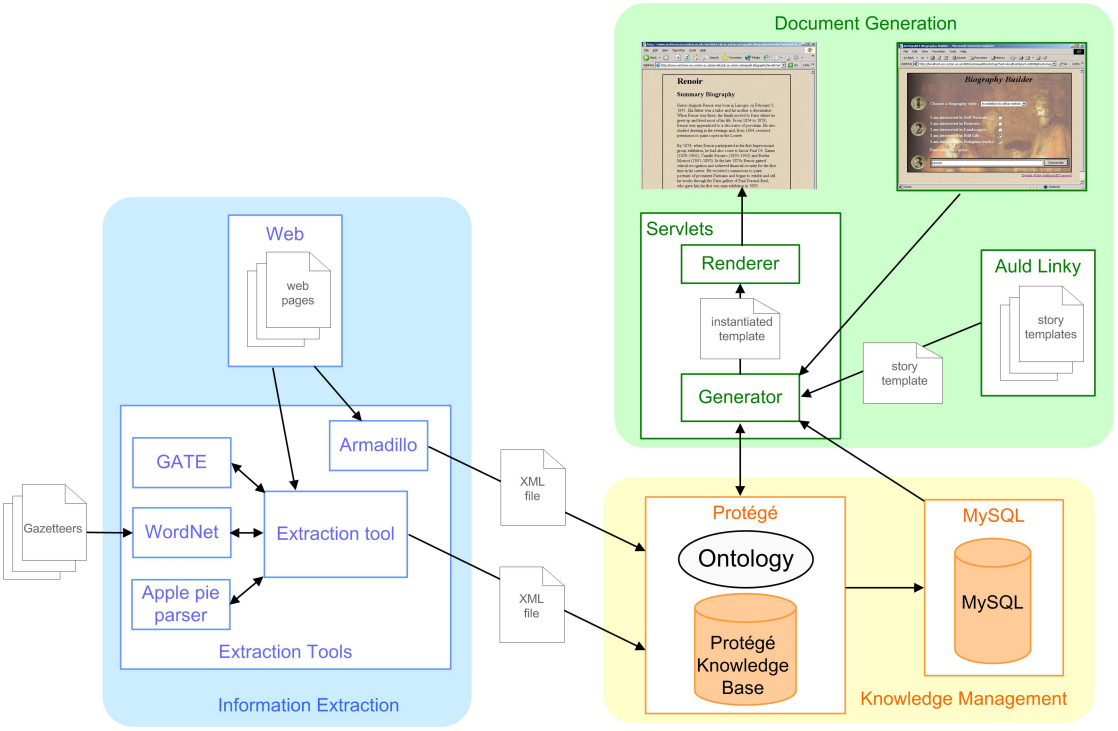


FIGURE 2.6: ArtEquAKTs architecture for generating biographies [117].

keywords are blended together to build maps of connected concepts. This is done using ALLOY; an implementation of Harrell and Goguens own conceptual blending algorithm [48]. Conceptual blending allows for GRIOT to incorporate diverse themes and form metaphors by blending the features of a desired content and a keyword being used as a metaphor, first to consider the suitability of the pair for constructing a metaphor and then to construct the metaphor itself. Providing proper preparation is made by the author this allows GRIOT to create convincing poetry, and the use of conceptual blending is both a novel method of metaphor generation and perhaps a useful tool (if somewhat simplistic) towards fostering narrative cohesion.

The adaptive hypertext system Topia [4] mentioned earlier could also be considered an author centric narrative generator centred on the representation of content off the Web. While Topia's narratives seek to serve a very specific function (search results), its generation of the presentation of results could be considered a narrative and the methods it uses are concerned with authorial simulation as opposed to content.

Some systems do model the contents of the narrative to be generated as part of story generation as opposed to representing existing content but still remain author centric. Universe [72] is an early author centric system that generates goal driven melodrama. The narratives are generated through a combination of character frames defining given characters and plot fragments detailing the actions of a given segment of plot. The character frames are detailed in Lebowitz's earlier publication [71], and are based on a model of several dramatic statistics representing each character. As the plot frame

calls for each character one is generated or selected from existing characters, generated characters are made from combinations of predefined stereotypes (taken from melodrama clichs) that define values for the characters statistics. The plot fragments explored are grammar based and provide a certain goal with a series of possible actions and candidate roles to be played by characters [73]. The statistics of the characters that fill these roles determine how the plot unfolds and the resulting narrative is generated. However it is to be noted the presentation for Universe is very simplistic and presented as a simple report of each plots major points as the goal of the project was the generation of plot plans as opposed to rich text.

Work by Hovy [63] similarly maintains an author centric perspective simulating plot decisions of the author but generating its own content. Hovy's work generates monologue narratives from a single perspective on the subject of a series of simple predefined narrative events. It has a very strong linguistic focus and is tightly coupled with work in natural language processing in that it aims to generate coherent discourse around its plot points based on a series of linguistic rules. Hovy's linguistic focus gives his work some things in common with the less frequent story centric approaches identified by Riedl and Young [99] but it is the author centric aspects that are more principally interesting to this work. As well as constructing text around plot points Hovy's work constructs the structure and sequence of plot points, this is done to ensure the most coherent story. The *text plan* for each plot point expresses preconditions, constraints, and a decomposition of definitions and details in its content. As well as being used to seed the eventual language generation Hovy demonstrates that this may be used to ensure plot elements occur in a coherent order where preconditions have already been exposed and constraints are met.

Callaway's work on the StoryBook system and Author architecture [28] is a more modern example of author centric narrative generation system that generates its textual content. Callaway presents the 'Author' architecture for narrative generation systems as a marriage between plot based author centric narrative generation such as work by Lebowitz [71] and natural language processing for presentation, this approach is then demonstrated in an implementation of the architecture called StoryBook. Callaway's approach takes a story specification in the form of rich text detailing characters and events that occur such as the finished results of a TaleSpin [85] story, he creates a simple similar system for this purpose. It automatically extracts the actors, environments, and events from this text to build an ontology representing the story as a fabula basis. This is then used to build sentential specifications for each chronological event in turn which are processed by a three stage natural language generator (Sentence Planner, Revision Component, Surface Realiser) to generate a richer text narrative representation of the story. Fundamental to Callaway's work is the concept that narrative generation while demonstrating sophisticated models of plot has yet to make progress in other crucial areas of storytelling such as prose. As such his significant contribution is in natural

language processing rather than models of narrative; within his three stage process he presents rules governing variety and complexity that generate the more evocative language suitable for storytelling.

Card Shark [16] demonstrates an author centric system that neither generates its content nor gathers it from the Web. Card Shark's material is written by an author who writes cards containing narrative sections complete with constraints on sequence based on numeric order and the content of preceding items. The narrative generation system itself is in the rules by which these cards are placed in that each 'player' has a random selection of narrative sections and plays a section of their choice that conforms to the current constraints. The resulting narrative is a stream of cards selected by the users, Bernstein describes this as a 'Sculptural Hypertext' in that the resulting links are present through the cards that stay after the unplayed cards are removed. The system itself is a simplistic set of rules modelling the causal cohesion process of an author, and as such can be considered author centric.

Card Shark's core ideas have been utilised in an automated generator in the Thespis system [16], where automated 'actors' follow a Card Shark like systems playing out a story. Thespis relies on characters in a piece being predefined with a set of possible actions with constraints much like the sections in Card Shark. Thespis takes predefined characters and then following the rules and constraints of their possible actions plays out the hypertext drama in a two dimensional space that the user is free to explore. Although the rules which govern how these actions occur are still causal the remodelling of them as actions for individually modelled characters gives Thespis a character centric aspect and it could be considered a compromise approach as explained below.

Author centric approaches typically perform very well at generating specific narratives in a well-defined context, such as the artists biographies of ArtEquAKT [117], as the systems are purpose built for strong coherent narratives within these genres. However this makes it very difficult to create systems that can create varied stories of different genres as they would necessitate different structures and rules for each genre and could become very formulaic.

2.2.3.3 Compromise Systems

Many narrative generation systems often seek a compromise approach in order to counteract the weakness of using one approach or another. Some systems such as Universe [72] will only make slight compromises, such as the ideal story drama curve approach in Facade or the choice to model characters in Universe, but others make much larger steps towards marrying the two approaches.

The Virtual Storyteller [110] is a narrative generator that at first seems to be a character centric approach. Like other character centric approaches it uses a set of agents to

simulate the actions of characters in a fictional story in order to generate narrative as a report of the actions performed. However where Theunes approach differs is the use of a director agent similar to Facade [80]. The Virtual Storyteller's director agent vets actions taken by characters to ensure they conform to a set of rules governing the quality of the plot. These are then passed to a narrator agent who converts passed actions into a text discourse that might be presented in a range of different ways. The intention is to create stories with high character believability but also high plot cohesion. However the director is reliant on its knowledge base of constructed rules and as such to generate truly varied stories necessitates the creation of a broad range of rules for different genre. Swartjes and Theune further developed the Virtual Storyteller with subsequent projects [108]. The director agent became a 'plot agent' that as well as interacting with character agents interacted with a world agent representing the setting. This world agent had access to information about the narratives environment but also its particular genre through a knowledge base in the form of an ontology. This allows the director (or plot agent) to separate between generic storytelling rules (such as preventing the antagonist killing the protagonist in the first scene) and genre specific rules. This allows potentially for a genre ontology to be constructed separately and for the virtual storyteller to generate narratives for any genre specified. The Virtual Storyteller continues to be developed and through it Theune may have brought some of the strength of author centric approaches to a character centric approach but with it also its limitations in its reliance on the quality of its knowledge base and its rigidity to defined genres.

FearNot! the interactive narrative system discussed earlier created by Aylett and Louchart [8] [9] contains elements similar to both Facade[80] and the Virtual Storyteller[110]. Its system is primarily concerned with the delivery of preconstructed narrative (in the wider VICTEC project FearNot! was involved to provide social education on bullying) but the specific narrative delivered was determined by character agents and actions taken by the interacting user. Similar in the way Facade delivers preconstructed dialogue and actions FearNot! delivers episodes; a similar technique utilised at a larger scale. The agents role in adapting the narrative presented could classify FearNot! as a character centric narrative generator, however its attention to structure makes it a compromise approach. The structure of prewritten episodes inherent to the system means that the agents are tied to the structure of the plot. Also the inclusion of a 'Stage Manager' agent whose rules and goals are authorial in nature along with a role in vetting character actions is very similar to the director of the Virtual Storyteller.

There are examples of other interactive narrative generators that have sought a compromise approach such as the work of El-Nasr[41][42] on proposing a narrative generator informed by theory from performing arts and filmmaking called Mirage. El-Nasrs approach is similar to both Facade[80] and FearNot![8] in that intelligent agents deliver different pre-constructed material for their characters based on the interactions of the player and the authorial rules of a central director agent. El-Nasrs 'beats' are more

similar in scale to the material delivered by Facade[80] in that they tend to be individual actions and pieces of conversation as opposed to the larger episodes delivered by FearNot![8] giving a more directly interactive feel. El-Nasrs key contribution is however more in how the resulting narratives are generated as opposed to dealing with issues facing interaction, her architecture aims to support the presentation of generated narratives with different dramatic techniques. Mirage currently runs only a single story, an adaptation of the Greek tragedy Electra, in which the user participates as a character in a 3D environment. El-Nasr identifies presentation as a key stage in narrative generation and in order to better present the desired narrative the director agent in Mirage, as well as vetting character actions, makes decisions on visual composition mostly through lighting and the camera. This includes ensuring the camera is focused on the action and a set of lighting actions for various tasks such as heightening dramatic tension, matching current mood, and highlighting objects or characters of emphasis. In a later evaluation of this work[42] taking opinion interviews comparing Mirage with and without the visual composition techniques from experts within the performing arts El-Nasr notes a perceived improvement in the system's ability to effectively convey the narrative, though still with some criticism of choices made by the director. Mirage represents a different sort of compromise in narrative generation in that it is a character centric system paired with an author agent that not only considers structure but also visual composition. Its focus and findings show an emphasis on the importance of presentation in effectively generating narrative.

While the previous systems are largely a character centric approach with attached author centric systems ACONF [99] is the inverse. While Riedl and Young perform their assessment of existing narrative approaches [99] they also present their own solution to the problem; Actor Conference (ACONF). ACONF is principally an author centric approach similar to Universe [72] in that the story is based around a singular overall goal (such as a successful bank robbery). A knowledge base of possible scenes/actions and their causal connections is then analysed by an earlier developed planning system called Longbow [120] to construct possible sequences of plot segments that might achieve the goal. A designated number of intelligent character systems are then added to the system, each in turn representing an expert system on the personality and capabilities of a character in the plot. Characters contain their own causal knowledge bases of actions they can successfully perform (including the outcomes of these actions) and use their own versions of Longbow [120] to break down the plot segments and successfully fill them with actions. This is done collaboratively and when one set of actions cannot be filled by a given character other characters fill in incomplete segments with actions (Riedl and Young refer to these partially populated plot segments as 'narrative hypotheses'). The fully generated plan of actions that represents the completed generated narrative is then passed to Mimesis [119][121]; a narrative presentation system that takes the actions from an ACONF plan and its characters and renders the story using the UT game engine similarly to the work done by Cavazza [33], an example screenshot of a Mimesis

presentation can be found in figure 2.7.

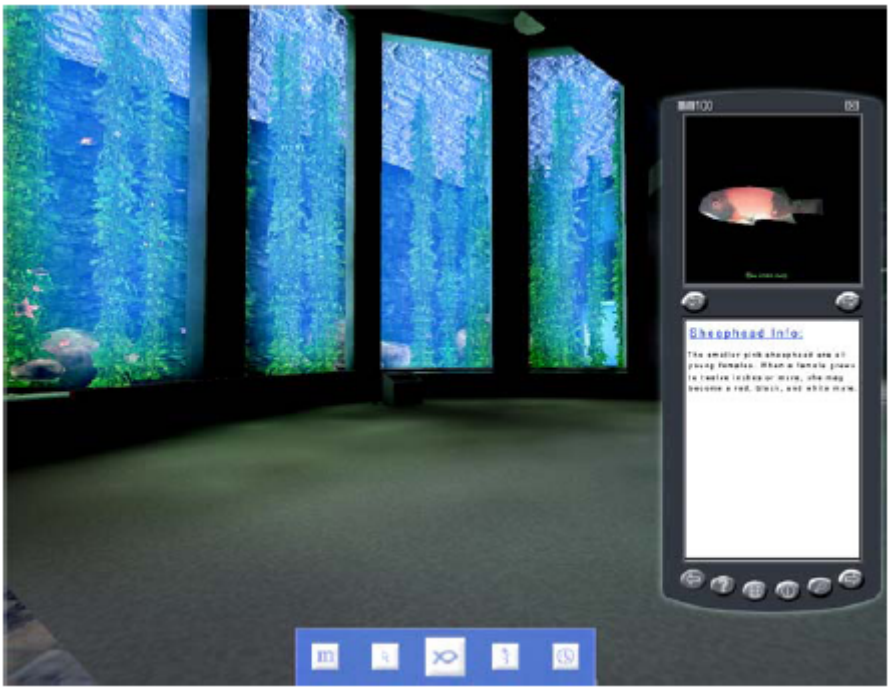


FIGURE 2.7: A screenshot of the Mimesis narrative presenter [122].

Subsequent work from this group has explored other aspects of plot construction and ordering that might play a role in narrative generation such as the work on Suspenser by Cheong and Young [36]. The Suspenser project aims to model how suspense is built in a narrative and incorporate it into planner based narrative generation using Longbow [120]. Cheong uses a definition of suspense as the reduction of options for a protagonists successful outcome, noting the importance that options are maintained throughout but increasingly decrease as the suspense mounts. As such the Suspenser system when ordering a plot of potential narrative elements exposes a multitude of options early and then increasingly selects actions that fail the preconditions of options. In evaluating Susepenser based on user perceptions of suspense Cheong finds that it can be controlled to delivering suspense at different levels (high and low) and can also at its highest settings generate stories of greater suspense then humans in some cases, though it is to be noted this is not to a degree of statistical significance.

Recently Riedl has completed an exploration into what might represent future direction of compromise approaches to narrative generation in his comparison work on interactive narrative systems [98]. In this work Riedl is predominantly interested in character centric, emergent, approaches and the effect of introducing author centric elements, such as the director agent included in so many systems or as he refers to it 'drama management'. Riedl performs a low tech experiment using improv actors instead of agents in a system to play out scenes for an audience either based solely on their given characters or also with the overall dramatic direction of an external author/director.

Evaluating the effect based on audience perceptions of his own earlier identified notions of character believability and plot coherence [99] Riedl finds very little impact is made by the external direction, but that core desirable individual features of the narrative did come across stronger. Riedl concludes that while author centric elements such as a director may not improve plot coherence as he expected they are useful for guiding the content of generated narrative. He also states that future compromise approaches may want to explore agents as actors rather than characters, where the character agents themselves contain author centric rules for creating the desired narrative rather than relying on an external director agent.

These systems experienced mixed success with both reporting the generation of successful narratives. However they suffered from similar problems to character centric approaches, while the addition of measures to ensure the narratives structure is engaging does have a positive effect the engaging narrative can at times still fail to emerge from the result and the systems can be reliant of stories that are heavily predefined at the request stage rather than being entirely generated. Furthermore these systems while exploring further refinements of plot and character exposition still largely demonstrate an ignorance of sub textual concepts such as themes.

2.3 Term Expansion

As part of building structuralist models of narrative I will be considering the connection and relationships between concepts and how they are in turn connected to individual terms either in textual language or representing a feature in multimedia (such as a tag). The machine based expansion of connections between terms and concepts is something that has already been explored in the information retrieval field of term expansion. Term or query expansion allows systems to improve the relevance of information requests. By expanding the terms in user's queries or the candidate terms against which they are being matched a greater number of successful matches may be found. There are a variety of methods that can be used to achieve such expansion by assessing different relationships between a term and other terms. In this section I will explore existing research into how terms might be connected and how they are expanded to related terms to provide context to my own semiotic expansion detailed later in the thesis.

2.3.1 Lexical Systems

Perhaps the most straightforward method of term expansion is to use a thesaurus, expanding a term using synonyms and other similar words. WordNet [89], a large general purpose thesaurus developed by Princeton University provides a good basis for a system undertaking such an expansion with a large variety of terms and many different kinds of

lexical relationships drawn between them. Voorhees conducted an initial investigation [116] on the generic effectiveness of lexical query expansion using WordNet as a basis for different lexical relationship and using the TREC collections[3] as test search data. However Voorhees work shows that there is little advantage to such expansion, finding only minimal improvement on very small queries and no improvement on larger ones.

Buscaldi experienced more success using WordNet in his work [27] exploring the more specific problem domain of searches for geographical locations (again using the TREC collections but only a subset based on geography). Buscaldi used meronyms and holonyms (textual relationships that denote a term refers to something that is part of or contains the referred to term, e.g. London is a meronym of England, the holonym is the inverse) to expand queries based on geographical locations so as to find more broadly relevant locations and improve the overall relevance of the search. However, while such a lexical technique works well for geographical searches, this is such a specific area that it is difficult to expect this sort of expansion to be successful everywhere, not all terms can be expanded usefully with meronyms.

2.3.2 Co-Occurrence

Another form of term expansion is co-occurrence. Co-occurrence is a statistical method involving analysing the semantic similarity of two terms based on the frequency with which they occur together in a document. Co-occurrence can be used in automatic keyword extraction, such as in Matsuo's work [81], but also for query expansion as I will discuss in this section. In such systems a corpus of potential results is analysed and terms attached to documents in the corpus that co-occur frequently with the terms used in the query are used to expand it. This method of expansion is automatic and has returned impressive results. This has been demonstrated in work such as Buckleys [26] where as part of the Smart system he employed co-occurrence evaluating its generic text search performance using the TREC collections. Buckley found a 7-25% improvement in relevance when Smart utilised co-occurrence for query expansion. There are some important factors to consider when expanding in this way such as what corpus to use to train the expansion, and how to measure the likelihood that a co-occurring term is useful for expansion. Work done by Xu [118] (again exploring generic text query effectiveness with the TREC collections as test data) on co-occurrence found that a local corpus produced more relevant results then a global one, emphasising the importance of training the expansion with a limited set of documents of ascertained relevance. Peat [96] performed an intensive study of different measures of term similarity in co-occurrence and highlighted that the frequency of term occurrence in the corpus is an important factor in determining whether its co-occurrence with a query term is significant. He

[3] A benchmark collection of documents, standard textual queries, and value judgements of relevance established by the TREC conference. Can be found at http://trec.nist.gov/data.html as of 7/6/11

demonstrated that by using clustered, important, low frequency terms relevance could be further improved.

Measuring the similarity between terms in co-occurrence is done in several different ways. Mandalla [78] identified three main equations; Mutual Information (MI), the Dice Coefficient (DC) [60], and the Tanimoto Coefficient (TC) [101], finding the best results with MI. The equations for these are displayed below, where for MI P(a) is the probability term a occurs, P(b) is the probability term b occurs, and P(a, b) is the probability the two terms co-occur. For the other equations F(a) is the frequency a occurs, F(b) the frequency b occurs, and F(a, b) the frequency the terms co-occur.

$$MI = log\frac{P(a,b)}{P(a)P(b)}$$

$$DC = \frac{2F(a,b)}{F(a)+F(b)}$$

$$TC = \frac{F(a,b)}{F(a)+F(b)-F(a,b)}$$

These are all slightly different methods of measuring similarity but are essentially looking at a similar ratio of the frequency a term co-occurs with another term over the frequency of each term occurring with or without the other. Core to these methods of measurement is the concept that a term which occurs alongside another for the majority of its occurrences can be considered likely to be similar. Detecting the actual co-occurrence itself is a different challenge where terms can be said to co-occur if they are adjacent, or share a sentence or paragraph. Hearst however has shown greater success using a multi paragraph variable window of co-occurrence size [59]. The text windows or 'passages' used by Hearst are based on discussion topics within the text, subdividing the text based on the detected main topic of discussion by measuring the distance of given terms from candidate topics. However, as is the case with the research present in this thesis, co-occurrence detection becomes significantly more simple when not dealing with rich text. When using folksonomies or other corpus with tagging meta data detecting co-occurrence of tags is as simple as detecting that a document has both the terms as tags.

Mandalas work [78] undertook a comparison between a variety of term expansion methods, including a generic thesaurus, co-occurrence, and Head-Mod as well as combinations of the three. The third method mentioned, Head-Mod [61], is a method similar to co-occurrence in that it is based on detecting similarity through usage. However it is not based on statistical occurrences but on grammatical usage within a corpus, expanding using terms used in similar grammatical patterns. Mandalas results suggest that co-occurrence based on a local corpus is the most effective method of term expansion. However all of these methods are sometimes subjective to a dangerous side effect of term expansion that is frequently discussed in work such as that by Zhou [123]; query drift.

Query drift is where repeated term expansion through terms falsely considered to be relevant allows the results of the query to become tainted with irrelevant subjects, and has been experienced with all methods discussed so far.

Some work has been done to mitigate the effects of query drift and improve the quality of co-occurrence. In his work on improving co-occurrence with local feedback Mitra [90] demonstrated that using feedback to train the corpus used for co-occurrence can improve the relevance of co-occurrence, amongst other things reducing drift. Whereas Carpineto [31] used statistical elements of information theory in order to add weighting to terms selected for expansion, allowing terms that were statistically very likely to be relevant to the query term to receive a greater focus in the resulting query after the expansion.

Co-occurrence appears to be the strongest method for term expansion in improving relevance of queries. However it is a solely statistical basis for inferring what as users intentions were when using a term rather than based on any semantic understanding. As such it is vulnerable to query drift and its effectiveness highly dependent on the quality of the corpus used to train it. Work done in other fields is now beginning to use models of expert knowledge as a basis for expansion for queries in specific fields such as the work done by Fu [46] which uses an ontology to expand and improve geographical queries similar in objective to the co-occurrence work done by Buscaldi [27]. However any success experienced by such approaches is in some ways hindered by the fact that such models are not generated automatically.

2.3.3 Ontological Approaches

While considering the semantic difference between terms and their expansion it is important to consider the semantic web and the completed work in the area on building connected vocabularies and ontologies. Berners-Lee's initial concept of a machine readable web [15] shared much in common, motivationally speaking, with this work (though on a much less specific scale) in its goal to provide models of information that led to computer understanding of content being processed. At the heart of this was the notion that while data could be stored in a variety of different ways across the web, different representations could be compared and contrasted using an expert created document, formally specifying the relationships between terms, known as an ontology. Ontologies typically related to a specific area, and specified a taxonomy of the objects and properties for that area along with rules for inferring information. The power of inference could potentially allow for terms to be expanded across many chained relationships, rather than a single measure of semantic difference. In the work revisiting this idea by Shadbolt et al. [104] the progress towards a fully semantic web is covered, including a review of the technologies supporting the creation and use of ontologies. They explain that while successful progress is being made towards a semantic web there are still further questions to be answered and challenges that the underlying concepts face. This

includes that the ontologies which support it must be developed, and maintained, by committed practice communities of experts, a cost that is shared by the manual authoring approach to my own model suggested later in this thesis. Other questions are also raised including how different ontologies sharing a similar space might be aligned and mapped, and how trust and provenance for semantic content might be established.

Some work has already been done looking at how ontologies might be mapped on to each other and combined. This includes work such as the LODE project [105] which looked to build an interlingua model containing axioms expressing mappings between a collection of already existing event ontologies. In this project Shaw reviews the important modelling factors facing the concept of an 'event', such as temporal interval, location, participation, causal influence, and composition. This leads to a shared vocabulary in the form of a linked data model to allow for inference over multiple event ontologies. The model is evaluated both by modelling event data from Wikipedia[4] and testing for interoperability between existing event models, in which it performs successfully in both tests. The LODE project subsequently was taken forward by Troncy [114] who investigated how combining the representation of events with media representations using the Media Ontology could lead to a representation of events linked with images, videos and other media. Troncy explored what connected particular media to an event based on similar factors to LODE, these connections allowed for the automatic annotation of an event with relevant media. It is possible to see how event ontologies might be used to describe a narrative in terms of plot events within its content. This could be useful for systems concerned with narratives deliberately lacking a subtext or bias, such as historical records or encyclopaedia, where the explicit content of the narrative is crucial to the system, rather than any implicit subtext.

It is possible to use the semantic connections within an ontology as a basis for term expansion for queries. The connections between concepts might be used as a basis for term/query expansion such as in the work done by Fu [46] however, ontologies are normally purposefully created for specific fields by a small group of experts, fully exploring a small group of concepts. Ontologies such as the Gene Ontology [6] are used to expand terms used in queries relevant to their subject or glean further meaning from terms used in media related to their subject but are not much use in more generic scenarios.

While such ontologies are normally authored by a specific group of experts there are cases where these knowledge representations are generated in a more collaborative way by wider communities. Ontolingua [43] seeks to provide a set of tools for communities to collaboratively create ontologies allowing for wider groups of users of varying expertise to reuse existing ontologies, or extend existing ontologies, as well as create their own to work towards larger representations of knowledge. Ontolingua also supports these ontologies in achieving consensus over conflicting contributions. Projects such as VoCamp's [5]

[4] www.wikipedia.org
[5] http://vocamp.org/wiki/Main_Page

also work towards generating knowledge representations from larger communities rather than specific groups of experts, in this case consensus on vocabularies for specific areas of expertise that can be used by knowledge representations such as ontologies. It is also possible to automatically generate representations of knowledge, simulating the authoring process by processing large amounts of relevant information into elements and relationships that can be used to create an ontology. This was explored in work by Tho [111] where amongst other things an academic work ontology was created from a set of publication information. However even this requires a hierarchy of defined concepts and a corpus in a particular area and as such is limited to the creation of knowledge representations for a specific area rather than more generalised cases.

2.4 Background Conclusions and Summary

Narratology offers workable structures of narrative through structuralism [13]. These offer explanations of how to model themes (such as in thematics [113]), and how users might interpret the language within a narrative (through semiotics [102]). Narrative systems have explored a wide range of approaches to both generating and presenting narrative [20][99], sometimes using narratological theory [117], and have experienced mixed results. However most narrative systems are concerned with the chronology of events and content within a narrative [99], what we might call the plot, and give little attention to the subtext, including themes, something asserted as crucial by Tomashevsky [113].

The inclusion of a model of themes in narrative systems might lead to a fuller system understanding of narrative and potentially better results. For this to work we would need to define a structure, and explore how definitions of themes might be constructed as instances of the model. Such definitions would express how the different concepts present in a theme are related and how they connect to the language of the narrative itself. Term expansion as an area explores a range of different relationships between terms as a means for computationally understanding semantic similarity in language [116][81]. Potentially these methods might offer a way for connecting language to thematic elements by semantically measuring the distance between language used and themes in question. However, it is also possible the best results might be limited to a manual capture of semiotic relationships.

In the following chapter we explore a structure for a thematic model, as well as the method for how to define themes within it.

Chapter 3

The Thematic Model

Authors use themes to communicate a subtext within a narrative. This subtext may be an agenda or simply an emphasis of a particular part of the narrative or even simply an emphasis of the authors own style. This subtext gives a narrative direction beyond merely communicating a chronology leading to deeper narratives with authorial purpose [113]. Within this chapter I present my thematic model; a model of themes based on structuralist theory of thematics and incorporating semiotic relationships. This model could support narrative systems [50] such as narrative generators by allowing authorial understanding of, or rules concerning, thematic content. Primarily such rules and structures would be of use to author centric approaches to narrative generation, but could also be used by compromise approaches such as those that use a 'director' agent. Definitions of themes using this model expand the given themes into their components and associated features, effectively performing a semiotic term expansion; a process which we explore and evaluate in this chapter.

3.1 The Model

Existing systems work in thematics often looks at themes as a classification of content, performing keyword extraction in order to classify a document into a particular topic such as in work by Maria [79]. These Information Retrieval (IR) projects often seek to demonstrate that thematic classification can show a marked improvement over traditional keyword search but approach the concept of themes from the perspective of the core content of a document rather then its subtext. Projects such as the work undertaken by Paradis [95] bring thematic extraction closer towards discourse theory investigating human readers observations of what makes a documents theme but are still tethered to the idea of themes as topical classifications of content. More recently work on theme and mood extraction such as work by Bischoff [17] looks at extracting themes from multimedia (music in the case of Bischoff) and trying to support thematic tagging of work.

However Bischoffs definition of a theme; common usage of a piece of music (such as 'for travelling' or 'early morning') does not grasp the subtextual meaning that I mean to capture with my own thematic model. Harrells use of 'thematic domains' [56] is closer to what I propose but lacks any semiotic structure, while each represents a conceptual definition of the theme they are simply a collection of associated terms without notion of denotation or connotation and constructed without a particular method. My model seeks to represent the thematic subtext of a narrative and the concepts that influence the narratives content rather then classifications of the content itself, this is done within a structured model based on semiotic relationships.

3.1.1 Structure

To do this I go back to Tomashevsky's structuralist work on thematics [113]. Features within the narrative denote Motifs and from these Themes can be identified. I assume a situation where a story is compiled with many small segments of narrative that are structured together, in this case the selection of these small atomic segments and their content are key to communicating a theme. I use the term Narrative-Atoms or *Natoms* to describe these segments; small atomic pieces of narrative that cannot be further broken down by a given system and still make sense, this might be a single photo or paragraph, or a sentence and fragment of an image, depending on the natom in question. It is to be noted that my concept of a 'narrative atom' is directly connected to the media of presentation as supposed to a story element in the structuralist sense such as in work by Bremond [21] where he uses the term 'narrative atom' to refer to narrative functions (such as Propps). The content of these natoms is rich with information, however only some of it is visible to a machine (such as generated meta data and authored tags), I call these visible computable elements *Features*. Features might take any number of forms in that they can be any computer identifiable information on the natom. For the work in this thesis we commonly use tags but they might also be computer vision detectable features of an image, automatically created meta data, or features detected from other methods of analysis. Natoms contain any number of features which may or may not work towards connoting a theme in a story. Features can each denote a *Motif*, a basic thematic object that has connotations within the story, for example the feature *cake* denotes the motif of *food*. These motifs in turn connote broader *Themes* in the context in which they are presented, for example *cake* in the context of a gathering may connote *celebration*. These themes, when combined with other themes or motifs could in turn be used to further connote other themes, for example *wedding* might also connote *celebration*. This forms the foundation of my thematic model of a narrative:

- Natoms contain features
- Features denote motif's

- Themes are connoted by other themes and motifs

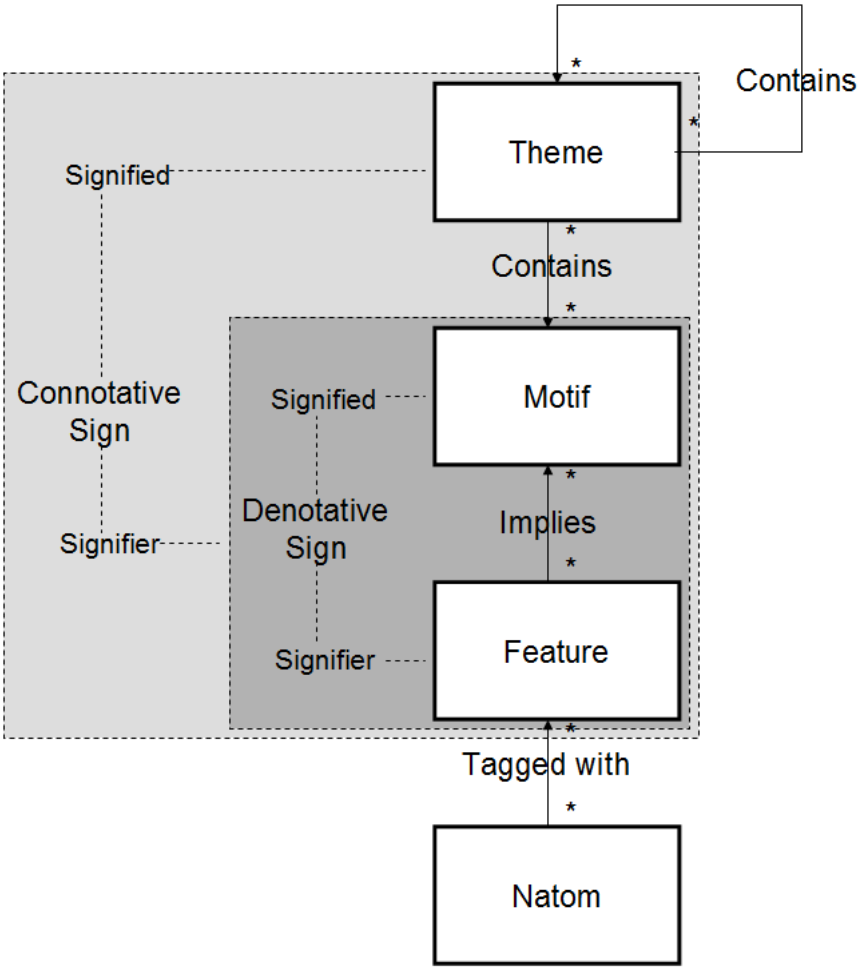


FIGURE 3.1: The Thematic Model

The model is shown in Figure 3.1, which also shows how the parts of the model map to Barthes' idea of denotative signs as the signifiers for connotative signs. Features denote Motifs because motifs directly define the feature (normally as a generalized version of it). Themes are broader concepts communicated over the entirety of the narrative, typically by numerous motifs. Themes are never denoted as they are complex entities not connected to a single item but become apparent after the presentation of a range of entities in the context of each other, as such a theme is a connotation of the motifs, and by extension the features, within the narrative. An example of how these entities can be connected to features in natoms is later shown in section 3.1.3.

While this model does not represent the basis of a complete model of narrative it could potentially be used in conjunction with other models to enhance narrative generation. As is discussed later in this chapter it could be used to select natoms to be used within a discourse. As such we could use themes constructed from this model to influence the story selection in author or character centric systems to give them a thematic subtext.

When a narrative is formed a part of the story is selected and then presented as a discourse[13]. We can consider virtual collections of resources as our story, and should we want to create a discourse to tell a story of Tuesday it might select all the natoms (photos, blog entries, etc.) of that day. Using an appropriately populated thematic model we could examine the features of those natoms in order to identify motifs and thus potential themes. Natoms that connote these popular themes could then be selected or emphasized to create a final discourse with more direction and cohesion. If the virtual collections were very large we could look for natoms that supported particular themes, for example, by using public photo collections to create a discourse (a photo montage) with the themes of family, winter and Christmas.

Because features can be tagged using a variety of different language for such a system to work every motif object would need a broad list of features that could denote the motif. In turn themes will also require some way of knowing what motifs are suitable for them, as a connotation is more complex and subjective then a denotation deriving a method for forming the relationship is more complex. While it is possible that some connotations might be stronger than others the model makes no provision for this as it would be extremely complex to form a way to reliably measure this strength and apply it. This is however something that could be incorporated into future versions of the model should a reliable way of measuring it be formed. The effectiveness of the thematic model will inevitably be connected to the quantity and richness of the thematic definitions formed in the terms of these elements and their relationships to each other.

3.1.2 Rules

The elements and relationships present within the model are governed by a series of rules ensuring that definitions made within the terms of the model are valid. These rules ensure that the relationships are used correctly and that when a definition for a theme is made the elements that comprise it are correctly based within the structure. The rules largely relate to the foundation of elements and relationships noted above but also govern the use of *justifications*. When a connotation relationship is formed between an *element* (either a theme or a motif) and a theme, a justification for the connotation is also added explaining why one connotes the other, and no one theme may be connoted by two elements with the same justification. This was included to aid the process of forming definitions. As definitions at this point are largely authored by hand, justifications help the author consider the role of potential elements in connoting a theme and help them consolidate the wide variety of relevant features into motifs formed around the key roles. In plain text these rules could be articulated as such:

- An element may be either a theme or a motif, not both, and all themes and motifs are considered elements.

- A feature is not an element, nor can an element be considered a feature.
- A denote relationship is always between a feature and a motif, and all motifs must be denoted by at least one feature.
- A connote relationship is always between an element and a theme, and all themes must be connoted by at least one element. Also all connote relationships must include a justification.
- No two connote relationships may exist with the same theme and justification.

```
theme(X) :- -motif(X), -feature(X), element(X), connote(Y, X, J).
motif(X) :- -feature(X), element(X), denote(Y, X).
connote(X, Y, J) :- element(X), theme(Y), justification(J).
denote(X, Y) :- feature(X), motif(Y).
:- connote(A, X, J), connote(B, X, J).
```

FIGURE 3.2: Rules of the thematic model expressed in DLV

These rules could be articulated more formally. I have been experimenting with a variety of constraint based languages for making a formal expression of the rules and have settled on DLV [1], a language that extends datalog [2]. Such an expression could also be later used in the creation of an automatic validator for instances of the model. The rules can be expressed formally in DLV as shown in figure 3.2.

3.1.3 Example

Figure 3.1.3 shows a simple example of how a collection of natoms connotes a theme in the terms of the model, in this case a passage of text[3], and two photographs that could be interpreted as connoting the theme of winter. The features presented are present within the given natoms, it is feasible that the natoms would be tagged with them or that they might be automatically extracted from them. These features literally denote the motifs of snow, cold, and warm clothing. As snow demonstrates many different features might denote the device of snow but in this case thematically they serve the same effect. Finally in the context of each other these motifs connote the concept and theme of winter.

3.2 Authoring Definitions

A possible draw back to the thematic model is its reliance on manual semiotic term expansion which unlike automatic term expansion approaches relies on hand written

[1] http://www.dbai.tuwien.ac.at/proj/dlv/ as of 13/7/11
[2] A query and rule language similar to Prolog
[3] text from William Shakespeare's Blow, Blow, Thou Winter Wind

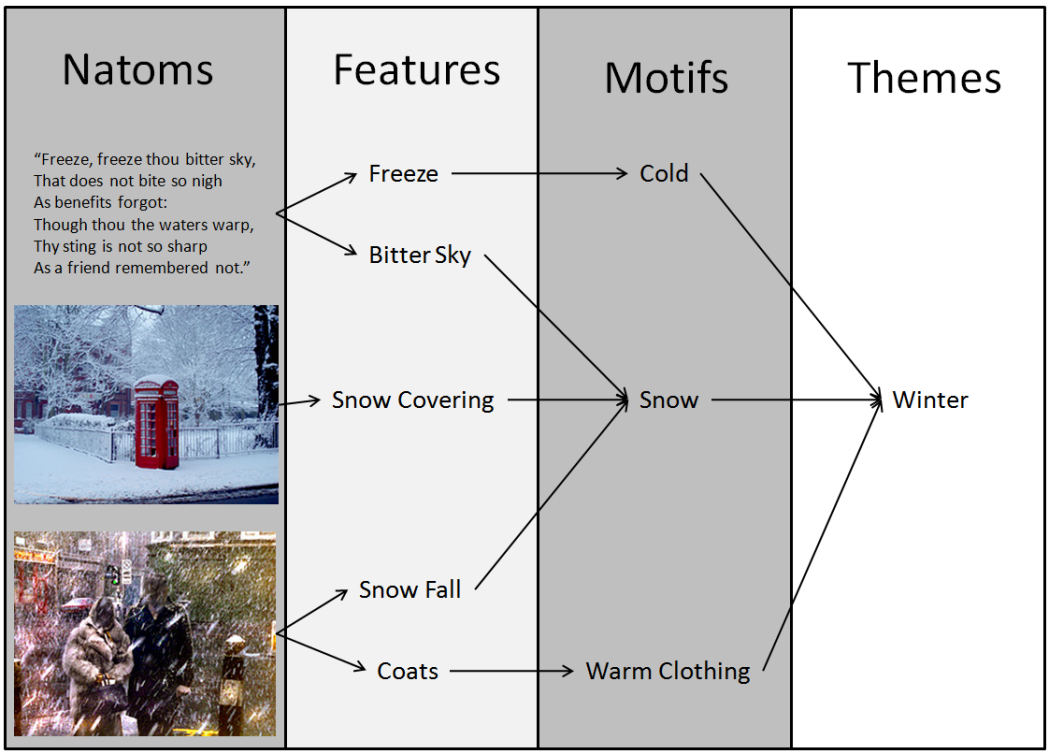


FIGURE 3.3: A Worked Example

definitions of themes (instances of the model). Initially I authored each theme used in initial experiments myself according to the rules of the model. A fully working system would need access to a very wide network of themes and motifs, something that is impractical for us to author by hand, it would need to be either automatically generated or contributed to by a wide community of users. As the thematic model relies on capturing peoples subjective views it is very difficult to automatically generate such a thing from existing resources on the web. As such the strongest course of action seems to be to develop a way for everyone to contribute to a wide base of definitions a system would use.

3.2.1 The Guide

In order for users to contribute to the system they will need to either fully understand the model or have a guide or tool that walks them through the creation process and allows them to create definitions based on the rules of the model. To begin with I elected to develop a guide that anyone could use to create valid definitions of themes in the terms of the model. In order to do this first I needed to analyse the process I went through in applying the rules of the model to create a definition. To do this the theme 'danger' was expanded and each decision made in the process of doing this recorded, having completed the process there seemed to be three key parts/stages.

1. Identifying component elements
2. Expanding sub-themes
3. Removing associated or weak elements

Identifying component elements
This part (stage in the authoring process) revolves around the extracting of key elements of the theme in question and classifying each as either a theme or motif based on the semiotic rules of the model. However it is to be noted that for someone with no experience in abstract modelling extracting what the key elements are of a conceptual theme is not a trivial task. To explain the process in a more accessible way to non experts it was split into two, a word association exercise and then a classification one. This way the contributor could happily engage in simple and more familiar word association to initially expand the theme and then classify the expanded words as themes or motifs based on the rules of the model. Unfortunately the free flowing imprecise nature of this process leads to much more verbose definitions with a lot of overlap between elements, so to counter this the contributor groups together similar elements or those that serve a similar purpose into a more well defined model.

Expanding sub-themes and removing associated elements
These two stages cover expanding the sub-themes into further themes and motifs and then removing the weaker associated elements. An observation of my own decisions in this process shows that the process of expansion is seldom done to completion, in order to save time on a long and recursive task the expansion of a sub-theme is often cut short if it becomes apparent that it will later be labelled associated and removed. As such it is important to note in the guide that these stages should be under taken simultaneously but they otherwise should be simple enough to be undertaken by a non-expert.

This leaves us with a final break down of five stages for defining a given theme in the terms of the model:

1. *List Associated words*: The contributor spends some time expanding the specific theme into a list of associated words to get a list of related concepts.
2. *Classify as Themes or Motifs*: The contributor then makes two lists using the results of stage 1 based on the rules of model classifying each as either a theme or a motif.
3. *Group elements*: The contributor groups together similar elements or those that share a similar purpose into a single element based around the shared purpose or a generalisation of the features they share.

4. *Expand Sub-Themes*: The contributor takes remaining theme elements and expands them as they have done the initial theme. Care is taken to consider stage 5 when doing this in order to save time.

5. *Remove associated elements*: The contributor removes each theme or motif that is not entirely relevant to the root theme.

An example of these steps being followed can be found as a part of the finished guide we created, which is included in Appendix A. Having deconstructed the process the first version of the guide was created, this included an introduction with a short explanation of the model including some specific examples and then a paragraph for each stage explaining what had to be done to complete a definition.

Although we are adopting a manual approach where connotations defining themes can be directly captured from people it is worth considering what aspects of this process could be, potentially, automated. The first stage is perhaps the most obvious candidate for automation, where initial word association might be done automatically in a number of ways. This would include using thesauri and lexical expansions such as WordNet, or using co-occurrence on sufficiently large corpus of documents, or a relevant ontology. All of these approaches have their draw backs when it comes to term expansion, as discussed in the previous chapter, but it's possible that if used appropriately they could at least be useful for the first stage. Similarly term expansion could be potentially used for stage 3 by measuring the semantic similarity between justifications or element names and consolidating sufficiently similar elements. Stages 2 and 5 are perhaps the most difficult to automate, as they require a complex understanding of connotation, which is subjective to individuals (thus our manual approach). It is possible that some very specialist ontologies may already capture some relevant connotations that could be used in these stages, but they are likely to only be useful to defining very specific themes. Automating stage 4 would be a case of simply repeating earlier automated stages. Further exploration of an automatic authoring approach is not covered in this thesis due to identified limitations, while future work in this area could be useful it is considered outside the scope of this research.

3.2.2 Refining the Guide through Expert Review

Having created the first version of the guide based around my deconstruction of the process into a series of steps it was important to test and refine it using human users in order to insure the explanation was sensible and that my articulation of the process would yield valid definitions. To do this four expert reviews were organised. Four individuals experienced as computer scientists in modelling abstract concepts had the model and process explained to them briefly and were then given the guide to work through and create a definition for the theme of 'danger'. As they performed the process they were

interviewed on the decisions they made and what they thought of each section of the guide. This was then used to refine the guide before the next expert review which was conducted in the same manner but with a different version of the guide. The purpose of this process was to iteratively refine the guide to make it easy to use and to also improve the way it articulated the rules of the model so that the definitions it yielded were as valid as possible. Finally a review was performed with a non-expert, a person with an understanding of the idea of themes but no experience in modelling concepts in order to provide an insight in refining the guide for non-experts.

3.2.2.1 Findings

Having completed the expert review process a final version of the guide was developed. Throughout the process a variety of issues occurred with the guide, which with each iteration I attempted to find solutions for. The majority of these issues could be classified under three types; Language, Model, and Stage 3.

Language issues included a variety of problems with semiotic terminology, which at times confused the process. Model issues included problems in attempting to articulate the structure and rules of the thematic model, which was sometimes misinterpreted leading to invalid definitions. And finally, Stage 3, refers to problems with the third stage of the guide where authors group together related elements. As the process of building the definitions is such a subjective one it is difficult to form exact rules for how to go about this and as such authors struggled with how to approach this problem.

There were a variety of solutions I attempted to solve these issues during the process of developing the final guide. Unsuccessful solutions included:

- **Glossaries** aiding users in understanding semiotic language were ineffective. Users remained concerned as to whether they were correctly connecting elements despite the inherent subjectivity in the process.
- **Elaborate diagrammatic examples** lead the authors to simply try and recreate the diagram for their theme rather then follow the process independently. In such situations it became apparent that the authors needed more guidance but that the influence of examples should be minimised.

However some solutions were successful at mitigating the problems experienced by the authors:

- **Replacing semiotic terminology** with plain English descriptions of what was expected of the users. This relieved the formal feel that was making the particular subjective early stages difficult. For example, by using “implied” instead of “connotes”.

- **Simple textual examples** were used throughout the guide which gave authors some idea of what was expected at the end of each stage without influencing their decisions too much.
- **Forms** were added to the guide with instructions for use. These guided the users in what was expected of them at each stage and to help them realise how their subjective definition fits into the rules and structure of the thematic model.
- **Justifications** for connotation relationships were added as a part of stage 3. These encouraged the authors to consider the role each element played in connoting another element. This gave a focus around which to consider which elements should be grouped together in different ways making the stage more systematic.

The resulting guide was found to be satisfactory after a final review using a non-expert user revealed no significant problems. The process of creating the guide highlighted some important difficulties in formalising a process to capture subjective knowledge, however the successful solutions experienced great success with subsequent expert reviews after they were made leading to better definitions. The resulting guide can be found in Appendix A.

3.2.3 Evaluation of the Guide

Having created a guide the next step was to test to see whether a community could use it collaboratively to contribute definitions of themes. To do this I arranged for an experiment where a selection of 15 non experts (in the sense that they had no experience of modelling abstract concepts) would use the guide to form definitions in the terms of my model for one of five predefined themes. Their definitions would then be analysed to ascertain whether they were first of all valid (according to the model), and if not what part of the process had been wrongly interpreted and led to an invalid model. This would uncover whether it was possible to capture peoples subjective understanding of themes in a usable form and also whether the guide was sufficient to enable the process.

3.2.3.1 Methodology

The participants were all volunteers from the English department of Southampton University. This made it likely that they were inexperienced with the process of formally modelling a concept such as a theme as elements but at the same time having a well-grounded understanding of what themes are. The themes selected for the participants to define were ‘winter’, ‘spring’, ‘family’, ‘celebration’, and ‘danger’. The first four themes were selected as themes used in the original development of the model and ‘danger’ was selected as the theme used in the process of creating the guide. This meant that if necessary I would have definitions created by an expert for the themes available to compare

results to while still providing a selection of varied, sometimes contrasting, themes of various levels of complication (based on our own experiences of defining them).

The participants were invited to attend one of three focus sessions, each of which was approximately an hour long. In the session the students were given a very brief introduction and created their definitions. It was important to give the participants an understanding of what they were going to be doing but at the same time, as part of what I was testing was that non experts could create definitions using only the guide, I didn't want to train them in the process in a way outside what the guide offered. To accomplish this the sessions began with a five minute introduction explaining what was asked of them but no more. Participants were then given a guide and given a theme to define, the themes were distributed from a deck to insure it was random who received which theme but that there would also be an even distribution of themes attempted across the experiment. Participants were invited to ask questions but answers were given strictly to clarify the task not to influence the decisions made in their definition which were collected and filed for analysis when the participant felt they had finished. The experiment was also passed through the departments ethics committee and granted approval, the approved application can be found in section E.1 of appendix E.

3.2.3.2 Results

The following table pictured in figure 3.4 summarises the findings, displaying which definitions were valid and what stages users had in some way struggeled with.

Definitions were labelled as valid as long as they structurally complied with the models rules, note in some cases this led to models that were of dubious semantic quality, but were nonetheless valid. Whether the definition was valid or not the notes and forms returned by the participant were analysed for signs as to which stages they struggled on or had questionable results for, the relevant stage to which the participants struggled is also noted on the table.

The results show that just over 50% (8 out of 15) produced valid definitions and each theme was defined in a valid manner at least once. However it also shows that all the participants except two struggled with the process or produced questionable results for at least one of the stages. Of these stage 4 seems to cause the most problems, followed by stages 2 and 5. It is also to be noted that every participant who produced an invalid definition struggled with stage 4 and this was often the reason why their definitions were invalid.

Participant	Theme	Valid?	Problem Stages
1	Spring	No	4
2	Family	No	4
3	Danger	Yes	2
4	Winter	Yes	3
5	Celebration	No	4
6	Family	Yes	4, 5
7	Spring	No	4, 5
8	Danger	Yes	4, 5
9	Celebration	No	2, 3, 4, 5
10	Winter	No	2, 4, 5
11	Winter	Yes	
12	Danger	No	2, 4, 5
13	Spring	Yes	2
14	Family	Yes	2, 3, 4
15	Celebration	Yes	

FIGURE 3.4: Summary of guide evaluation experiment results

3.2.3.3 Analysis

The results lead to a variety of conclusions; first of all the definition process appears to be either an expert task, or the guide by itself is insufficient to support users. Nearly half of the participants produced invalid models and of those that did produce valid ones only two didn't struggle in some fundamental way with a part of the process. However the fact that valid models we're produced demonstrates that it is indeed possible to capture people's subjective definitions of themes although the results suggest that the guide is perhaps insufficient if we were to rely on a wider community building a base of definitions.

The difficulties encountered by participants seem to largely be connected to their inexperience of modelling a concept as a series of elements rather than to do with the thematic model, however it is to be noted the modelling of such subjective information faces its own challenges. The abstract ideas involved in identifying key features, modelling, and classifying them even with the help of the guide seems difficult to those not used to it. For example, the stage participants most frequently had trouble with, stage 4, was mostly due to difficulties surrounding the idea of recursive expansion rather than specifically the model. Definitions, such as the one produced by participant 1 shown in figure 3.5, often expanded sub-themes on the first layer but not subsequent sub-themes. This led to many models that were invalid simply due to being incomplete as sub-themes were left with no elements connoting them. This can also sometimes lead to definitions

including only partially expanded themes as shown in the definition from participant 6 in figure 3.6. While these might be valid they are of poor quality including themes not fully explored.

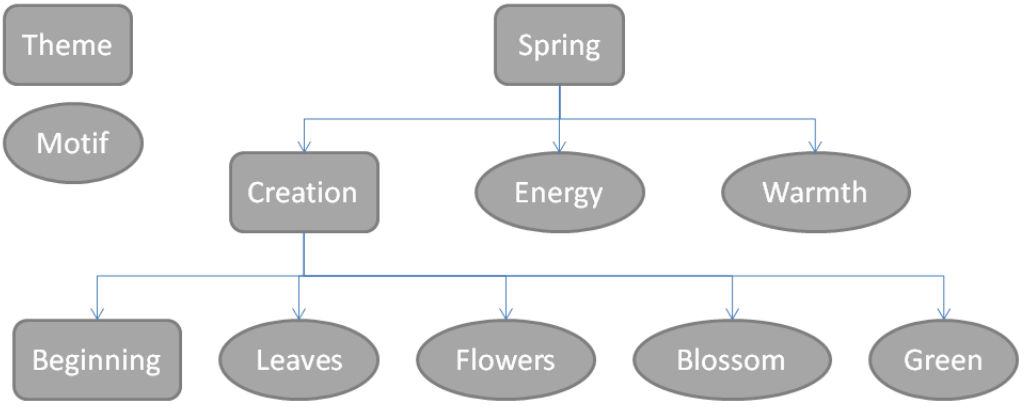


FIGURE 3.5: Definition from participant 1

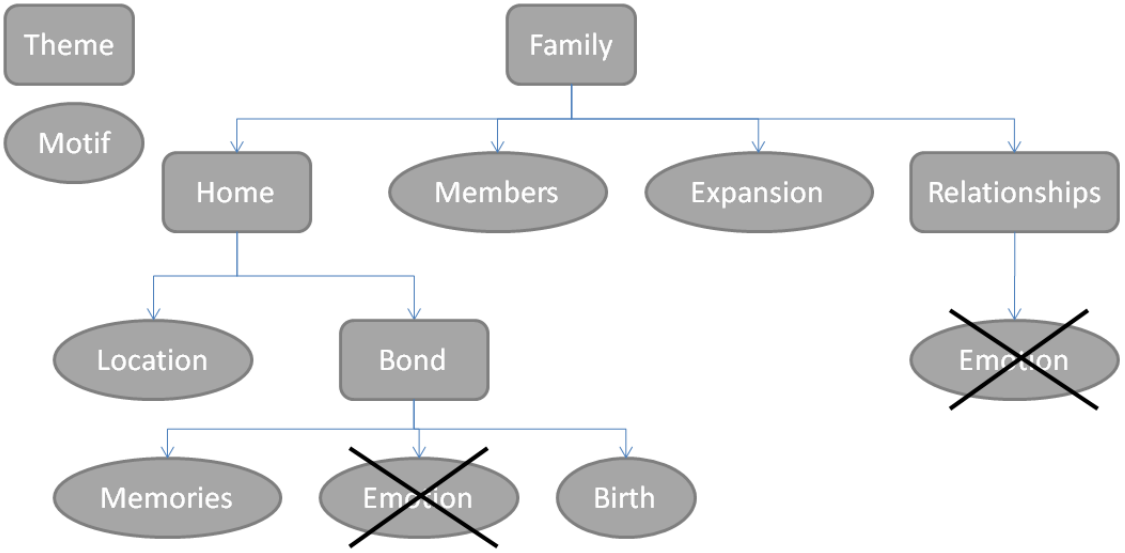


FIGURE 3.6: Definition from participant 6. The crosses denote elements the participant labeled as associated.

This also, as a knock on effect, caused some participants to struggle with stage 5 as elements that would likely be labelled associated and refactored/removed were not due to them not being fully expanded and as such their relevance was left unexplored. This is best shown in the definition from participant 12 shown in figure 3.7 where 'oppressive control' which could quite possibly include elements not relevant to the parent theme is left in and unexpanded. However other problems in stage 5 also seemed to stem from a lack of understanding of modelling concepts. In the definition from participant 6 shown in figure 3.6 the participant labelled the element 'emotions' as irrelevant but did not remove or refactor its parent elements of 'relationships' or 'bond' (and by extension 'home'), it is possible that this is a simple oversight but it could also show that par-

ticipants fail to understand the idea of parent and child relationships between elements within a model leading them not to understand that in such a case the parent is affected.

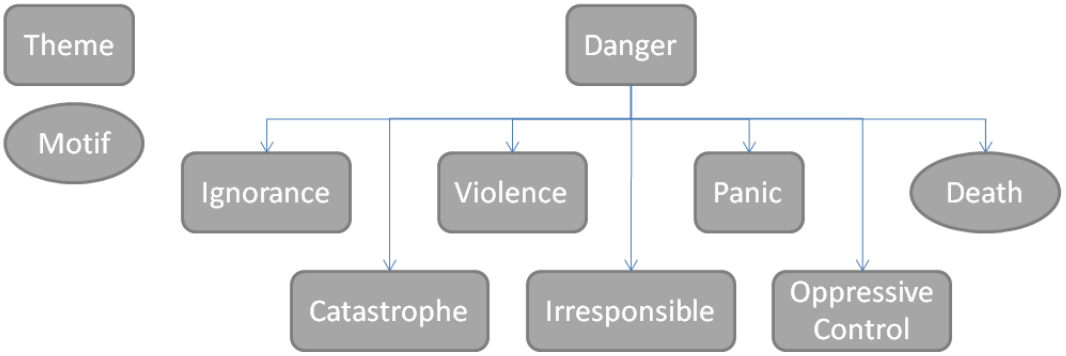


FIGURE 3.7: Definition from participant 12

However, not all the problems participants encountered can be attributed to participants failure to understand modelling concepts. There are much more subjective results that reveal users potentially struggling with the rules of the model. In some cases participants show a potential misunderstanding of the rules explained in stage two by making classifications of elements as themes or motifs that would be difficult to defend. The definition used by participant 12 in figure 3.7 shows this in that 'Death' is classified as a motif despite the fact that stage 2 states that high level concepts should be classified as themes. It is however to be noted that it is difficult to draw conclusions here, as this stage is such a heavily subjective process where the line between theme and motif may seem blurred. However results do seem to suggest that there is much potential for conflicts between definitions from this stage, and that some participants definitely struggled with it. In a similar way a few participants struggled with stage 3, failing to group together similar elements and ones that shared the same justification, figure 3.8 shows the definition from subject 4 who has failed to group together hats and scarves despite them both sharing the justification 'worn during', although in this case a rule of the model has been directly over looked in some cases problems in this stage can be just as subjective as those with stage 2 when participants were trying to decide whether elements were similar enough to be grouped.

From this analysis we can draw three main findings:

- Non-experts can struggle with the principles surrounding modelling a concept that are core to this process.
- Participants failed to realise when they had broken a rule of the model presented to them through the guide.
- Some of the heavily subjective decisions necessary in the process are likely to lead to conflicts in definitions and potentially definitions that are valid but of

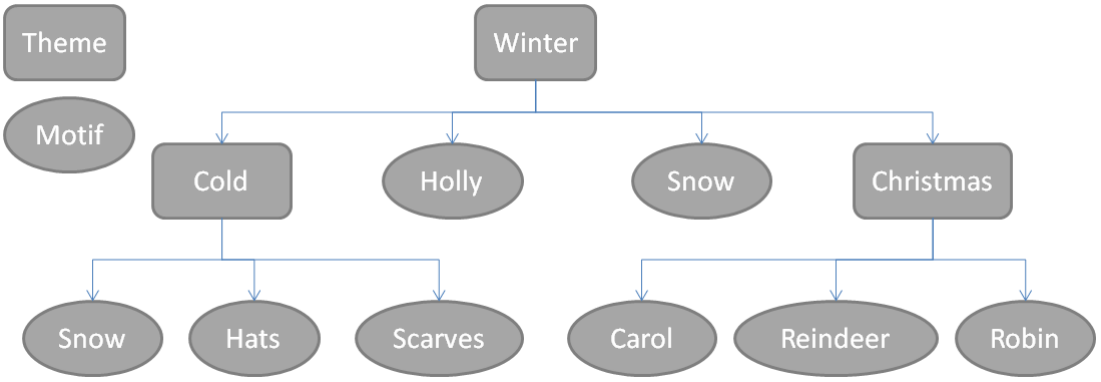


FIGURE 3.8: Definition from participant 4

semantically poor quality.

The first two are related in that the lack of understanding of modelling principles is, in a majority of cases, what causes rules to be broken. This shows that the guide by itself is insufficient as it cannot ensure that users follow the rules of the model. While a full investigation of the challenges facing authoring in semiotic models is beyond the scope of this thesis there are a number of potential solutions we can discuss.

A potential solution to this would be the building of a tool to accompany the guide. Such a tool would use a formal schema of the rules of the thematic model to validate definitions on the fly directing users to errors in their creations. This kind of tool could potentially be applied in many different ways; as well as standalone authoring tool it could potentially be used alongside tagging systems to allow them to thematically expand tags they used for their media, validating the subsequent definitions.

This would however not solve the final issue as the subjectivity of these potential errors makes them impossible to detect. However, if a large community base of definitions was to be built, a set of super definitions based on the submissions of all participants, then it is possible that such anomalous conflicts could be detected by frequency of occurrence and removed. For example, referring again to the definition of participant 12 in figure 3.7, if the majority of definitions did not classify 'death' as a motif then such an assertion by a contributor would not be accepted into the super definition. Formation of a super model could also solve the problem of valid models that are only partially expanded such as that of participant 6 in figure 3.6 as subsequent definitions could fill out and complete any partial expansions. The super model would then form a definition by consensus that could be used by any system utilising the thematic definitions, such as the TMB (the Thematic Montage Builder, explained fully in chapter 4). Potentially the way these definitions are built could allow for weighting of thematic components based on the frequency of their occurrence in the set of all authors definitions for that theme, this might improve the performance of systems such as the TMB but modifications would first be necessary. The supermodel would also make partial definitions

useful, as although these definitions by themselves are not complete they do contain correctly formed elements with valid relationships to others, as such the contribution to the supermodel would not need fully valid complete definitions but could simply use a validator to only use aspects of the definitions that were correct. It is important in this light to remember that while many of the definitions created with the guide were invalid the formation of an authoring tool could reduce this significantly and the use of a super model would make even invalid definitions useful.

3.2.4 Conclusions on the Authoring Process

This experiment sought to assess whether it was possible for untrained humans to capture the thematic definitions necessary for systems to use my thematic model. By exploring the systems and structures of any potential implied meaning beyond terms used for tagging we gain a greater understanding of the way people might use tags on the Web that we can use to improve systems that draw understanding from these tags. Work I present later in this thesis in chapters 4 and 5 shows that expanding queries on a thematic basis can improve the relevance of results of certain queries, this suggests that the implied subtext behind the terms used in queries and the tags people use are important to providing relevant results. The thematic models heavy basis in semiotics makes it very practical to use for semiotic term expansion and initial evaluation of its performance (as discussed later in chapters 4, 5, and 6) seems promising however its reliance on themes being defined by hand means a practical method of capturing these definitions from a contributing community is essential.

While existing knowledge representations might be used as a basis for term expansion they are often limited to the contribution of a small number of experts, and even when this is overcome by using methods that allow for contribution of a wider community or automatic generation they are still normally limited to a specific area. As such something like thematic expansion on a semiotic basis faces very different challenges. The information being modelled is more subjective and less an exercise in modelling specific facts, and there are no experts on what different terms connote. In trying to capture such information it's important to facilitate contributions from a wide range of people, as well as model subjective information from a massive range of concepts in a formal structure of elements and relationships.

The authoring experiment suggests that by using the guide developed it is indeed possible to capture peoples subjective definitions of themes but also that the guide is often insufficient. While over half of the participants produced valid definitions many of them were of a low quality and all participants except two displayed that they had struggled with at least one part of the process. Much of what the participants found difficult seemed to be with the process of modelling an abstract concept in the terms of elements and relationships but there were also points where participants struggled with

the specifics of the thematic model itself.

This would suggest that the authoring process is to some extent an expert task and that the guide, while making it possible in some cases to produce valid definitions, is insufficient on its own to produce quality definitions on a large scale. However, as we all make semiotic connections between terms, it is the current process itself of creating the definitions that is expert, not the understanding of the expansion, and as such the process needs to be further improved that it can be undertaken successfully by a wider community. There are several possible courses of action that could mitigate these problems. The creation of an authoring tool backed up by a formal schema describing the rules of the thematic model could guide authors in applying the rules to their definitions more closely than the guide by itself. It would also be able to draw their attention to when a definition contains mistakes by validating in on the fly so that they can correct it. The quality of final definitions used by systems could also perhaps be improved by the creation of a system that forms definitions submitted by the community into super definitions. This would solve the problem of incomplete definitions by filling them out with the assertions other authors have made as well as improving the semantic quality of the definitions by resolving conflicted definitions by way of popularity. Furthermore, a system that forms super definitions would make even invalid definitions useful, so long as they contained some correctly formed elements and relationships.

This experiment has shown that there are challenges with creating useable semiotic definitions in theory that are similar to those faced in ontology or taxonomy creation, despite the cognitive differences between subjective thematic models and objective ontological ones. We have also discussed how these problems might potentially be mitigated through support tools or the community-driven creation of super-definitions. A thematic approach to term expansion affords a system a greater understanding of what users imply when they make a particular query or choose a particular term for a tag, and in my future work we hope to explore how this subtextual understanding can be utilised at a larger scale to help applications find information that is more relevant to their users.

3.3 Thematic Integration with Narrative Generation

One of the important potential contributions of a thematic model is that it may be used to give narrative generation systems a machine understandable representation of theme [51]. This can be used to imbed or emphasise themes within the narratives they generate, potentially leading to stronger thematic cohesion. Referring back to the division of narrative generation illustrated in figure 2.4 in chapter 2, we can explore the possibility of a thematic systems involvement at different levels of narrative generation; story, plot, and presentation. Themes are intangible concepts, a subtext rather than a

core focus of the narrative, and for this reason it seems at first that narrative generation would benefit from thematic involvement at the presentation level. Here, themes could be connoted or emphasised through how the system presents core features of relevant motifs, emphasis through presentation of these features leading to a more prominent theme. However this is a process that could potentially fail if there were no features relevant to the desired theme in the given narrative, the system might find that at the presentation level a thematic system might only be able to offer from a subset of themes (those already present in the narrative at some level).

At the story level of narrative generation a thematic systems involvement would be in some ways the opposite of its involvement at a presentation level. Instead of offering elaboration on existing narrative features at the story level a thematic system would generate additional narrative elements based on a shopping list of required motifs for the desired themes. This way themes would become apparent through the presence of certain story elements that connoted the desired themes. This could potentially fail however if the systems plot generation did not make use of the thematic story elements or they were not properly exposed possibly leading to absence of key motifs. Also, such an approach could damage the generated narrative more than help it potentially flooding the system with elements irrelevant to the plot. For some systems story generation integration is not always an option, at least on a fully autonomous level, with many systems generating plot out of pre written and defined story elements. These semi-automatic approaches to story generation require a very different approach perhaps with thematic guidance on the creation of these elements as supposed to influencing the automatic generation process in others.

At the plot generation level thematics could play a role in the story selection of narrative elements as well as the way relationships build. The first part of this would be similar to how the TMB prototype builds photo montages (as described in chapter 4) in that a list of desired motifs would be compiled and this would be used to thematicly score potential story elements and as such influence their selection and inclusion in the plot. Also, the relationships between story elements and actions of elements could in turn be factored in as features that denote motifs, as such potential actions at the plot generation of the story could be thematicly scored influencing what occurs. For example a story in which violence is a desired theme might see the protagonist kill the antagonist rather than banish or imprison them. However, like inclusion at the story level it is possible that heavy thematic involvement could damage the plot itself, making its involvement a dangerous balancing act, potentially forcing plot actions that damage the quality of the narrative. Furthermore like involvement at the presentation level, a lack of complete control over the story elements could potentially restrict available themes.

Character centric narrative generation approaches start by simulating the content of the narrative itself, modelling characters, locations, entities, and events. While this potentially makes it less discrete at the story generation level through the process of

plot and presentation, character centric approaches still go on to generate natoms that contain features and by extension denote motifs. An integration would have to seek to ensure that certain features were planted in order for the themes to become apparent in the finished narrative. To do this involvement at the story level seems easiest as this is where the elements present within the narrative are generated. However, as a more semi-automatic approach with predefined elements is more common than automatic generation at this level in character centric approaches it could be difficult to integrate a thematic approach with the prewritten characters. At the plot and presentation levels an integration seems more possible, potential character actions and story events can be thematicly scored to influence actions taken to be conducive with desired themes and then presented in a way that emphasises the relevant thematic content. Character centric generations frequent use of game engines means that integration at the presentation level may be easier where knowledge of the entities present in a particular scene is much more exact then in natural language. However as already discussed a reliance on integration at these levels potentially limits the available themes.

Author centric approaches are heavily based on structures and largely concerned with the authoring process rather than modelling the content of the narrative. The story generation process for some is about composing a pool from large collections of potential natoms, often from the Web, based on their relevance to required parts of the narrative structure. At the story level author centric thematic integration could be a relatively simple process of scoring potential segments to be included in the pool of potential natoms based on their thematic relevance. At the plot level, integration could be similar to character centric approaches, in that elements selected for exposure could be chosen based on their thematic qualities potentially needing thematic selection rules to be written for the system. Should an author centric system present a stream of natural text from other sources it could be difficult to integrate at the presentation level as the systems understanding of the content is limited by its meta data understanding of the natoms. However, for those systems that use templating or selected pre authored text presentation, using thematics becomes more feasible where techniques such as omission or emphasis (spatially or visually) can be used to highlight relevant or hide irrelevant segments to help connote a theme.

The possibilities apparent from this investigation are summarised in the table in figure 3.3. Decisions and selections made in generation may be influenced thematically by making the objective of the decision thematic as well as for plot objectives. Further thematic integration can be achieved through emphasis at the presentation or plot levels and other presentation choices such as style may have an influence that could be worked in favour of desired themes.

		Narrative			
		Story		Discourse	
				Plot	Presentation
		Semi Auto	Auto		
Generation Approach	***Character Centric***	*Uses Prewritten Characters.* Could be supported by thematic writing tools.	*Goals and Rules are generated.* Generation process influenced thematically.	*Characters and Elements Selected. Rules applied. Actions chosen.* Thematically influenced selection of characters and elements as well as actions chosen.	*Generation of media. Revealed narrative chosen.* Emphasis of relevant elements and relationships. Choice of style.
	Compromise	*Uses Prewritten Characters and purpose built Director Agent or an author rule set.* Could be supported by thematic writing tools and have a thematically influenced author/director.	*Both characters and author goals and rules generated.* Goal and rule generation thematically influenced.	*Characters and Elements Selected. Rules applied. Actions and Events chosen. All Influenced and vetted by Director Agent or Author rule set.* Thematically influenced selections of elements and actions. Thematic vetting by director/author.	
	Author Centric	*Material generated or collected.* Thematically influence which resources collected/generated.		*Populating/Building Structure.* Influence selection and positioning of resources. Influence structure built.	

FIGURE 3.9: Summary of potential integrations of a thematic system with narrative generation

Chapter 4

Themed Photo Montages

In the previous chapter I presented a thematic model of themes inspired by Tomashevsky's notion of themes and based on semiotic relationships between terms. In this chapter I present a software system called the Thematic Montage Builder (TMB) that uses thematic models for term expansion, and show how it can be used to build themed photo montages (based on a user query) from tagged photographs from the image site Flickr[1]. I also present a comparative evaluation that compares montages created with the TMB with those created with simple keyword search.

4.1 TMB

To evaluate the effectiveness of thematic definitions constructed using the model a prototype was built which utilised them. The prototype built simple photo montages around titles which contained a subject and desired themes. In doing this I can ascertain firstly, if definitions made in terms of the model could be used by a system to successfully embed that theme and secondly to compare it to other systems in order to measure any improvement in relevance to thematic queries. As a photo montage builder that utilised the thematic model this initial prototype was called the Thematic Montage Builder or TMB.

The prototype itself was written in Java with a simple JSP front end. As a source of natoms (photos in this case) the photo sharing system Flickr was used due to its large amount of readily available and well tagged items. Folksonomies such as that made available by Flickr have been demonstrated to offer meta data on items of a higher semantic value as opposed to collections with automatically generated data [2]. As such, Flickr's tags are likely to reflect the features of each item. The definitions of themes that the prototype would use were written in XML, each file representing a

[1] http://www.flickr.com as of 21/07/2011

thematic element (either a theme or a motif). Definitions for themes listed the motifs with which they shared a connotation relationship and definitions for motifs listed the features that denoted them. For this prototype, four root themes were authored by hand following the defined authoring method. The themes selected for the initial experiment were *Winter*, *Spring*, *Family*, and *Celebration* and their structural definitions can be found diagramatically in section B.1 of appendix B.

The prototype generated montages by taking a desired length (number of images), a desired content (subject of images), and a desired list of comma separated themes. These are explictly seperate fields that the TMB uses, content defining the search string for building the initial image corpus, and for themes a comma seperated list of already defined themes used to form the montage. The TMB searches Flickr for the desired content and forms a base corpus (effectively a fabula) using the top 30,000 images. The thematic quality (its relevance) of each image with respect to the desired themes is then calculated and the top N images are returned where N is equal to the desired montage length.

The thematic quality of each image is calculated based on the features present. Each tag is considered to be a feature and using this, each image's *component coverage* and *thematic coverage* is calculated. How these are calculated and how thematic quality is calculated from them is presented in equations (4.1) below. TQ is thematic quality, TC is thematic coverage, CC is component coverage, T is the number of desired themes, C is the sum number of components (elements, themes or motifs, that directly connote a theme) of all desired themes, and t and c are the number of themes or components respectively for which the image has a relevant feature. A feature is considered relevant if it directly denotes a motif that is either a component or through a chain of connotation later indirectly connotes the component or theme requested.

$$\begin{aligned} TC &= (t * 100)/T \\ CC &= (c * 100)/C \\ TQ &= (TC + CC)/2 \end{aligned} \tag{4.1}$$

In summary the final thematic quality is expressible as a percentage and is based on how many of the desired themes the image is relevant to as well as how relevant it is to each themes top level thematic components.

4.2 Evaluation Against Flickr

Within this section I explain the methodology, results, and analysis for our evaluation of the TMB producing themed photo montages compared to Flickr's keyword search.

4.2.1 Methodology

For this experiment I sought to evaluate how effective a system utilising the thematic model would be. The experiment would need to evaluate whether the system was able to generate results that successfully connoted the desired themes. It would also be important to compare these results to those produced by an existing system to see what value, if any, the thematic approach afforded. As the TMB is based in part on Flickr I elected to compare it to Flickr's keyword search. As well as comparing the thematic relevance of both approaches for individual images I was keen to see how well the thematic system performed in a more narrative context of many 'natoms'; in this case a photo montage. From this we see two key objectives:

- To evaluate the effectiveness of the TMB in selecting images connoting desired themes in comparison to simple keyword search.
- To evaluate the effectiveness of the TMB in the more narrative context of generating montages.

The evaluation asked participants to rate images individually and in sets (montages) according to how relevant they were to a given title. These titles deliberately contain a content subject and a desired theme(for example, 'London in Winter' where 'London' is the content and 'Winter' the desired theme). The images and sets were generated for four different methods to be compared:

- *TMB*: Using the TMB and the Flickr API to search by subject and select by thematic quality
- *Flickr*: Using Flickr to search by subject and theme
- *BaseL(ow)*: Selecting images from Flickr at random
- *BashH(igh)*: Using Flickr to search tags by subject and filter manually

In this way I hoped to compare the performance of the TMB with keyword search on Flickr, and place both of these methods in context by comparing them to the base cases of random and hand-picked samples. For each test the user would be presented with two titles and the images for the test for that title. Depending on the test these would

be presented either individually or in groups. The users were then asked to rate the images (either individually or as a montage) from 1-5 on their relevance to the title. To ensure the data was representative I chose titles composed of contrasting themes and fabulas (such as taking a fabula built on the content of 'factory' and theming it with the theme of 'family') as well as regular or complimentary themes and fabula pairings. In each test the users rated images for two titles, one with regular pairings and the other with contradictory. At the end, the users were also asked to rate images for titles that included more than one theme in order to measure the system's performance under more complicated requests. This resulted in four tests;

- Test 1: Images presented individually for two titles both containing one theme (one title with a contradictory theme/fabula pairing, and the other with a regular pairing)
- Test 2: Images presented in their montages for two titles both containing one theme (one title with a contradictory theme/fabula pairing, and the other with a regular pairing).
- Test 3: Images presented individually for a single title containing two themes.
- Test 4: Images presented in their montages for a single title containing two themes.

In order to make the evaluation fair I presented the single image tests first, so participants would not already have associated them with a group. The images on the single image tests were also randomly shuffled and for the group tests I randomised the order in which different montages as well as the images within the montages appeared. I also added a restriction on image groups that no more than one image would be allowed per Flickr author - this is because image sets published by a single author are often taken as a part of a set and have natural cohesion and would artificially seem to be stronger, more coherent, montages. Finally users were only allowed to take the evaluation once, a unique evaluation link for each user was given out per email address and only the first result for a given ID was recorded. The methodology for this experiment was granted ethics approval from the universities ethics committee, the approved application can be found in section E.2 of appendix E.

Having finalised the methodology I selected the titles for which to generate the montages based on the themes that the TMB currently handled. The titles selected met the needs of the experiment; four single theme titles including two with regular theme/fabula pairings and two with contradictory pairings and two multiple theme titles. The titles chosen for single themes were 'London in *Winter*', '*Celebration* and Earthquake', '*Spring* Picnic', and '*Family* Factory' and for multiple themes 'My *Family* in New York at *Winter*', and '*Celebrating* the New House in *Spring*'. For these titles the TMB used

the thematic definitions created for the prototype which are presented diagramatically in section B.1 of appendix B.

To test the technical and logistical aspects of the experiment a pilot study was performed with a small number of participants. My pilot study was performed with 22 users. While this is a relatively low number of people it still gave us a large amount of data, as each user was asked to rate 40 images and 4 groups for each of the 4 sources. This resulted in 880 data points for single images and 88 for groups, enough for early indications of quantitative significance to emerge (which I measured with a t test). The pilot study was important to test the stability of the evaluation and also find if there were any significant improvements that needed to be made to the test. On completion of the pilot I found the test to be stable and required only minor rewording of instructions to make the test easier to understand. The results of the pilot study are present in a published paper [52], and offered very similar initial results and conclusions as the full experiment as explained below.

4.2.2 Results

The full evaluation had 108 test subjects. All of the images used in this experiment can be found in section D.1 of appendix D. The mean rating of natoms from the TMB is higher than from a keyword search (Flickr) in both single and grouped (montaged) images. Figure 4.1 and Tables 4.1 and 4.2 show the data and t-tests for single images. Figure 4.2 and Tables 4.3 and 4.4 show the data and t-tests for grouped images. The hypothesis that the TMB selects natoms more relevant to the title than a keyword search is true with less than 0.0005 probability of error for both group and single images.

Set	1	2	3	4	5	Total
TMB	1437	944	857	609	393	4240
Flickr	1558	1019	812	513	346	4248
BaseL	3176	638	277	122	53	4266
BaseH	588	725	957	998	988	4256

TABLE 4.1: Single Images Rating Frequency of TMB and Flickr experiment

Set	Mean	SD	Variance
TMB	2.428	1.329	1.766
Flickr	2.310	1.295	1.678
BaseL	1.414	0.832	0.693
BaseH	3.252	1.350	1.822

t=4.227, df=8486, p=0.0005

TABLE 4.2: Single Images Rating Statistics of TMB and Flickr experiment

At first glance the difference between the TMB and Flickr only appears to be slight however it must be seen in the context of the difference in results between a best case scenario (human selection: BaseH) and a worst case scenario (random selection: BaseL).

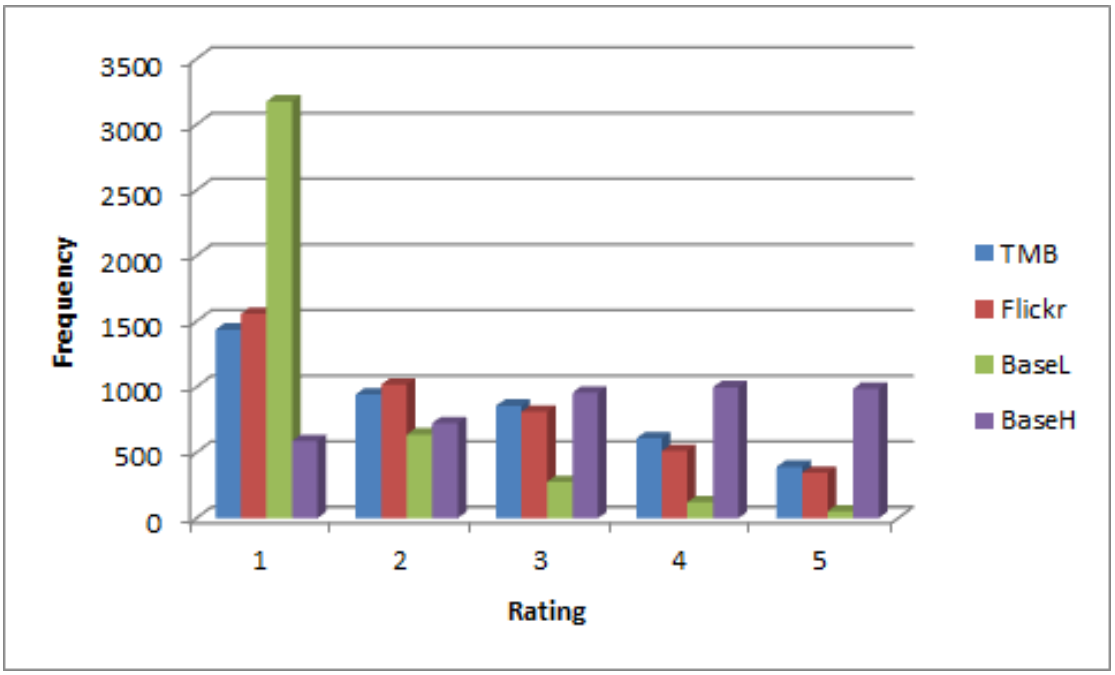


FIGURE 4.1: Single Image Rating Frequency of TMB and Flickr experiment

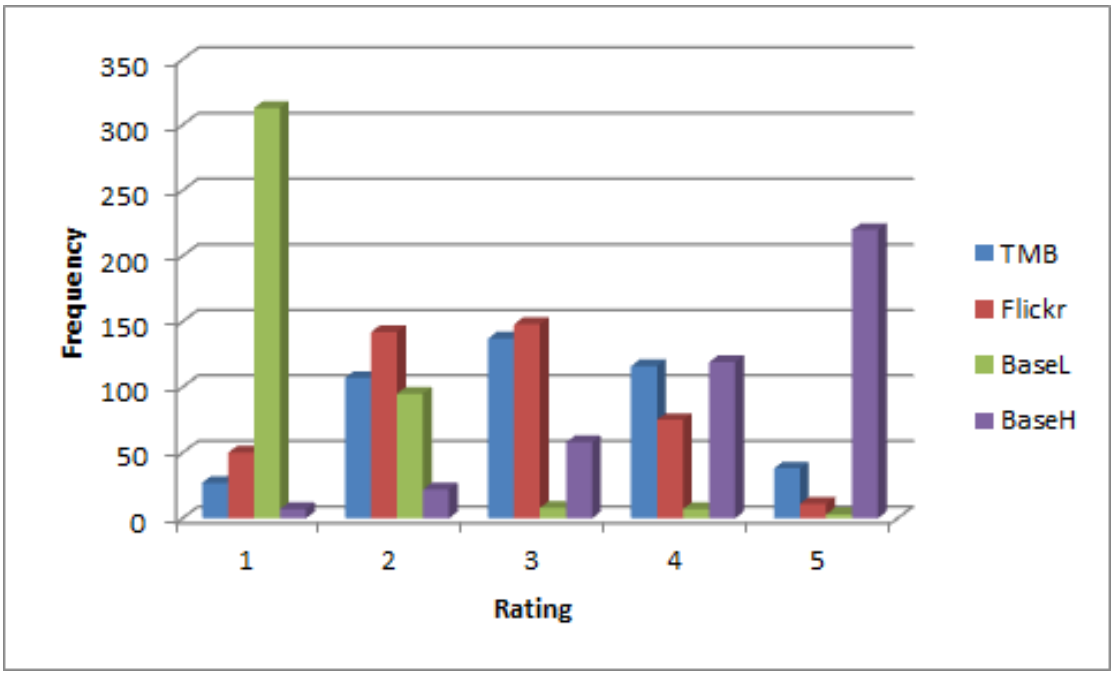


FIGURE 4.2: Grouped Image Rating Frequency of TMB and Flickr experiment

Set	1	2	3	4	5	Total
TMB	27	106	135	116	37	421
Flickr	50	141	147	73	11	422
BaseL	311	93	8	7	3	422
BaseH	7	22	57	119	217	422

TABLE 4.3: Grouped Images Rating Frequency of TMB and Flickr experiment

Set	Mean	SD	Variance
TMB	3.071	1.064	1.133
Flickr	2.654	0.983	0.967
BaseL	1.336	0.668	0.447
BaseH	4.225	0.979	0.958

t=5.902, df=841, p=0.0005

TABLE 4.4: Grouped Images Rating Statistics of TMB and Flickr experiment

Figures 4.3 and 4.4 show the relevant means in a way that they can be compared (it is worth noting these graphs, and all similar graphs in this thesis, do include standard error bars, but that they are too small to be seen.). These ranges are rather smaller than we might expect, and in this context the improvement given by TMB is rather more impressive.

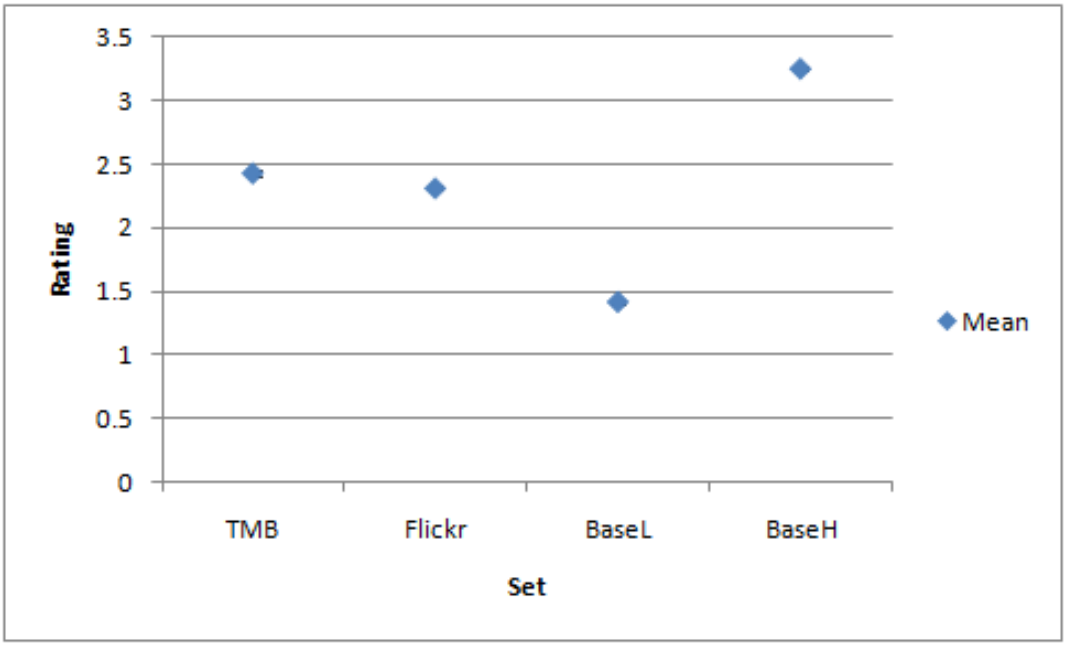


FIGURE 4.3: Single Image Mean of TMB and Flickr experiment

As expected the results also show that the TMB proves better in a montage context with significance where it can build themes over a group of natoms, a t-test shows this hypothesis to be true with only a 0.0005 probability of error. The data shown in table 4.5 reveals that while both a keyword search and TMB improved when their natoms were presented as a group the TMBs improvement was much more significant, the hypothesis that the TMBs improvement was greater than the improvement of a keyword search in a group context is shown with this data to be true according to a t-test with less than 0.0005 probability of error.

As explained I also wanted to observe how the TMB performed under a range of situations so deliberately included titles that had contradictory theme fabula pairings as well as titles which included multiple themes. Tables 4.6 and 4.7 and figures 4.5 and 4.6

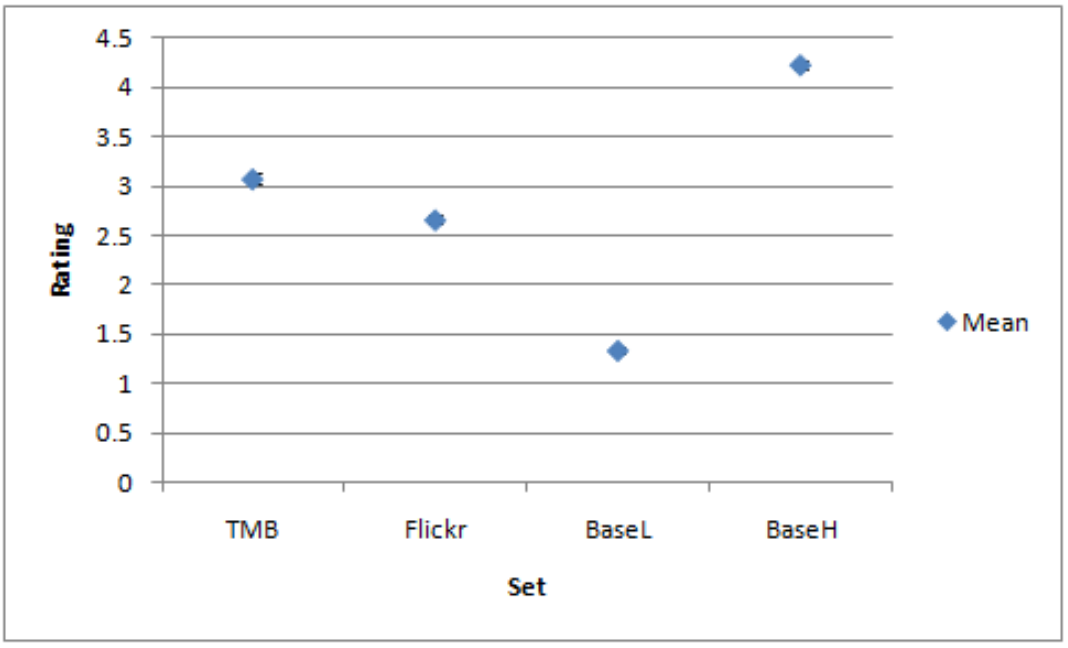


FIGURE 4.4: Grouped Image Mean of TMB and Flickr experiment

Set	Mean	SD	Variance
TMB	0.650	1.420	2.018
Flickr	0.356	1.447	2.095

t=9.328, df=8486, p=0.0005

TABLE 4.5: Grouped Images Improvement Statistics of TMB and Flickr experiment

show the contrast of results for single and grouped images respectively between titles with multiple themes and those with just one theme where as the tables 4.8 and 4.9 and figures 4.7 and 4.8 show the contrast of results for single and grouped images between titles with contradictory theme fabula pairings and regular pairings.

Set	Mean	SD	Variance
TMB Single Theme	2.456	1.399	1.958
Flickr Single Theme	2.496	1.337	1.790
BaseL Single Theme	1.404	0.842	0.710
BaseH Single Theme	3.235	1.427	2.037
TMB Multiple Theme	2.399	1.253	1.571
Flickr Multiple Theme	2.122	1.223	1.496
BaseL Multiple Theme	1.425	0.822	0.676
BaseH Multiple Theme	3.268	1.267	1.606

TABLE 4.6: Single Images Single and Multiple Themes in Title Contrast Statistics of TMB and Flickr experiment

Set	Mean	SD	Variance
TMB Single Theme	2.981	1.135	1.288
Flickr Single Theme	2.849	0.971	0.943
BaseL Single Theme	1.292	0.659	0.435
BaseH Single Theme	4.037	1.109	1.230
TMB Multiple Theme	3.164	0.983	0.968
Flickr Multiple Theme	2.471	0.962	0.926
BaseL Multiple Theme	1.383	0.673	0.453
BaseH Multiple Theme	4.415	0.787	0.619

TABLE 4.7: Grouped Images Single and Multiple Themes in Title Contrast Statistics of TMB and Flickr experiment

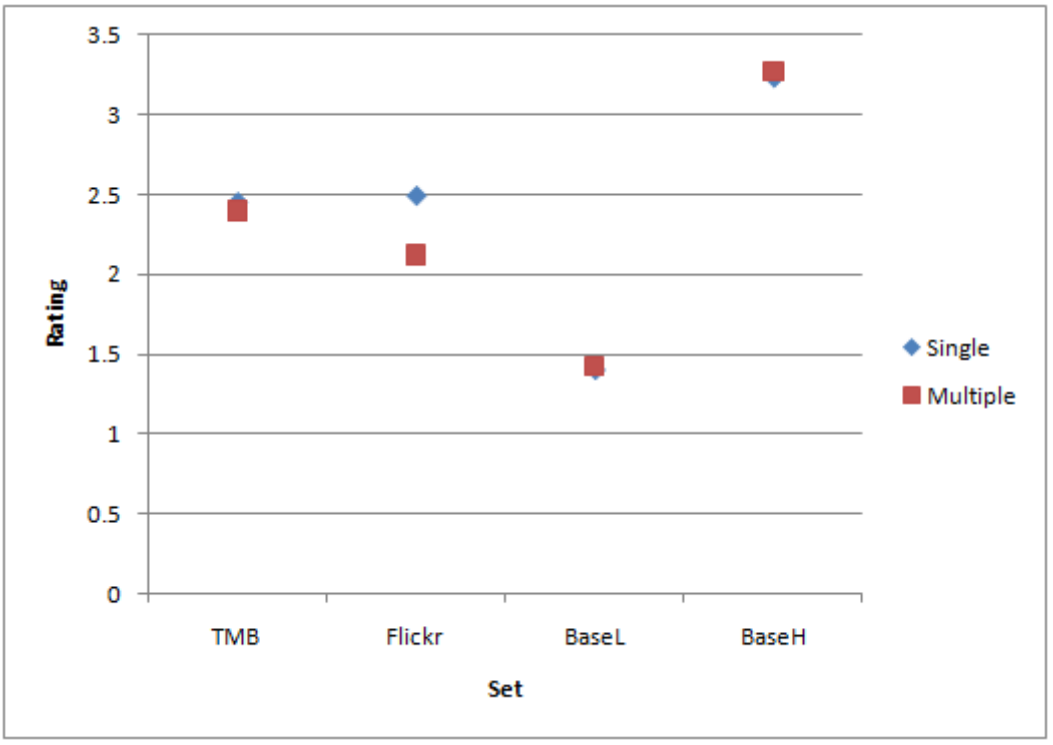


FIGURE 4.5: Single Images Single and Multiple Themes in Title Comparison of TMB and Flickr experiment

Set	Mean	SD	Variance
TMB Regular Theme Fabula Pairing	2.563	1.336	1.785
Flickr Regular Theme Fabula Pairing	2.258	1.282	1.644
BaseL Regular Theme Fabula Pairing	1.409	0.822	0.675
BaseH Regular Theme Fabula Pairing	3.546	1.265	1.601
TMB Contradictory Theme Fabula Pairing	2.028	1.224	1.499
Flickr Contradictory Theme Fabula Pairing	2.464	1.322	1.749
BaseL Contradictory Theme Fabula Pairing	1.429	0.864	0.746
BaseH Contradictory Theme Fabula Pairing	2.370	1.204	1.451

TABLE 4.8: Single Images Contradictory and Regular Theme Fabula Pairing in Title Contrast Statistics of TMB and Flickr experiment

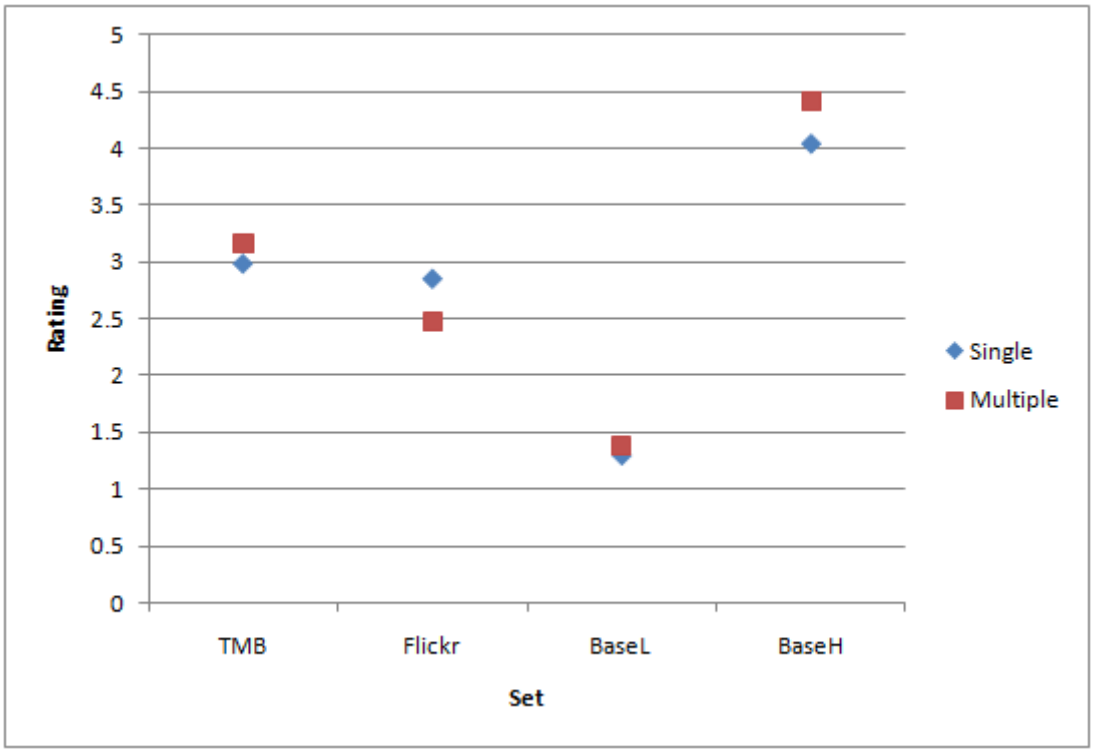


FIGURE 4.6: Grouped Images Single and Multiple Themes in Title Comparison of TMB and Flickr experiment

Set	Mean	SD	Variance
TMB Regular Theme Fabula Pairing	3.332	0.988	0.977
Flickr Regular Theme Fabula Pairing	2.578	0.966	0.934
BaseL Regular Theme Fabula Pairing	1.337	0.656	0.431
BaseH Regular Theme Fabula Pairing	4.534	0.737	0.544
TMB Contradictory Theme Fabula Pairing	2.292	0.894	0.799
Flickr Contradictory Theme Fabula Pairing	2.905	1.001	1.001
BaseL Contradictory Theme Fabula Pairing	1.339	0.702	0.493
BaseH Contradictory Theme Fabula Pairing	3.301	1.034	1.069

TABLE 4.9: Grouped Images Contradictory and Regular Theme Fabula Pairing in Title Contrast Statistics of TMB and Flickr experiment

4.2.3 Analysis

The results show that in both single and grouped images Flickrs keyword search performed worse in titles with single or multiple themes with the one exception of single images and single themes, where Flickr out performed the TMB. The results also show that in both single and grouped images Flickr performs better than TMB for contradictory theme fabula pairings and TMB performs better than Flickr for regular pairings. What this tells us is that the TMB is better at consolidating a wide variety of themes than simple keyword search, this can be attributed to the way the TMB consolidates a thematic request into a single list of relevant features, while keyword search might look for each theme separately the TMB will scores natoms with both themes present highly.

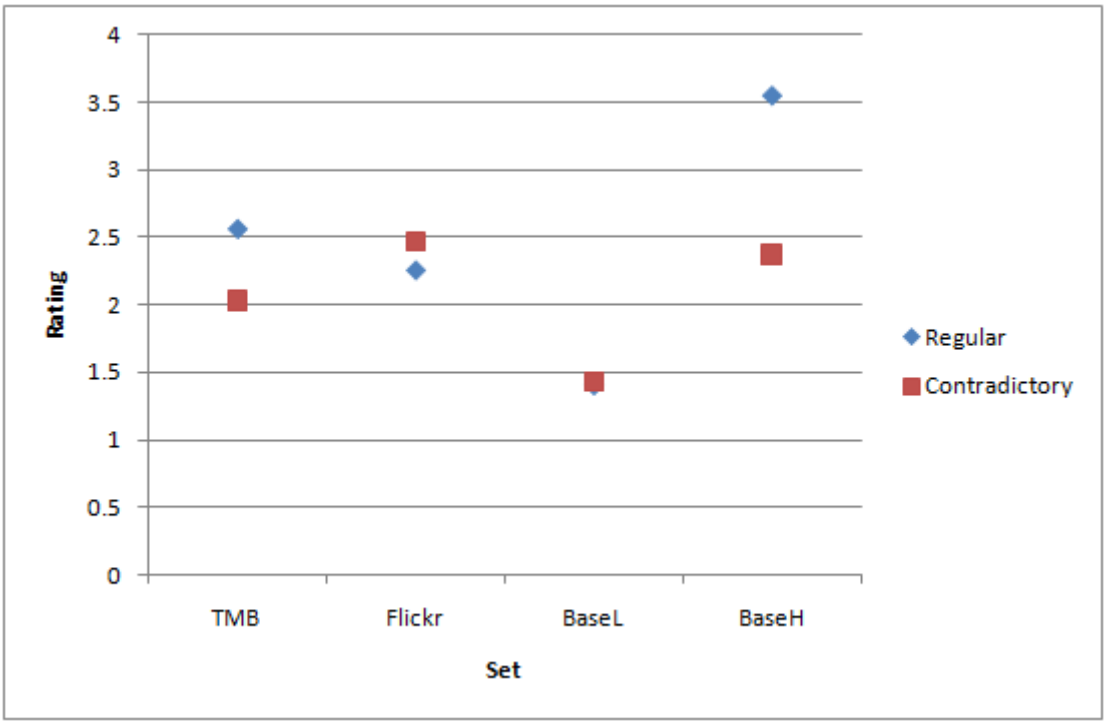


FIGURE 4.7: Single Images Contradictory and Regular Theme Fabula Pairing in Title Contrast of TMB and Flickr experiment

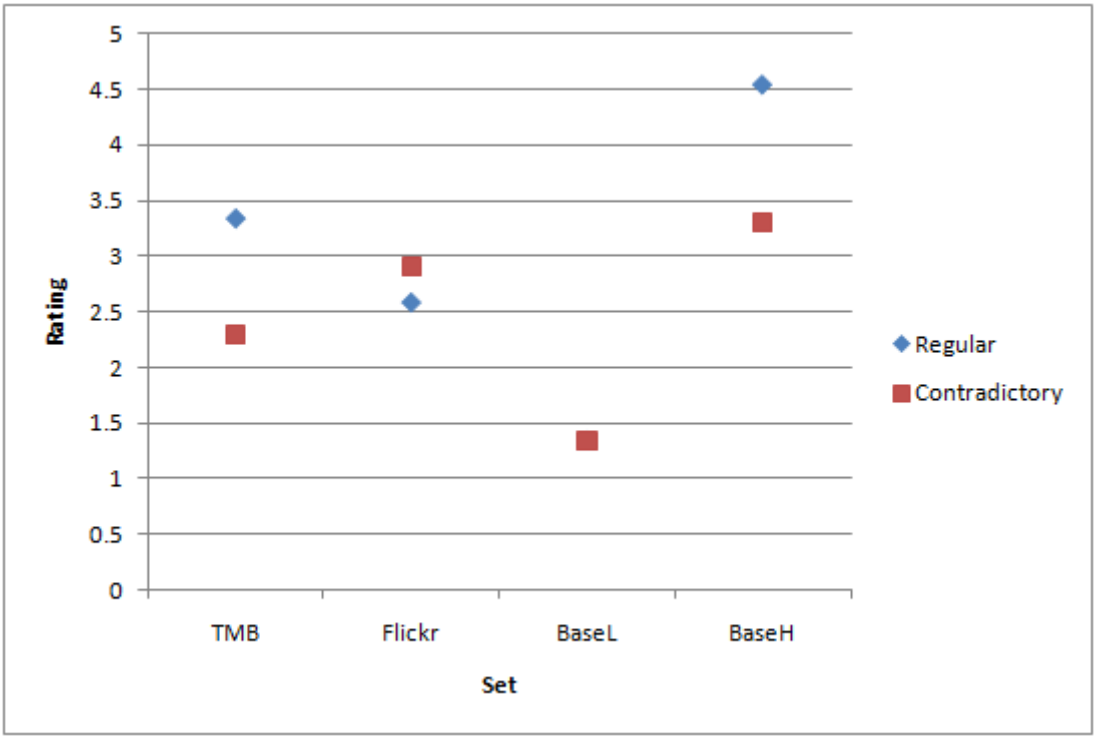


FIGURE 4.8: Grouped Images Contradictory and Regular Theme Fabula Pairing in Title Contrast of TMB and Flickr experiment

That the TMB performed worse for contradictory pairings is also no surprise, because of the way this experiment has been designed with the TMB scoring a finite fabula it's

possible that having built the Fabula for a specific content that it would not contain any features to use in connoting a contradictory theme where as the keyword search retrieves each word separately from a much wider pool of images and as such can pick images for each idea.

These results offer encouraging observations. The TMB seems to be performing better than a keyword search with some significance and furthermore it seems the TMB is very strong within a group context, this could lead us to believe it could perform similarly strongly within a more narrative context (where there are groups of natoms). It is also encouraging to see the TMB is able to consolidate multiple themes as well, although it is an accepted constraint that such a system performs less strongly with contradictory theme fabula content pairings.

4.3 Conclusions

This experiment has shown that a system using the thematic model was able to return results with a higher relevance thanks to its understanding of the thematic part of the query than that of a simple keyword system. It has also shown that it is particularly capable at putting together a montage of images and capable of marrying together multiple themes. Although it is to be noted that in circumstances where the content contrasted heavily with the desired themes it struggles to find the motifs it needs.

This leads us to two significant conclusions. Firstly, that themes are concepts that become apparent through the presence of particular motifs which are semiotically connected to the theme and are more complicated than a keyword associated with them. Secondly, that a system using a thematic model which understands this semiotic hierarchy of concepts is more capable of understanding the thematic value of content. The TMB's improved performance over keyword search demonstrates both of these; the higher relevance of its results over the keyword search showing a greater understanding of themes and the success of its use of full semiotic definitions of these themes over simple keywords suggests that these are perhaps concepts more complicated than a single keyword.

However, such a system is still reliant on very subjective definitions authored by hand, the quality of the results is based on the quality of the definition and it cannot handle themes that have not been previously defined. If a set of super-definitions were to be collected together from a community of contributors using the method defined in the previous chapter we might find a wide variety of definitions that have been refined over multiple authors could be available.

There is further evaluation needed of this system; at this stage the thematic approach has only been shown to be more relevant than simple keyword approaches. It is pos-

sible that a similar effect might be achieved by expanding a keyword with a variety of sophisticated methods of term expansion with a more statistical basis, such as term co-occurrence, without the need for understanding of the semiotic composition of a theme. To this end the following experiment seeks to explore the difference between utilising term expansion on a semiotic basis, such as the thematic model, and more automatic statistical expansion, producing more thematically cohesive and relevant results.

Chapter 5

Photo Montages and Term Expansion

The experiment in the previous chapter demonstrates that a thematic system (in this case the TMB) is capable of making themes become apparent in the context of a photo montage. The experiment also demonstrates that it is better at this then a simple keyword search system (in this case Flickrs search) in the sense that users found montages generated by the TMB more relevant to thematically charged titles.

It can be argued that the reason for its greater performance lies in the models that support the system. By expanding a single word that represents a theme into a selection of subthemes, motifs, and features the system is able to improve the relevance of its thematic representations in the same way that query expansion improves the relevance of search results. In authoring definitions a manual form of semiotic term expansion is occurring and due to the normalised nature of the component coverage metric used by the TMB the structure of this expansion creates a weighting on different terms.

If the thematics system's success lies in this expansion it is important to ascertain if a manual semiotic expansion is the best method of expansion for representing themes. It is possible, for example, that another form of term expansion might also out perform a keyword search as well as not having the manual drawbacks of the semiotic expansion. To answer this question I need to re-perform the original experiment measuring the relevance of images to thematically charged titles but rather then compare it to a keyword system compare it to another term expander.

5.1 Co-Occurrence Montage Generator

As discussed in the background section in [78] Mandala compared and reviewed the performance of a range of term expansion methods for query expansion. While a com-

bination of approaches was used to get the best performance the strongest individual approach was co-occurrence, a method of term expansion that has been demonstrated numerous times to improve the accuracy of search results. As such I identified the co-occurrence term expansion method implemented by Mandala as a suitable candidate as a term expansion method for a system to compare to my own.

In order to keep the comparison fair, the co-occurrence system would operate with the same rules as the TMB. A corpus on the subject of the montage would be compiled and the system would then expand the term representing the desired theme to identify the objects in the corpus with the highest thematic quality. The top N of these images, where N is the desired size of the montage, would then be returned as the montage.

The system rates the semantic similarity of two terms within the corpus based on how frequently they occur, and co-occur. For this system if the terms co-occurred as tags for a particular image in flickr this was recorded as a co-occurrence. Based on these two frequencies the semantic similarity of the two terms may be calculated in a number of different ways: Mutual Information (MI), the Dice Coefficient (DC) [60], and the Tanimoto Coefficient (TC) [101]. The equations for these are displayed below, where for MI P(a) is the probability term a occurs, P(b) is the probability term b occurs, and P(a, b) is the probability the two terms co-occur. For the other equations F(a) is the frequency a occurs, F(b) the frequency b occurs, and F(a, b) the frequency the terms co-occur.

$$MI = log\frac{P(a,b)}{P(a)P(b)}$$

$$DC = \frac{2F(a,b)}{F(a)+F(b)}$$

$$TC = \frac{F(a,b)}{F(a)+F(b)-F(a,b)}$$

All three metrics work on the same principle of how frequent the terms co-occur over how frequently they occur independently, making terms that occur infrequently but always alongside each other the most similar. Mandala performed further tests on the accuracy of co-occurrence when using these different coefficients and found that precision was highest when using Mutual Information, although the margin was very slight. For this reason my implementation also uses Mutual Information as a measure of semantic similarity.

Having done this the system can create a vector for a pseudo document (a model representing a theoretical ideal document with tags proportional to their similarity to the desired term). This is based on the semantic similarity of every term used as a tag in the corpus to the term for the desired theme, where each term is a dimension. The thematic quality of each image is then calculated as the Euclidean distance of a vector

describing the image (where the frequency of each term comprises its distance along that dimension) from the vector describing the pseudo document. In the case where multiple themes are used the half-way point between the pseudo document for each theme is used. Also, when detecting the presence of a term, basic stemming is used so that plurals and other minor variations of the same term are all still detected.

This created a montage generator similar to the TMB in that a desired theme and content could be specified along with montage and corpus size and a montage would be returned that contained images relevant to the desired content that were also themanticly relevant to the desired theme. The difference being one was using the manual semiotic expansions in the form of the thematic definitions and the other performing an automatic expansion based on co-occurrence.

5.2 Evaluation Against Co-Occurrence

By comparing the performance of a co-occurrence based system to that of my semiotic system I seek to find which method of term expansion best represents themes within a piece of media. The semiotic system is based on literary theory of how themes become apparent in narrative, their structure, and the way we interpret different connotations and denotations from features. However it is possible that frequent similar use, measured by co-occurrence, could be similar to this. If co-occurrence could effectively simulate semiotics by performing as well as or out performing a semiotic system at generating thematically relevant montages then it would offer an automatic alternative to using the thematic model to represent themes. Alternatively if the TMB is able to outperform co-occurrence then for tasks where thematic relevance is important it may justify the existing overhead of authoring the thematic definitions.

5.2.1 Methodology

As for this experiment I was seeking to compare the TMB to a co-occurrence based system in the same context as the TMB was compared to keyword search the methodology has been kept as similar as possible. The experiment displays images to participants under a title composed of both a content keyword and theme(s) such as London in Winter (images about London with the theme of winter). Both systems as well as two base cases generate ten image montages for each title and participants view the images both individually and grouped together as a montage and rate their relevance to the titles. The experiment itself is divided into four tests; two tests for titles with a single theme, and two for titles including multiple themes to test to performance of the systems in both situations. For both set of titles the first test displays the images individually at random under the title they were generated for and the users are asked to rate their

relevance to the title from 1 to 5. The second test for each set of titles groups the images together in their montages, once again under the titles, and asks the participant to rate the relevance of the images as a group. The two base cases are used to give the results context and are a low base case of randomly selected images which are taken from the most recent images uploaded to Flickr and a high base case of images selected by a human compiling the best montage they can for the given titles from images in Flickr.

The same titles were used as the previous experiment to maintain continuity and so that a cross comparison of results would be possible if necessary. As with the previous experiment the titles sought to explore how the systems performed under titles including both single and multiple themes as well as titles with themes that complimented the content of the corpus or fabula as well as ones that were contradictory to it. As such four single theme titles were used; two regular theme fabula pairings and two contradictory theme fabula pairings, as well as two multiple theme titles. In the tests requiring single theme titles users were given one regular paired title and one contradictory one alternating to the other two titles on the next participant. As before the titles were London in *Winter*, *Spring* Picnic, Earthquake and *Celebration*, *Family* Factory, *Family* in New York at *Winter*, and *Celebrating* the New House in *Spring*. The same thematic definitions as the previous experiment were also used, as found in section B.1 of appendix B.

As before the systems are subjected to the same rules and restrictions. Each montage may not contain more than one image by the same Flickr user as images uploaded as part of a set have natural flow and may artificially seem more relevant than they are. All montages were generated in the same afternoon to ensure they were using as similar a state of Flickr as possible. And when the images were presented individually they were randomised so as to prevent the identification of which images belonged together in montages. This experiment methodology was granted ethics approval by the universities ethics committee, the approved application can be found in section E.3 of appendix E.

5.2.2 Results

The experiment received a smaller participation than its predecessor with 57 participants, this was however still enough to have statistical significance in my findings. All of the images used in this experiment can be found in section D.2 of appendix D. The thematic system outperformed the co-occurrence based system both in individual images and with montages. Figure 5.1, and tables 5.1, and 5.2 show the frequency data and statistics for single images, and figure 5.2, and tables 5.3, and 5.4 show the same data but for the images grouped as montages. The hypothesis that the TMB selects images rated more relevant for the given titles then the co-occurrence based system is true with a 0.0005 probability of error both for individual images and montages.

Once again while this improvement might seem slight it is important to view it in the

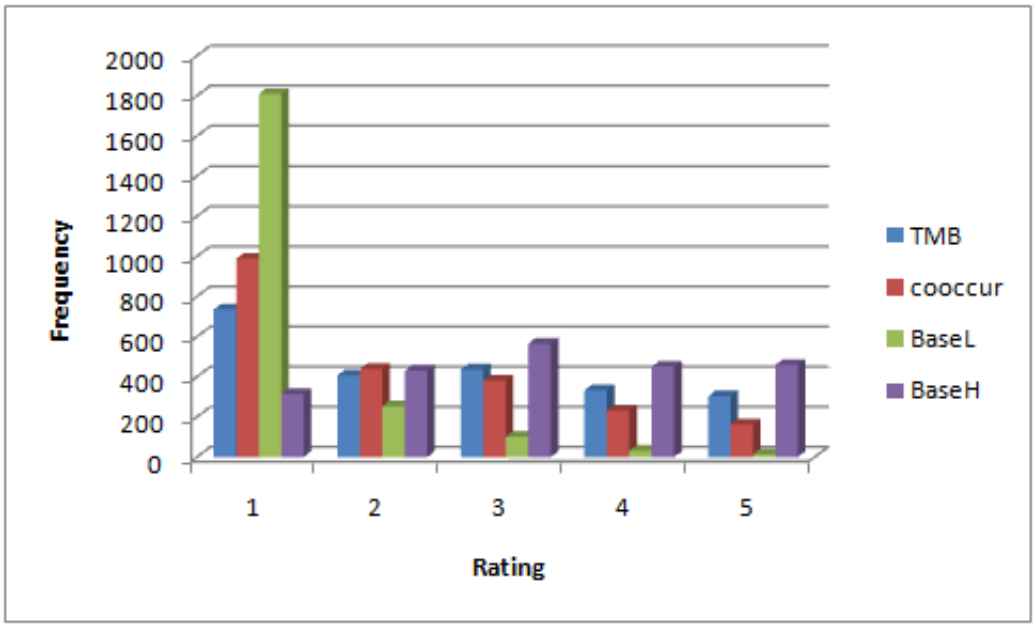


FIGURE 5.1: Single Image Rating Frequency of TMB and Co-Occurrence Experiment

Set	1	2	3	4	5	Total
TMB	736	407	437	332	302	2214
Co-occur	989	438	380	232	162	2201
BaseL	1809	253	100	31	12	2205
BaseH	313	430	565	451	459	2218

TABLE 5.1: Single Images Rating Frequency of TMB and Co-Occurrence Experiment

Set	Mean	SD	Variance
TMB	2.574	1.424	2.029
Co-occur	2.155	1.297	1.683
BaseL	1.269	0.661	0.437
BaseH	3.141	1.330	1.770

t=10.221, df=4413, p=0.0005

TABLE 5.2: Single Images Rating Statistics of TMB and Co-Occurrence Experiment

Set	1	2	3	4	5	Total
TMB	5	39	74	76	30	224
Co-occur	32	75	64	40	14	225
BaseL	180	38	4	0	1	223
BaseH	0	7	32	77	106	222

TABLE 5.3: Grouped Images Rating Frequency of TMB and Co-Occurrence Experiment

context of both base cases. Figures 5.3 and 5.4 show the mean relevance ratings of the four different methods of selecting the images. Standard error was calculated but is too small to display on these graphs. Both graphs show the thematic system outperforming the co-occurrence system, and the margin of improvement which at first might seem small is more impressive considering the margin between entirely random images and

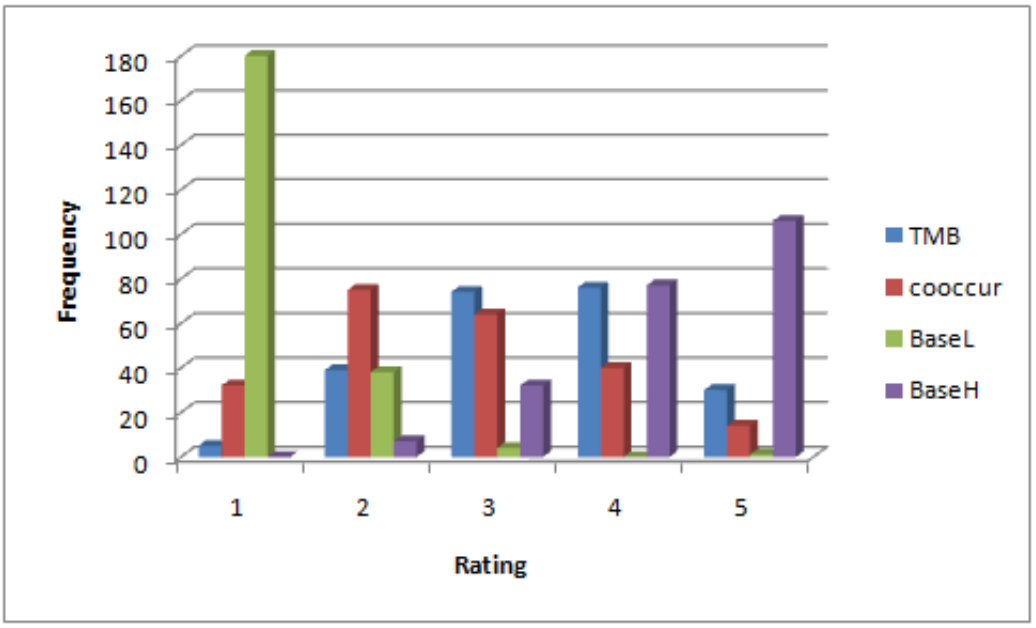


FIGURE 5.2: Grouped Image Rating Frequency of TMB and Co-Occurrence Experiment

Set	Mean	SD	Variance
TMB	3.388	0.996	0.992
Co-occur	2.684	1.111	1.235
BaseL	1.224	0.515	0.265
BaseH	4.270	0.823	0.678

t=7.053, df=447, p=0.0005

TABLE 5.4: Grouped Images Rating Statistics of TMB and Co-Occurrence Experiment

images purposefully selected to make the best montage possible.

As with the previous experiment the images selected by both systems were rated higher when presented as a montage. As shown in table 5.5 The average improvement in relevance rating from rating given as single image to rating given as a montage however is higher for images selected by the TMB then those selected by the co-occurrence based system. It is also to be noted that while it would be expected that a system performing better would receive a proportionally better improvement the improvement experienced is higher then what would be proportionally expected. The hypothesis that the TMB experiences a stronger improvement from individual images to grouped images is true with a less than 0.0005 probability of error.

Set	Mean	SD	Variance
TMB	0.809	1.457	2.123
Co-occur	0.530	1.484	2.203

t=6.297, df=4413, p=0.0005

TABLE 5.5: Grouped Images Improvement Statistics of TMB and Co-Occurrence Experiment

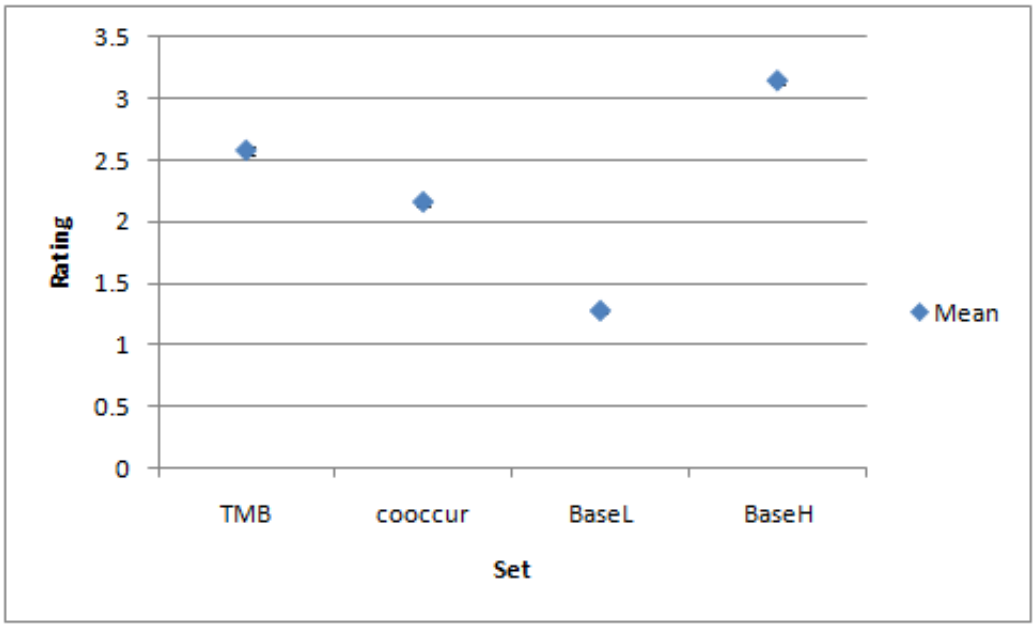


FIGURE 5.3: Grouped Image Rating Frequency of TMB and Co-Occurrence Experiment

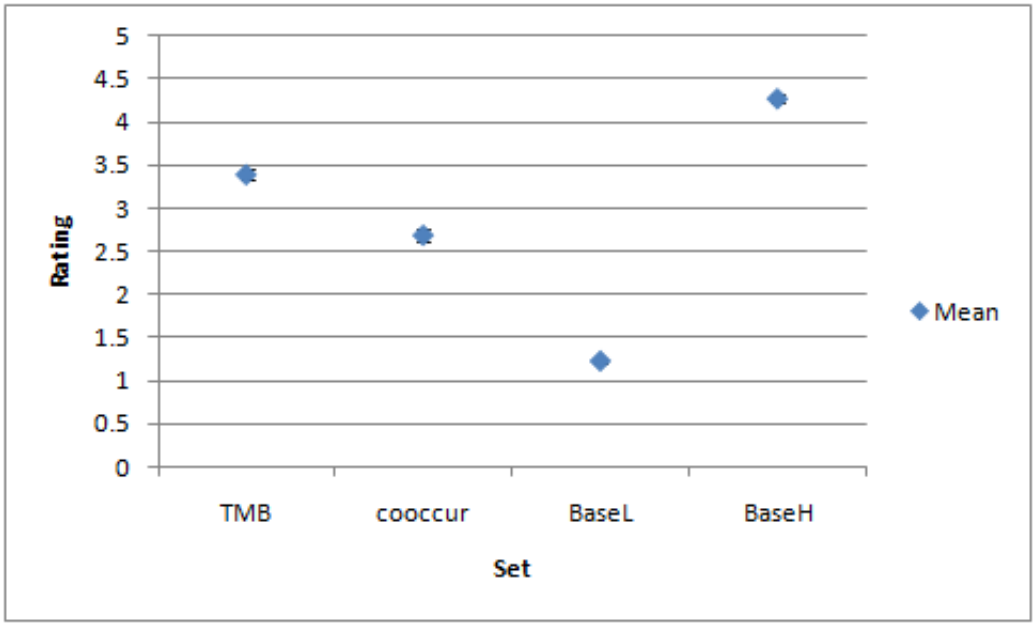


FIGURE 5.4: Grouped Image Rating Frequency of TMB and Co-Occurrence Experiment

As before I also recorded how both systems performed for titles that contained a single theme as well as those with multiple themes. This is shown in tables 5.6 and 5.7 and figures 5.5 and 5.6. Also once again I recorded how each system performed for titles with a contradictory theme fabula pairing as well as those with a regular pairing. This is displayed in tables 5.8 and 5.9 and figures 5.7 and 5.8.

Set	Mean	SD	Variance
TMB Single Theme	2.744	1.515	2.297
Co-occur Single Theme	2.555	1.439	1.069
BaseL Single Theme	1.261	0.671	0.450
BaseH Single Theme	3.163	1.417	2.009
TMB Multiple Theme	2.403	1.305	1.703
Co-occur Multiple Theme	1.749	0.983	0.965
BaseL Multiple Theme	1.277	0.652	0.425
BaseH Multiple Theme	3.118	1.237	1.530

TABLE 5.6: Single Images Single/Multiple Themes in Title Contrast Statistics of TMB and Co-Occurrence Experiment

Set	Mean	SD	Variance
TMB Single Theme	3.428	1.002	1.004
Co-occur Single Theme	3.159	1.130	1.278
BaseL Single Theme	1.153	0.409	0.167
BaseH Single Theme	4.209	0.868	0.754
TMB Multiple Theme	3.348	0.993	0.986
Co-occur Multiple Theme	2.205	0.861	0.741
BaseL Multiple Theme	1.295	0.595	0.354
BaseH Multiple Theme	4.330	0.776	0.601

TABLE 5.7: Grouped Images Single/Multiple Themes in Title Contrast Statistics of TMB and Co-Occurrence Experiment

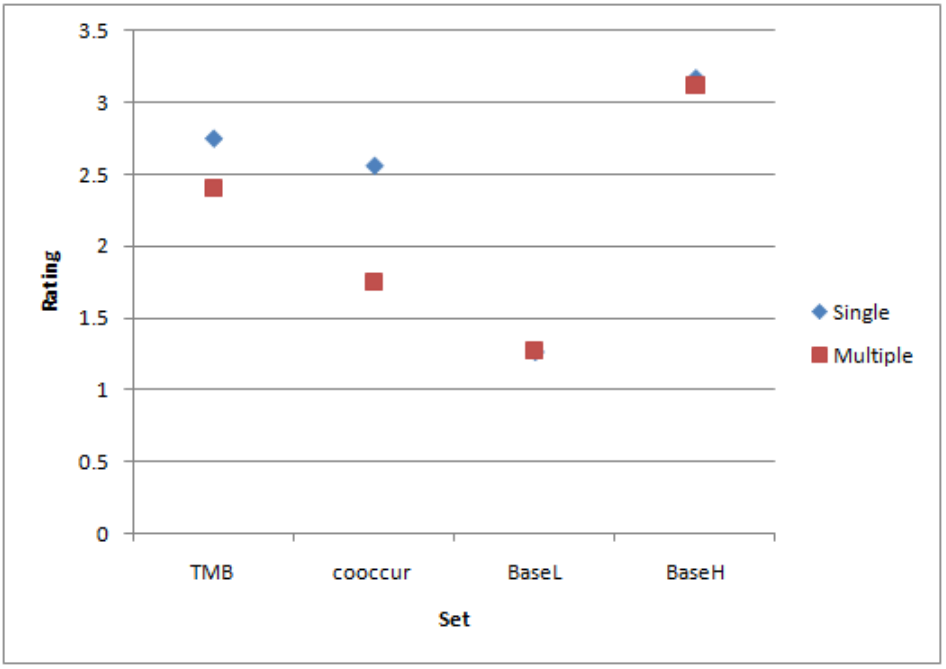


FIGURE 5.5: Single Images Single/Multiple Themes in Title Contrast of TMB and Co-Occurrence Experiment

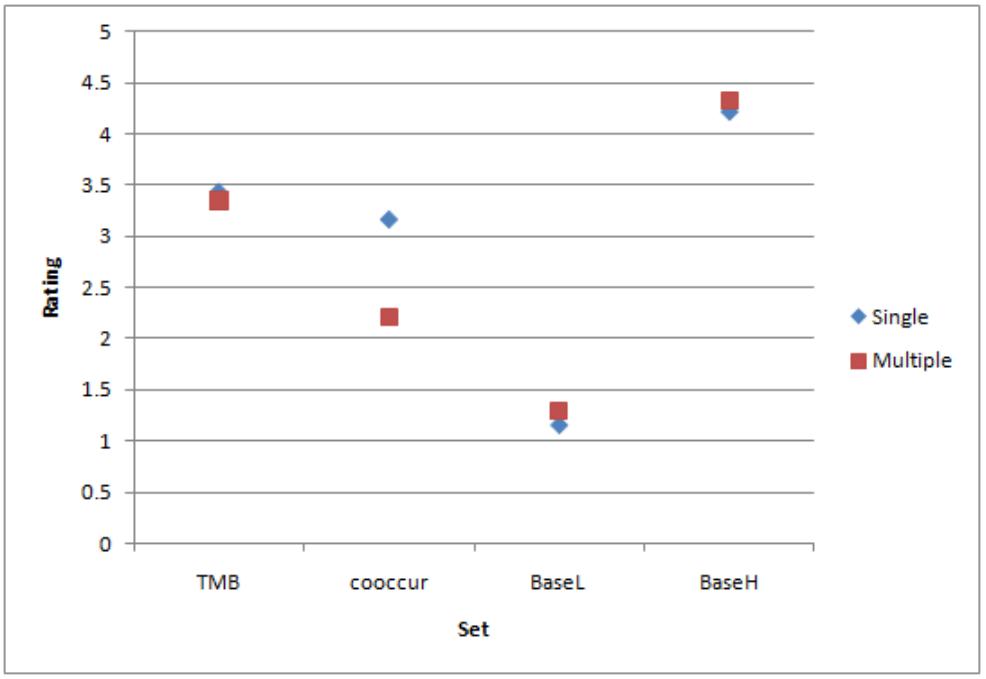


FIGURE 5.6: Grouped Images Single/Multiple Themes in Title Contrast of TMB and Co-Occurrence Experiment

Set	Mean	SD	Variance
TMB Regular Theme Fabula Pairing	2.706	1.445	2.088
Co-occur Regular Theme Fabula Pairing	2.157	1.273	1.620
BaseL Regular Theme Fabula Pairing	1.269	0.664	0.441
BaseH Regular Theme Fabula Pairing	3.369	1.282	1.644
TMB Contradictory Theme Fabula Pairing	2.177	1.281	1.642
Co-occur Contradictory Theme Fabula Pairing	2.149	1.368	1.871
BaseL Contradictory Theme Fabula Pairing	1.269	0.654	0.428
BaseH Contradictory Theme Fabula Pairing	2.459	1.237	1.530

TABLE 5.8: Single Images Contradictory/Regular Theme Fabula Pairing in Title Contrast Statistics of TMB and Co-Occurrence Experiment

Set	Mean	SD	Variance
TMB Regular Theme Fabula Pairing	3.559	0.964	0.930
Co-occur Regular Theme Fabula Pairing	2.601	1.089	1.187
BaseL Regular Theme Fabula Pairing	1.232	0.524	0.275
BaseH Regular Theme Fabula Pairing	4.482	0.709	0.503
TMB Contradictory Theme Fabula Pairing	2.875	0.916	0.839
Co-occur Contradictory Theme Fabula Pairing	2.929	1.147	1.316
BaseL Contradictory Theme Fabula Pairing	1.2	0.487	0.237
BaseH Contradictory Theme Fabula Pairing	3.611	0.811	0.657

TABLE 5.9: Grouped Images Contradictory/Regular Theme Fabula Pairing in Title Contrast Statistics of TMB and Co-Occurrence Experiment

5.2.3 Analysis

The data from the experiment shows the thematic system again producing the most relevant images of the two systems. The semiotic underpinning of the thematic system

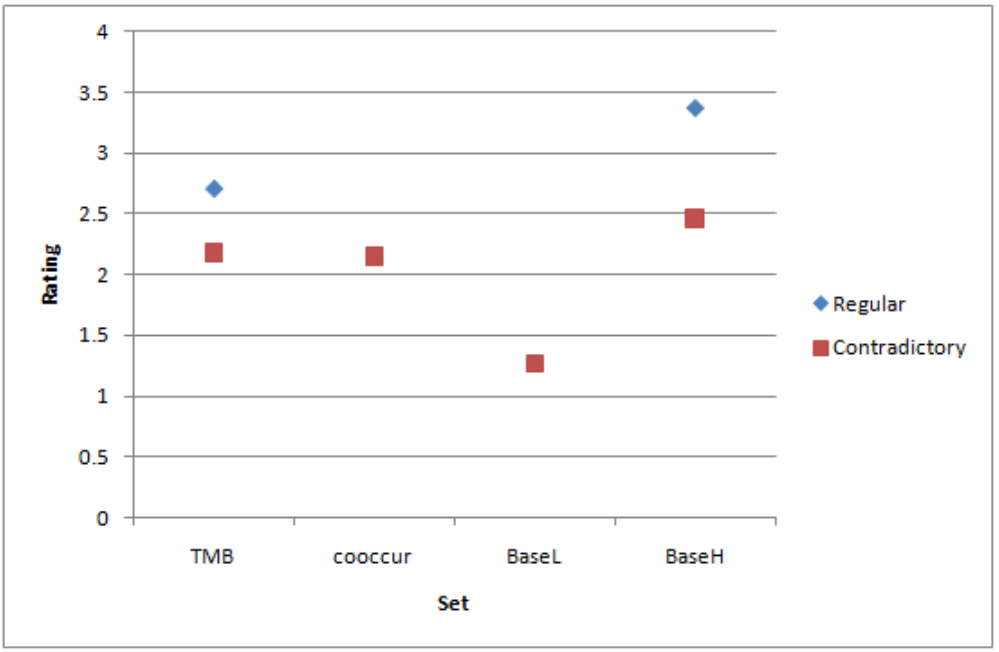


FIGURE 5.7: Single Images Contradictory/Regular Theme Fabula Pairing in Title Contrast of TMB and Co-Occurrence Experiment

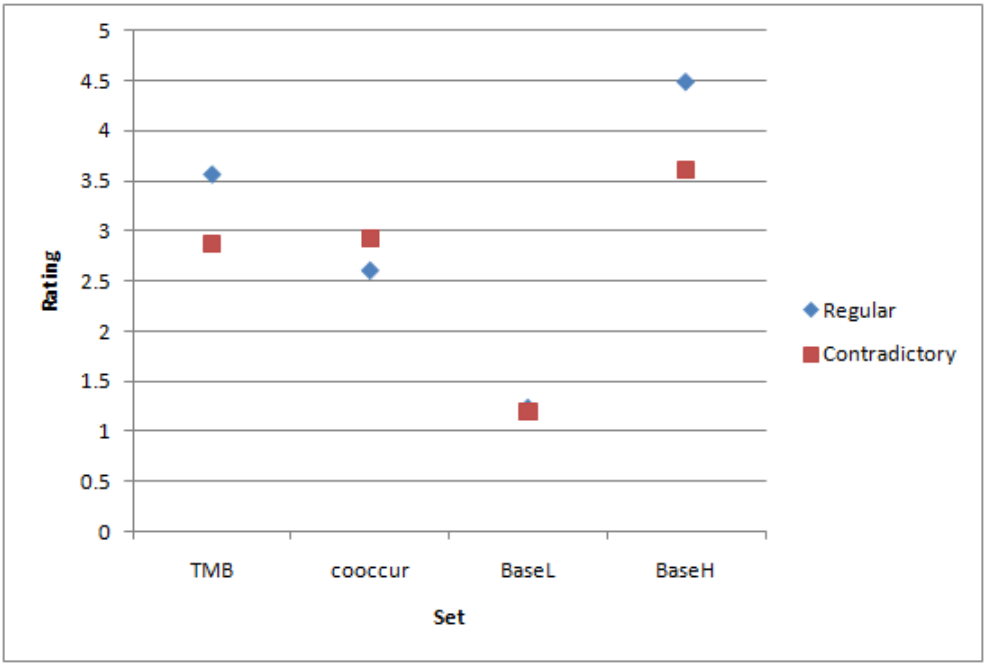


FIGURE 5.8: Grouped Images Contradictory/Regular Theme Fabula Pairing in Title Contrast of TMB and Co-Occurrence Experiment

appears to be a more effective means of expanding thematic content than co-occurrence. The relevance of TMB images was rated higher for both single and grouped images than the co-occurrence images, and the improvement from single presentation experienced by images presented as a montage was also greater for the TMB, all to a degree that can be considered statistically significant. While the improvement experienced may at first seem slight the standard error on the means shown is very small (0.027 - 0.074) and

in the context of the two base cases the improvement is much more important. The improvement from entirely random images to purposefully selected images by hand is 1.872 for single images and 3.046 for group images, the improvement from co-occurrence to TMB is 0.419 and 0.703 for single and grouped respectively. This means the small improvement shown by the TMB is actually 22.3% or 23.1% (depending on single or grouped images) of the improvement from worst to best case. In this light the improvement demonstrated cannot be considered insignificant.

Once again the TMB also showed it was more capable of selecting images for titles containing multiple themes, as before this can be attributed to the way thematic score is calculated emphasising images relevant to both themes and looking for common shared motifs. As before the TMBs weakest performance was when it was required to produce montages for titles with a contradictory theme fabula pairing in the title, due to the fact that the features representing the specific needed motifs are rarely found within the corpus. However in this case the co-occurrence system also struggled and instead of scoring higher than the TMB (as the keyword system did in the previous experiment) it performed comparably similar or worse.

Across the experiment as a whole the co-occurrence based system has performed significantly worse than expected, performing only slightly better for grouped images then the keyword search and worse for single images, this is exaggerated by the fact that the TMB performed stronger than before as well. There are a number of factors that may have contributed to this. Flickr, as a user contributed folksonomy, is constantly changing and evolving and as such its state at the time of the experiment might have been less favourable to the co-occurrence system. However, in order to remove that as a possibility both experiments would need to be carried out on multiple states of Flickr and this is not feasible within the scope of this research. Also it is possible that the co-occurrence system became affected by query drift as mentioned earlier in the background and discussed in [123]. This is to some extent born out in the image sets generated by co-occurrence where we can see it has drifted in one example from winter to snow to snowdrop (the flower). It has also been noted in work such as that by Xu [118] that the best results from co-occurrence come when it is trained using a local corpus of ascertained relevance to the query being expanded. While I was training using a local corpus it was not specifically relevant to the element I was expanding, for that to be the case the corpus would (as an example) have to be populated with a Flickr search for 'London in Winter' rather than just 'London'. If this is the case it is possible co-occurrence is less effective for expansion of terms for which it is more difficult to acquire a training corpus of ascertained relevance such as a theme.

Another possibility important to consider is that the TMB, while seemingly performing better, may be particularly well suited to a particular title and is having its average dragged higher by this individual case. In order to analyse this a little further table 5.10 displays the mean rating for each title from both the TMB and the co-occurrence

systems for single images whereas table 5.11 does the same for montaged images. Both tables also show the improvement in relevance made by the TMB (negative numbers representing instances where co-occurrence performed better).

- Title 1: London in *Winter*
- Title 2: Earthquake and *Celebration*
- Title 3: *Family* Factory
- Title 4: *Spring* Picnic
- Title 5: *Family* in New York at *Winter*
- Title 6: *Celebration* of New House in *Spring*

Set	Title 1	Title 2	Title 3	Title 4	Title 5	Title 6
TMB	3.892	2.253	2.097	2.684	2.371	2.435
Co-occur	2.690	2.563	1.710	3.254	1.601	1.900
TMB Improvement	1.202	-0.310	0.387	-0.569	0.770	0.535

Table 5.10: Mean Rating by Title for Single Images of TMB and Co-Occurrence Experiment

Set	Title 1	Title 2	Title 3	Title 4	Title 5	Title 6
TMB	4.172	3.000	2.741	3.778	3.345	3.351
Co-occur	2.793	3.600	2.185	4.037	1.857	2.553
TMB Improvement	1.379	-0.600	0.555	-0.259	1.488	0.797

Table 5.11: Mean Rating by Title for Grouped Images of TMB and Co-Occurrence Experiment

As can be seen by this data the TMB has scored significantly higher for titles 1 and 5 which were 'London in Winter' and 'Family in New York at Winter'. It could be possible on this initial observation to say that the TMB was simply drastically more effective at retrieving results for titles including 'Winter' and that this has dragged its overall average up. However if we remove the mean ratings for both titles including winter entirely we find the TMB still has a higher mean than co-occurrence for both single and montaged images. Showing 2.380 for the TMB and 2.267 for co-occurrence for single images and 3.243 for the TMB and 2.992 for co-occurrence for grouped images. It is also still statistically significant, even excluding the winter titles, to say the TMB performed better than the co-occurrence system with a t of 2.247 (p=0.01, df=2741) for single images and a t of 1.952 (p=0.05, df=278) for grouped images.

The TMBs success at 'winter' as a theme to an extent can be explained by the nature of the experiment in that I am using photographs. 'Winter' is quite a visual theme, its motifs such as 'snow', 'ice', and 'christmas decoration' are both visually very easy to identify and also something people often take photographs of and as such the system will likely have a wealth of relevant photos to choose from in the corpus. Other themes with less obvious visual motifs associated with them might be more difficult to represent

as an image montage. Despite the lower scores of other themes though the thematic system has still shown to be more effective consistently then other methods, generating montages more relevant to titles containing desired themes.

An interesting observation from both experiments is the thematic systems affinity for the presentation of images as a montage. While all systems experience an improvement, in this case the TMBs improvement is disproportionately higher. As discussed previously a theme in the narrative sense of the word becomes apparent over the course of a narrative through a succession of motifs, as such connoting a theme in a single image (or natom) is difficult where as a spread of motifs over a series of natoms, in this case the photo montage, allows the theme to become apparent. This would explain why all systems experience a higher rating for montages. It is possible that the thematic system, having taken the narratology based approach of looking for a breadth of relevant motifs is constructing a fuller representation of the theme in its montages. The other methods are either looking for a single keyword or only terms that share a usage instance with the expansion term itself.

5.3 Conclusions

In the previous two experiments I have compared the ability of a thematic system to generate themed photo montages against two other systems: keyword search and a system based on co-occurrence term expansion. Both experiments shared a similar methodology where for a series of titles, each containing a desired content and desired theme(s), the systems generated a ten photo montage. A series of participants then rated the images both individually and in their montages on their relevance to the titles they were generated for, all of this done in context of two base cases. The experiments have shown the following:

- It is possible to use definitions created in terms of a thematic model to generate simple photo montages relevant to a desired theme.
- A system using thematic definitions creates montages rated more relevant than those offered by both basic keyword search and co-occurrence term expansion.
- The thematic system is still effective in situations demanding multiple themes but sometimes less effective if the desired content and theme clash.
- While all systems find montages more effective than individual images at representing a themed title the improvement experienced by the thematic system is greater.

While it has been shown that co-occurrence is effective at increasing the relevance of search results in a general sense for the specific task of queries in a thematic context

the TMB appears to be more effective. Themes are complex concepts comprised of a variety of motifs that sometimes will become apparent over the length of a narrative and not share a particular instance with each other, potentially reducing co-occurrence and other systems effectiveness. The refinement and removal of associated elements at the end of the thematic systems definition process also helps reduce query drift and the expansion of semantically irrelevant terms to which co-occurrence is still vulnerable.

While coming at the cost of having to manually make thematic definitions, the TMB is consistently outperforming alternative systems including an alternative term expansion system. It is able to deal with a range of content and themes and is not noticeably affected by situations calling for multiple themes. While it does struggle with situations where the content of the corpus it is using and the desired theme clash, the similarity of its performance in this area during this experiment to that of co-occurrence suggests that this is a weakness shared by potentially all methods of term expansion. If the terms that would share semantic similarity are not present there is simply nothing to expand it to. While all systems experienced a rise in relevance rating for images presented together as a montage the TMBs disproportionately higher improvement suggests it will be more effective in more narrative contexts where it has the room to express a breadth of motifs over multiple natoms.

If we accept as a result of these two experiments that the best method for representing themes and selecting thematic content is one like the TMB using thematic definitions created through semiotic term expansion then the next step is to assess the effect of the usage of such a system in richer narratives. As discussed previously part of the reason themes could be considered important in narrative is that thematic cohesion helps tie a story together giving the individual parts natural cohesion and flow together.

In chapter 3 I identified the presentation level of narrative generation as that most suitable for integration with a thematic system and considering thus far my work has largely dealt with images the logical next step is a system that represents narratives with images. As such the next experiment seeks to investigate the effect on narrative cohesion of a thematic illustrator that takes short stories and redisplays them with selected images. This will allow an insight not only on the actual value of thematic cohesion but to the effect of using my system to attempt to improve it.

Chapter 6

Thematic Cohesion

The previous two chapters have described two experiments that show a thematic system improving the relevance of a selection of images to thematic titles over two other methods. The definitions created through semiotic term expansion allow for the system to identify the thematic quality of tagged items with regards to desired theme(s) by looking for the presence of features denoting relevant motifs. As well as for thematic searching this might be used as a basis for thematically augmenting narrative systems such as narrative generators. By being able to understand the relevance of a tagged piece of media to a theme a decision on whether to include it or not in a narrative with specified themes may be undertaken. Through this we may improve the thematic cohesion of narratives generated or presented by various narrative systems by emphasising or embedding a thematic subtext.

It is important to ascertain what the effect of embedding this is on the user experience of consuming a narrative. Cohesion aids the flow of a narrative by binding its different elements together but underlying themes are just one part of that. In this chapter I present an experiment that seeks to explore whether the emphasis of desired themes improves an audience's perspective of thematic cohesion and whether this increases the cohesion of the narrative as a whole. In order to do this I will need to identify exactly how narrative cohesion manifests and how it can be measured.

In chapter 3 I assessed the presentation level of narrative generation as the most viable for thematic system integration due to its relevance to a range of different approaches to narrative generation and a reduced chance in the integration potentially damaging the plot of the resulting narrative. It was identified that the emphasis of desired themes in the narratives presentation was a viable way of approaching this and I look to do that in this experiment through illustrations. This leads to another key focus of the experiment where I seek to ascertain whether a thematic system is capable of emphasising a given theme through illustrations.

6.1 Variables of Cohesion

The objective of this experiment is to assess the impact of thematic illustrations on perceived narrative cohesion. In order to do this some tangible ways of measuring narrative cohesion must be established. The idea of cohesion as a way of tying narrative or text together to aide how it flows and aide the reader in understanding what is being told has been explored in a range of research fields where I can draw on some important variables related to narrative cohesion. Much work in the area considers narrative cohesion to be the linguistic cohesion of the presentation of the narrative such as the work by Hudson [64] and, in part, the Coh-Metrix project [49]. I, however, think that narrative cohesion is more than this and that subtextual elements such as themes and genre as well as the role of the narrator have a large part to play in building a coherent narrative as we will explore in this section. In their work on MetaLinks Murray et al. [91] distinguished between these two different types of cohesion as 'Narrative Flow' and 'Conceptual Flow' however while the presentation of text is fundamental to the discourse a 'narrative' in the way I define it is significantly more complicated and includes both of these aspects.

Narrative cohesion is explored by Hudson [64] from the perspective of investigating children's understanding of narratives based on what a child can explain or recall from strong cohesive stories that flow together as well as incoherent difficult stories. The coherence of stories constructed by children themselves is also analysed. Hudson's 'cohesive devices' are largely centred around the logical sense of the narratives and are partly based upon the coherence categories of John-Steiner and Panofsky in their own work on children's narratives [66]. Hudson measures the cohesive logical sense a narrative makes through the presence of conjunction (connective terms in four categories ranging from the simple such as 'and' to the temporal such as 'then'), prepositional phrases and relative clauses (relative explanations triggered by words such as 'who' and 'that'), and anaphoric reference (referring to earlier imparted information). It is clear from Hudson's work that how much logical sense a narrative makes is an important aspect of its cohesion, whether it be temporal, referential, or explanatory in nature. However as well as the specifics of the language used, it is possible that narrative coherence could also be extended to a narratives subtext.

As mentioned earlier Tomashevsky discussed in his essay the importance of themes to narrative [113]. He explains that narratives are given meaning by themes and that themes unite the separate elements of a narrative giving it cohesion. Work by Ferret [44] also explores the link between themes and cohesion, demonstrating how lexical coherence based on collocation can be used to show thematic coherence in segments of narrative. The collocation network is based upon earlier machine readable dictionary work by Kozima based on lexical similarity [67] allows Ferret to measure the cohesion between terms in a section of text in order to infer the discussion of a similar concept and

infer thematic cohesion. The network Ferret uses is similar to the thematic definitions used by the TMB in that it is a collection of terms and relationships that can be used for understanding thematic content. However, it is based in collocation (co-occurrence) not semiotics which I have found to be less effective in the cases I have tested. The notion of theme is core to cohesion from the structuralist point of view in that it binds the narrative together with meaning. It is also possible it could be closely connected with the coherent use of language and the logical sense of the narrative as demonstrated by Ferret.

Genre is a common classification of narrative based upon a set of reoccurring features drawn from content and plot that position a narrative culturally within the context of other narratives. Tomashevsky pointed out that the motifs present that lead to the presence of a theme are intrinsically connected to the genre of a narrative suggesting that the genre of the narrative was what limited the motifs available [113]. The Coh-Metrix project [49] worked towards creating a system for analysing the coherence of texts through several metrics. The metrics used were a variety of text based lexical analyses such as latent semantic analysis, term frequency and density, and concept clarity. The measuring of these metrics however was intrinsically based upon the pre identified genre of the narrative which they identified in their initial Coh-Metrix report [83] as important to coherence. The work done by the Coh-Metrix project to date supports the importance both of logical language used and identifiable genre to measuring the coherence of narrative. In his work identifying key features of narrative Bruner [23] also highlights the importance of genre to cohesion. Under his discussion on 'Genericness' he explains how genre is a way of 'comprehending narrative' by using enabling language to act as a guide. By conforming to convention the narrative guides the audience to subconsciously fill in gaps in the presentation and make sense of the content.

There are other aspects of narrative I have not yet considered that could be an important factor in narrative cohesion. In work by Booth [18] there is a description of the importance of the concept of narrator in narrative. Booth explains how the authorial voice may be used either implicitly or explicitly to deliver the plot and build a relationship of short or long distance with the reader. As the narrator is core to the telling of the story, coherence in how the narrator is presented is also important to the cohesion of the story itself. McAdams explains from the perspective of modern psychology that people become narrators in order to make sense of a series of events or stories, thus it is the presence of a narrator that leads to coherence in a story [82]. The consistency with which a narrator is presented in a story affects how the story is perceived as being communicated and may be used to build a relationship with the narratives audience, as such the strength of a narrator's presence in a narrative could be considered as a variable that affects the narratives coherence.

We have already discussed how the logical use of language may affect the coherence of a narrative however there are other linguistic choices made in the telling of a story that

might affect its coherence. Earlier I discussed how structuralists such as Barthes [13] and Bal [11] consider narrative to be comprised of layers, often of story and discourse or a close to equivalent model where story stands for content and discourse for how the story is told. Features of discourse have already been identified here; themes, genre, narrator, but these cannot be said to completely account for the language choices made in a narrative's discourse through the presentation of individual elements. The way different narrator's approach and use a style of language can have an effect on its coherence. Style can be said to be a composite of attitude, tone, and mood of a narrative, representing decisions made on the presentation of elements at the discourse level. But also style represents the conventions authors have set for themselves either in previous narratives or earlier in the narrative in question. The stylistic cohesion of a narrative could be said to be in part the extent to which an author when making decisions about language used convenes to their own conventions.

From this I identify five key variables of narrative cohesion; logical sense, thematic cohesion, genre cohesion, narrator cohesion, and style cohesion. These can individually be defined:

- **Logical Sense:** the connective language used to explain the content of the narrative.
- **Themes:** the concepts communicated implicitly throughout the narrative.
- **Genre:** the conformance to conventions that culturally contextualise the narrative.
- **Narrator:** the presence of a consistent perspective communicating the narrative.
- **Style:** the way narrative elements are presented within the discourse.

Measured appropriately, and considered together, these may be used as a basis to understand the level of cohesion within a narrative.

6.2 Example Method of Measurement

In order for such a series of variables to be effective I need to consider their measurement. In order to form a formal method of measurement we must first consider what could be considered evidence of the value of each variable by identifying positive and negative narrative features for each. In this section I cover an initial, subjective, exploration of this detailing some features that might later be automatically detected, and a manual example of how detecting these features might provide an understanding of the cohesion of two generated narratives.

Table 6.1 contains candidate lists I have initially expanded from my definitions of positive and negative features for each variable; each feature is based upon either existing work exploring the area or my own definitions of these variables. The presence of a positive feature within a narrative can be considered evidence to suggest strength for the relevant variable whereas the presence of a negative variable could be considered weakness. It is to be noted that some parts of this could potentially be automated, in particular Coh-Metrix [49] might provide a way for several measures of logical sense and perhaps genre cohesion. However for the purposes of this section I present the root features that could be identified by hand connected to each cohesion variable

Variable	**Positive Features**	**Negative Features**
Logical Sense	- Correct use of conjunction - Correct use of preposition and anaphoric reference - Story is chonologically presented	- Content is Obfuscated - Content is contradictory/ not causal
Theme	- Core themes are identifiable - Core themes are present throughout	- Sub-themes conflicts with core themes
Genre	- Story fits conventions of an identifiable genre - Story follows genre conventions throughout	- Story fits multiple genres
Narrator	- Story has dramatised/explicit/ identifiable implicit storyteller - Storyteller is present throughout	- Storyteller changes character/personality
Style	- Language and Style used fits authors own previous conventions - Tone and mood of each scene is relevant to its content	- Style is not identifiable

Table 6.1: Cohesion metric variables and features

The features presented in table 6.1 should be considered examples as a starting point, and by no means exhaustive definitions. Each of them represents evidence that as a starting point eludes to positive or negative reflection within a given variable. It can also be said that the inverse of a feature has inverse effects; for example the absence of a positive feature is in itself a negative feature, and vica versa.

6.2.1 Example Analysis

In this section I demonstrate how these variables might be applied to narratives created from narrative generation in order to capture an impression of the narrative cohesion of these stories. For the purpose of this exercise I use an early but prominent character centric narrative generator; TaleSpin [85], and a more modern author centric narrative generator; ArtEquAKT [117]. TaleSpin generates stories about predefined characters with predefined goals in predefined settings. The system resolves a simple assessment of each characters actions in turn and reports on each as a sentence in a simplistic

manner until their goals are resolved, it is in a sense simulating the characters actions and exposing them to form an emergent narrative. ArtEquAKT on the other hand automatically generates biographies on artists by linking together relevant narrative segments taken from web resources into a biography structure. For the purpose of this example I used a TaleSpin story about 'George Ant' as shown in figure 6.1 and an ArtEquAKT biography of Rembrandt as shown in figure 6.2.

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 THE MEADOW THROUGH A VALLEY TO THE RIVER. WILMA DRANK THE WATER. WILMA WASN'T THIRSTY ANYMORE.
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. GEORGE FELL INTO THE WATER. GEORGE WANTED TO GET NEAR THE VALLEY. GEORGE COULDN'T GET NEAR THE VALLEY. GEORGE WANTED TO GET NEAR THE MEADOW. GEORGE COULDN'T 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.

FIGURE 6.1: A TaleSpin narrative.

Rembrandt HARMENSZOON

Summary Biography

Rembrandt HARMENSZOON was born on 15 July 1606 in leiden,netherlands.

Rembrandt HARMENSZOON was married in 1634 to null.

Rembrandt Harmenszoon van Rijn Rembrandt (1606 - 1669) Rembrandt, the Dutch baroque artist, who ranks as one of the greatest painters in the history of Western art was born in Leiden on July 15, 1606. His father was a miller who wanted the boy to follow a learned profession. Rembrandt began his studies at the Latin School, and at the age of 14 he was enrolled at the University of Leiden. The program did not interest him, and he soon left to study art. First with a local master and then, in Amsterdam, with Pieter Lastman, known for his historical paintings. After six months, having mastered everything he had been taught, Rembrandt returned to Leiden, where he was soon so highly regarded that although barely 22 years old, he took his first pupils.

During the years of their mutual life Rembrandt created such masterpieces as The Abduction of Ganymede. (1635), The Angel Stopping Abraham from Sacrificing Isaac to God. (1635), The Feast of Belshazzar. (c. 1635), The Blinding of Samson. (1636), Danae. (1636), The Prodigal Son in the Tavern (Rembrandt and Saskia). (c. 1635), The Night Watch (1642) and others. The Night Watch , maybe is the most famous Rembrandt work, and his the largest one (12x15ft. 3.5x4.5m), was commissioned by a company of the Civil Guard of Amsterdam for its assembly hall. The painting is a ?capitulation of the ideals of Rembrandt first ten Amsterdam years, and is the last painting in which he strives for brilliant external effects. From now on he set himself the aim of recreating in visual terms the intangible essence of man, his inner life?In his last two decades Rembrandt simplified his compositions, preferring more classical and stable structure.

In 1625 the 19-year-old Rembrandt returned to Leiden and opened his own studio, which he shared with his friend of the same age, Jan Lievens. Rembrandt executed historical paintings, initially following Lastman models: Tobit and Anna (1626), The Ass of Balaam Talking before the Angel. (1626). His physiognomic studies, resulted in numerous self-portraits: Self-Portrait . (c. 1629), Self-Portrait with Wide-Open Eyes . (1630). During his lifetime Rembrandt executed more than 100 self-portraits. He also produced many engravings and etchings.

FIGURE 6.2: An ArtEquAKT narrative.

'George Ant' shows some of the limitations often seen in early text generation systems in that outside of the structure, sentences are generated regardless of the content in each

other. It cannot be said to make good use of conjunction with sentences kept entirely separate when they might be more coherently joined. However it is chronological, there is no contradiction, and the content is not only not obfuscated but directly explicit. Thematically speaking the story exudes no core themes, except for possibly 'survival' or 'debt' (which are very weakly exposed). The story does conform to the genre of a 'fable' in that it anthropomorphises animals/insects and has a moral lesson, and this is present through out. There is no identifiable storyteller process and indeed the disjointed way the text is presented gives it a very inhuman feel. Finally stylistically the story is consistent and its disjointed style is coherent with other works by TaleSpin. Based on the presence of these features 'Geroge Ant' could be considered to have a high genre and style cohesion, a mixed logical sense cohesion, and very low thematic and narrator cohesion.

The ArtEquAKT biography of Rembrandt has some similar traits but also demonstrates some differences. It has good use of conjunction but is not chronological (presenting a paragraph on his later works before his earlier works) and missuses anaphoric reference by referring to a 'capitulation of the ideals [in his] first ten years' without expressing what these are. The content couldn't be considered obfuscated but a few text generation errors make parts slightly unclear. The piece is absent of themes, though in maintaining the factual voice of a biography this may be deliberate. The narrative does strongly conform to the conventions of a biography with initial details of key dates, a discussion of his early life, and then the details of his work, for which the system has been purposely designed. There is no key identifiable storyteller presence tying the work together, though again this could be said to be a deliberate feature of the genre where the presence of a personal perspective might compromise the biographies impartial nature. Stylistic choices are coherent throughout and consistent with what is expected from ArtEquAKT. Based on my summary of these features we might find this story similar to 'George Ant' (High genre and style, mixed logical sense, and low thematic and narrator) however there are some differences to consider. First of these is that the genre of the ArtEquAKT narrative (biography) predisposes it to be weakly themed and without the presence of a narrator to maintain its factual nature, as opposed to the TaleSpin story whose genre (fable) might have benefited by the active inclusion of themes and a storyteller. Secondly though both had mixed features for logical cohesion each demonstrated different features, TaleSpin's more simplistic style avoiding the need for more complicated back reference or structure (on which ArtEquAKT's story failed) but at the same it lacked the flowing sentences and conjunction ArtEquAKT demonstrated (thanks largely to its use of preconstructed prose).

What this process demonstrates to us is that an evaluation of different generated narratives cohesion is possible using my five variable approach. The listed features allow for the spotting of individual parts of a narrative that might cause the story as a whole to be coherent or incoherent. However the incomplete, example, nature of such a list

demonstrates there is further work and discussion to be done on the identification of cohesion features. As part of the analysis the fact that high cohesion is not always positive was demonstrated; a lack of lower presence of particular cohesion variables can be used for specific narrative effect or as part of the conventions of a genre. The cohesion evaluation of ArtEquAKT could be considered near perfect for that narratives objectives as its factual nature and style could be weakened by the presence of heavy themes or a narrator perspective. Where as the evaluation of TaleSpins story, despite being similar, is less ideal as the objectives of the narrative might of been enhanced by the presence of both themes and a storyteller. How these different variables of cohesion are connected to different genres, and purposes of writing is a connection yet to be clearly established, and key to detailed evaluation of the role of cohesion within a narrative.

From this process we can draw some observations on the cohesion of generated narratives, though it is to be noted these observations are based only on what we have seen from two systems and do not represent a complete survey. What has been shown is that automatically generated narratives can demonstrate similar traits, as both were shown to have high genre and style cohesion and low thematic and narrator cohesion. Stylistically speaking a generated narrative is likely to show high cohesion as each passage of text is generated with a similar, if not the same, method. However it is possible that systems which rely on resources written else where, such as ArtEquAKTs use of online material, might experience a collision of styles in some cases where material from two very different writers is used. It is also possible that the conventions of a genre are easier to capture then that of the presence of a narrator (as suggested in my findings) but this is too simplistic an analysis to make such a claim and it is important to consider that one of the systems used is purposely built to follow the conventions of a genre and that we might have found something similar for the narrator variable had I used a system that has purposefully considered narrator presentation such as the virtual storyteller [108].

6.3 Cohesion Experiment

Having decided upon some key variables for measuring narrative cohesion an experiment was developed to answer my two initial questions:

- Do thematic illustrations selected by a thematic system improve the perceived thematic cohesion of a narrative?
- What effect does this have on the perceived cohesion of the narrative as a whole?

The first question will demonstrate whether thematic integration at the presentation level is sufficient enough to have a genuine thematic effect on the story while the second will show if this actually translates into a positive effect on the cohesion of the story in

other variables (logical sense, genre, narrator, style) as well. If an improved thematic cohesion makes a story flow better and be more easy to read then we would expect to see an improvement in thematic cohesion translate into a more general improvement of cohesion as whole.

6.3.1 Methodology

For this experiment participants filled in a web questionnaire on perceived narrative cohesion for short stories with illustrations. There were three short stories selected and three different methods of generating illustrations for the stories with each user seeing the three stories with illustrations generated from different methods. The illustration method - story pairings were stratified across users to get a spread of data for each method on each story.

The stories used in the experiment would be divided into logical sections with each section given an illustration. To facilitate this the stories would be stored as xml allowing them to be marked up where the different sections began and ended. The illustrations for the stories are based either on the content of the section for that illustration or the content and the stories theme, depending on illustration method as explained later. To facilitate this the xml model for each story would store a content keyword for each section as well as a theme for the whole story.

The stories used were selected from Steve Ersinghaus'[1] contributions to the 2009 100 days project[2] where he wrote 100 short stories. This was an ideal resource for the experiment with a large collection of stories with suitably complex themes, strong imagery that lent themselves to illustration, and an author that was happy to help and engage with the experiment. Fifteen of the stories were reviewed for their suitability for the experiment. Stories that were picked were ones which logically fell into 3-5 sections and were of an appropriate length (took less than 10 minutes to read). Also, to ensure the spectrum of naturally occurring coherence in the plot was covered, a story that was distinctly abstract (and arguably authored with deliberately low cohesion) was selected, as well as a story that was more deliberately strongly coherent, and a third that fell somewhere between. The three stories selected were:

- **Story 1 - The Point:** A quite abstract story about two people meeting.
- **Story 2 - The Night:** A dark and slightly horror based story about a boy and unseen terrors with strong visual imagery.
- **Story 3 - Computer Leon:** A very strongly flowing, dialogue based story about competition between computing professionals.

[1] http://www.steveersinghaus.com/
[2] http://onehundreddays.net/

All three met the criteria in that they could be easily divided into sections, were of an appropriate length, and provided a spread of different inherent coherence. The stories themselves can be found in appendix C.

The illustrations for the stories would be generated in three different ways:

- **Method 1 - Content and Theme:** Illustrations were generated based on the content keyword for each section and the theme selected for the story. This would be done using the TMB with a corpus based on the content keyword from Flickr and the theme designated for the story.
- **Method 2 - Content only:** Illustrations were generated based on a chosen keyword describing the content in each section. This would be done using a Flickr search for the keyword.
- **Method 3 - Human Selected:** A base case of illustrations selected from Flickr by a literature expert after they have been given time to analyse the stories and consider their themes and asked to select the illustrations they consider best for each section of the story.

A comparison between methods 1 and 2 would show whether thematic cohesion had increased due to the themed images and also whether this had resulted in a change in other cohesion variables. Method 3 on the other hand will give my results context with an intended best case scenario. The expert for method 3 was an English Masters graduate from Cambridge University with a history of involvement in both literary criticism and computer science research communities, and was completely independent of the research team.

A fair and impartial way of generating the meta data necessary for the experiment was developed. Before selecting images for method 3 our expert who had performed an analysis of the stories was asked to identify a keyword to describe the literal content of each defined section of the stories and also to list the themes they felt were present within each story. They were also asked to identify from their lists of themes for each story which they felt was the strongest theme. The strongest themes went into the story models as the listed theme for each story and the keywords for content identified were entered for the content keyword for each relevant section. Having completed this the newly identified strongest themes needed to be modelled into definitions for use with the TMB, while the other identified themes were not modelled they were still a part of the experiment as explained later. To keep the definitions of the identified themes impartial three volunteers independent of the experiment were asked to follow the thematic definition guide explained in chapter two to define the themes. While an expert in the model was present to collaboratively help in forming these definitions to ensure the models created were valid, creative control of the definitions was left solely

to the volunteers and all the themes and motifs comprising the model were identified by them. This would ensure that the terms used in the definitions as well as the story meta data were impartial and not biased towards any particular system and would also test the TMBs ability to deal with more complicated and subtle themes. The structural definitions created for these themes can be found diagramatically in section B.2 of appendix B. The stories and their identified themes are displayed in table 6.2.

Story	Strongest Theme	Other Identified Themes
The Point	Relations	Pun, Description Simile, Innuendo, Geometry, Microcosm, Macrocosm
The Night	Transgression	Discovery, Authority, Fear, Experience, The Unknown
Computer Leon	Competition	Business, Reputation, Service, Technology, Ignorance, Marketing, Cheating, Sabotage

TABLE 6.2: Identified Themes for the three stories in the cohesion experiment

Having completed my models of the stories, illustrations were generated for them using the various methods and added to the models. In the case of method 3 this meant simply appending the locations of the illustrations selected by the experts. For methods 1 and 2 illustrations were retrieved from Flickr using their relevant systems at the same time to ensure a similar state of Flickr for both. It is possible with Flickr being a user generated collection that individual images might be incorrectly labelled and inappropriately selected. While the effect of individual images was reduced in the previous two experiments by the large volume of images the participants viewed the number of illustrations viewed in this experiment is much less and as such the effect of a single anomalous image is potentially increased. To reduce the effect of individual images each system selected their top five images instead of one for each illustration and when participants viewed the illustrations a random image from this set of 5 would be selected to be the actual displayed illustration. This would mean that users were still in effect rating the effectiveness of images selected by the representative systems but that the impact of a single unusually effective or ineffective image was reduced. The images selected obeyed similar rules to before in that illustrations for a single story may not contain more than one image per Flickr user as images from the same set inherently carry their own cohesion. Also selected images were reviewed to ensure none were obviously offensive or of a dimensions ratio that made them impractical for display on the site (the rule set was every image could not have a dimension more than a 150% of the size of the other dimension), however ultimately no images were removed this way.

When agreeing to participate users entered an email address, this address was then given a user ID and assigned a type which dictated which pairings of story and method the participant would consume. The use of the address was to try and ensure a single user did not repeat the experiment as signing up with the same address gave the same ID and only the first completed participation of each ID was recorded. The participants were then mailed with a link to the introduction of the experiment which contained

their ID and type so that the experiment system could display and record correctly. A brief introduction explained what would be asked of them and provided a glossary of terms to ensure participants knew what I meant by themes, genre, narrator, and other terminology used in the experiment. Participants were asked when reading the story to also consider the illustrations. Once they had begun the participants were shown the first story with its illustrations and then asked to answer a short questionnaire (explained below). This process repeated for all three stories and then uploaded the results for that participant.

The questionnaire was designed to measure the perceived cohesion based on the five variables I had identified as related to narrative cohesion. There were five questions one for each variable; logical sense, theme, genre, narrator, and style . Each question was answered using a single Likert scale of 1-5 using radio buttons with the exception of question 2 which asked the users to rate each theme on a list of 23 themes (the entirety of the list of themes identified by the independent expert for all stories) in response to the question from 1-5. The questions were:

1. How logical was the story? E.g. did the story make causal sense to you?
2. Please rate the strength of the presence of the following themes in the story. E.g. how apparent was it that these themes were present? Were they subtle or overt? (Followed by a list of themes)
3. How strongly do you feel this story fits into an established genre?
4. How strong and consistent was the presence of an identifiable storyteller? E.g. Was the story told from a perspective you could easily identify?
5. Is the style, presentation, and language used to express the story consistent? E.g. is the story throughout presented in the same way or does it frequently change tone?

Each question also qualified to the participants what was the low/high ends of the scale of 1-5. For question 2 the list of themes was compiled of an alphabetically sorted list of all the identified themes for all the stories. As such the list displayed was the same for each story and it contained all of the themes that were identified as present in the story they were reading as well as the themes for the others that would presumably be weaker or not present at all. By sorting them alphabetically they were not grouped by story and as such it was not immediately obvious which themes belonged to which story.

The pages the stories were presented on were designed to be plain (so as not to communicate additional ideas to the readers) but with some sense of design presenting the stories in a central column with margins around the images presented on alternating

sides of the page so that the site was not ugly to the point of interfering with the cohesion of the stories. For each page the questionnaire was kept separate from the story and the entirety of the story was displayed with images on a single page in a constant style with the questionnaire below. While this could lead to a long page as noted by Gee [47] navigating can break immersion when evaluating a narrative and as I was measuring cohesion I was keen to avoid this. A screen shot of a narrative displayed through the system can be seen in figure 6.3 On completion their 1-5 ratings for the different variables of cohesion for each story as well as their ID, type, and the particular images they viewed were recorded. The experiment itself took approximately 20-30 minutes and was advertised to a variety of groups (including both the computer science and literature worlds). This experiment was also given approval by the schools ethics committee, the approved application can be found in section E.4 of appendix E.

The Night

A young boy no more than ten woke to the dark with the need to pee. In this world, peeing at night was not so easy a thing. His mother and father slept in the room at the far end of the hall, the bathroom in between, two picture windows that looked out onto the lawn and the lawns of the other houses and a park beyond. His sister, Betty, slept in the bed beside him.

He got out of bed and paused with amazement at the door to the hall: Mother or Father had forgotten to shutter the windows at dusk. He should, he told himself, go back to bed but the pee threatened to explode. He thought about asking Betty. But something in him said no and to look. To test what always had been said and warned against. The windows must remain shut up when the sun goes down. The doors must be locked at night and every night, no exceptions.

He'd seen photos. They were only photos: stars, the moon, eclipses. "I'll lock you out at night," Betty would threaten. "Stop," Father would say. "It isn't funny."

But could he just look? Would he? A quick look. Just take a quick glance, then pee, then back to bed, and shuffle through the whys in the morning when it was safe to do so, safe to ask questions, or safe to say "I risked and it wasn't true. None of it has been true, Father."

Last month he and Betty and snuck into the hallway and sat beneath the shutters, their hearts patting in their chests like crows wings. "No," the boy had said finally. "No way."

"We just shouldn't," Betty whispered. "We just can't."

So, why not just look now? He watched down the hall. The moon shown bright through the big windows. It was that Blue Time. Four or so. He'd seen the blue light only once before, early morning high in the mountains, where everyone sat on an enormous balcony, the hotel rising behind spatter-lit yellow from the random windows in the higher rooms. It had been so rare, Mother standing by Father at a stone lookout. They had drinks. Even together then they looked lonely and helpless and somehow guilty of a quiet collaboration in solitude and failure.

That time had been about this same time, so late and early, and everyone happy, a wakeful holiday, refusing in unison to sleep. He remembered how his mother's head had slowly fallen to Father's shoulder, his head to her head. So high they were, the clouds swirled below silver and humped, high and safe above a world where night never really came.

He felt the energy or thrill of impulse. He felt drawn from the door, drawn to the window. He moved his eye to the left of the frame, looked out for the first time at an alien night and encountered moon glow, the shadows of the nearer trees soft and silent on the bright white ground. Nothing, he thought. There was nothing, nothing to fear out there, all the rules untrue or just false.

The windows on the block were shuttered, blank spots against the gray sky. In the park, it must have been very far, he saw a dark shape drift rapidly out from the pines into a clear space on the grass, something that suggested spider but with only two or three legs, small so far away, then it was gone, like something painted on the ground by a breeze. He felt a pressure in his chest, something of an illicit sweetness at the bottom of his throat.

"They sense vision," he remembered hearing. "They ignore us if we don't see . . ." Grandfather said. They had been in Wisconsin with relatives. They had been eating in a windowless cafeteria. Grandfather told stories. He explained, while Mother and Father ate, ready to silence the speaker if he strayed. ". . . or look. I don't know when precisely they came, but it was just after your mother was born, and we had to change. All of us, even governments. We lost the night everywhere to their teeth, their nails. But we're safe if we don't look. Safe if we shut the night out. We no longer have the moon to ourselves.

"They're drawn to the human eye. They only take he or she who sees them. And so we board our windows. It's safest that way. Sleep through the night. Keep the windows shut. Love the daytime."

"Father," Father said. "Please."

The boy watched the night. Something dark slashed across a nearer street. Then the boy heard his Father.

"No," Father said. He was standing at the bedroom door. "My god, son," he said.

"Father," the boy whispered, his voice shaking. "I just had to pee."

"Oh my god," the Father said. "Oh my god."

FIGURE 6.3: A screenshot of 'The Night' as displayed to participants through the system

6.3.2 Results

The experiment received 66 participants and offers some conclusions to my research questions. Where possible the images used in this experiment can be found in section D.3 of appendix D. The users ratings for each question of the questionnaire were recorded and these were converted into cohesion scores for the different combinations of story and method. For Logic, Genre, Narrator, and Style the mean of the rating for the relevant question was used, for theme however my question was more complicated and

this warranted a more complicated scoring system. Thematic cohesion has been divided into three scores; Theme(S) representing the mean score for the strongest theme (as identified by our independent expert) for that story, Theme(I) representing the mean score for all the other included or present themes identified in that story, and Theme(E) representing the mean score for all the themes not identified by expert for that story. The values for these scores for different story and method pairings can be found in table 6.3.

Story	Method	Logic	Theme(S)	Theme(I)	Theme(E)	Genre	Narrator	Style
1	1	1.958	3.208	2.662	1.659	1.625	2.125	2.667
1	2	2.000	3.048	2.702	1.891	1.857	2.857	2.857
1	3	2.143	3.238	2.637	1.771	1.714	2.286	2.476
2	1	4.619	3.762	4.262	1.661	4.286	3.762	4.381
2	2	4.208	3.125	4.104	1.792	4.208	3.542	4.333
2	3	4.333	2.667	3.905	1.794	4.048	4.381	4.571
3	1	4.800	4.952	4.270	2.143	3.476	4.000	4.286
3	2	4.524	4.857	4.148	2.190	3.095	3.190	4.143
3	3	4.583	4.958	4.231	2.246	3.375	3.542	4.125
All	1	3.792	3.974	3.731	1.821	3.319	3.296	3.778
All	2	3.577	3.676	3.651	1.958	3.180	3.196	3.778
All	3	3.686	3.621	3.591	1.937	2.952	3.403	3.724

Table 6.3: Cohesion Ratings for Stories and Illustration Methods

The scores for story 1 ('The Point') can be found in figure 6.4, the scores for story 2 ('The Night') can be found in figure 6.5, the scores for story 3 ('Computer Leon') can be found in figure 6.6, and the scores for all stories averaged can be found in figure 6.7. The graphs show the position for each score for method 1 (Thematic), method 2 (Non Thematic), and method 3 (Human Selection Base Case) of illustration.

6.3.3 Analysis

The results lead to some interesting observations, first of all as we might expect the overall cohesion scores of the deliberately selected abstract story 'The Point' were lower than the other two stories (a total average of 2.351, as supposed to 3.702 for 'The Night', and 3.864 for 'Computer Leon') and the story selected for deliberately high cohesion scored generally higher, lending some vindication to the ability of my questionnaire to record cohesion scores. However conclusions based on the different methods for presentation are not straightforward with no method significantly and consistently raising cohesion above other methods.

Our first research question was 'Do thematic illustrations selected by a thematic system improve the perceived thematic cohesion of a narrative?', to answer this I need to consider how an improved thematic cohesion would manifest within the scores. As a story

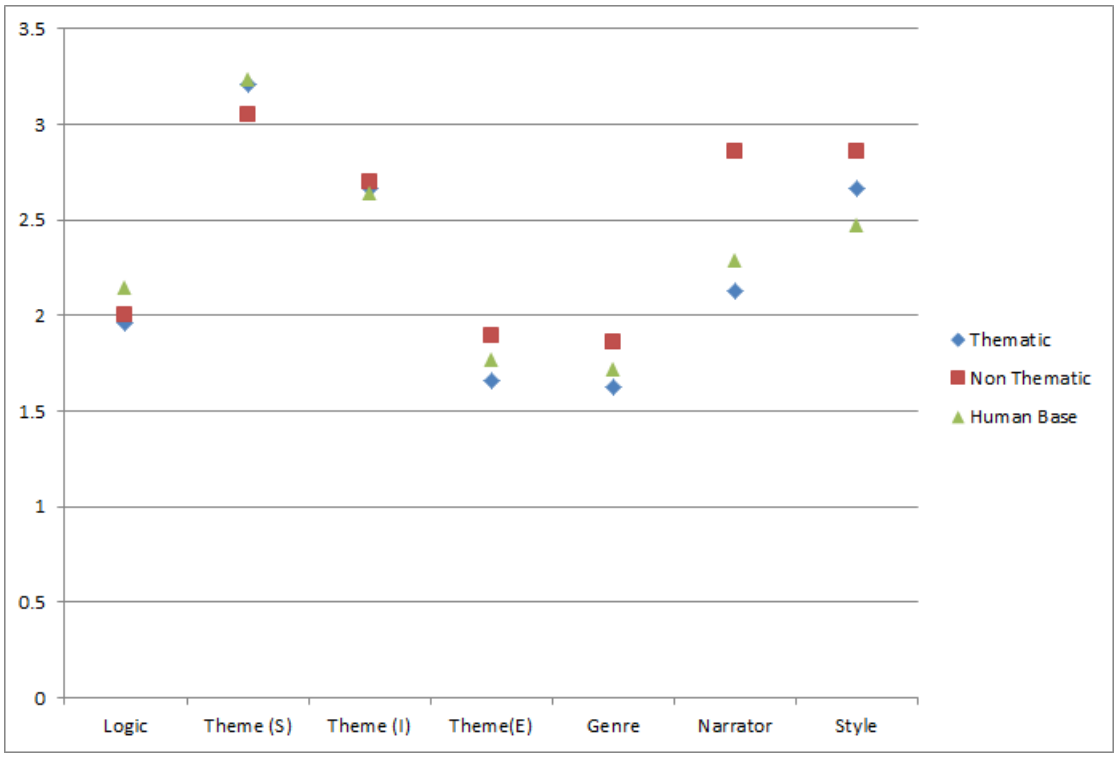


FIGURE 6.4: Cohesion scores for story 1: 'The Point'

becomes more thematically coherent its stronger deliberate themes would be identifiable throughout and false or unintended themes (what we refer to as 'thematic noise') would become less detectable. As such in my thematic scores we would expect to see Theme(S) rise and Theme(E) decrease for a successful increase in thematic cohesion. Analysing the overall data for the range of stories we find that the thematic approach (method 1) has indeed increased Theme(S) and decreased Theme(E) over the generative approach not using themes (method 2). However when putting this through a t test the hypothesis 'Method 1 scores Theme(S) higher than method 2' scores a t of 1.181 (df=130, p=0.2) whereas 'Method 1 scores Theme(E) lower then method 2' scores a t of 2.607 (df=2010, p=0.005) showing that while the decrease in Theme(E) is statistically significant with only a 0.005 probability of error the increase in Theme(S) is only significant with a 0.2 probability of error. The reason for the first hypothesis scoring a low t could be in part due to the low number of degrees of freedom; the design of the experiment meant only the ratings for a pair of strong themes could be compared whereas there were many more themes excluded to be compared yielding a higher number of data points. It is possible that had the experiment received a larger number of participants that this trend would have continued and become statistically significant. Indeed in previous experiments [50][52] the thematic elaboration of the TMB has always been slight as it is here and significance has been achieved due to the larger number of data points. What this shows us is that thematic emphasis definitely reduces thematic noise in a story and possibly raises the strength of the core theme suggesting that thematic cohesion has been improved in some way but that no definite conclusions can be made with this data on

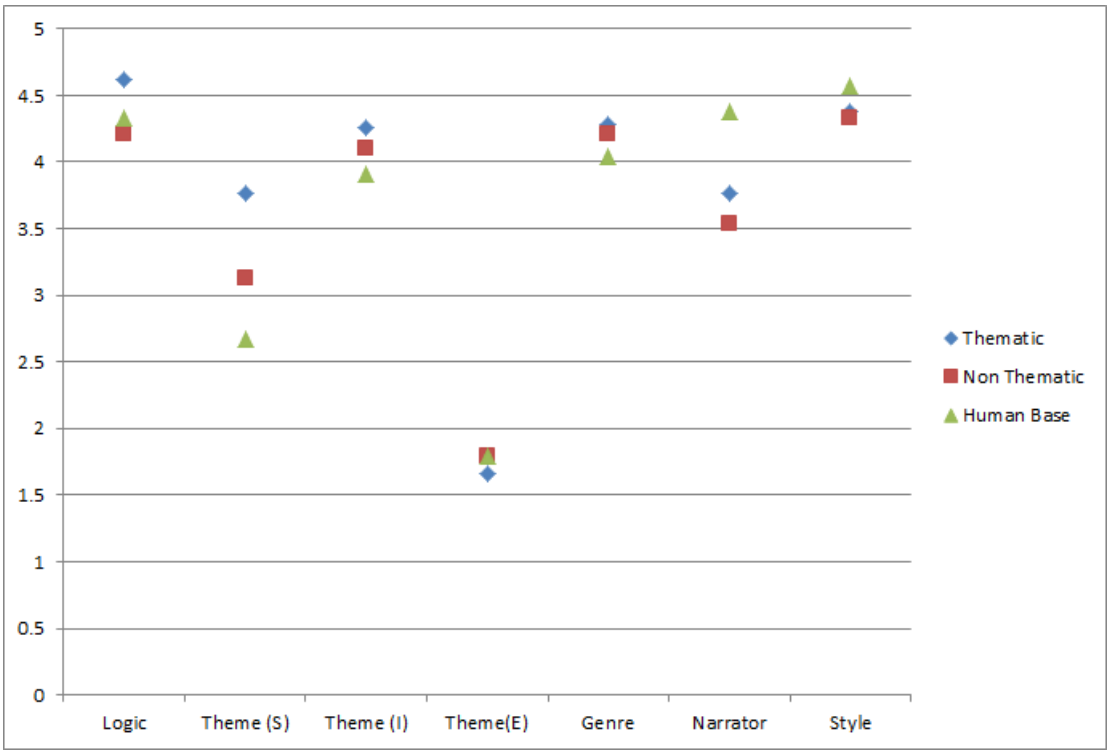


FIGURE 6.5: Cohesion scores for story 2: 'The Night'

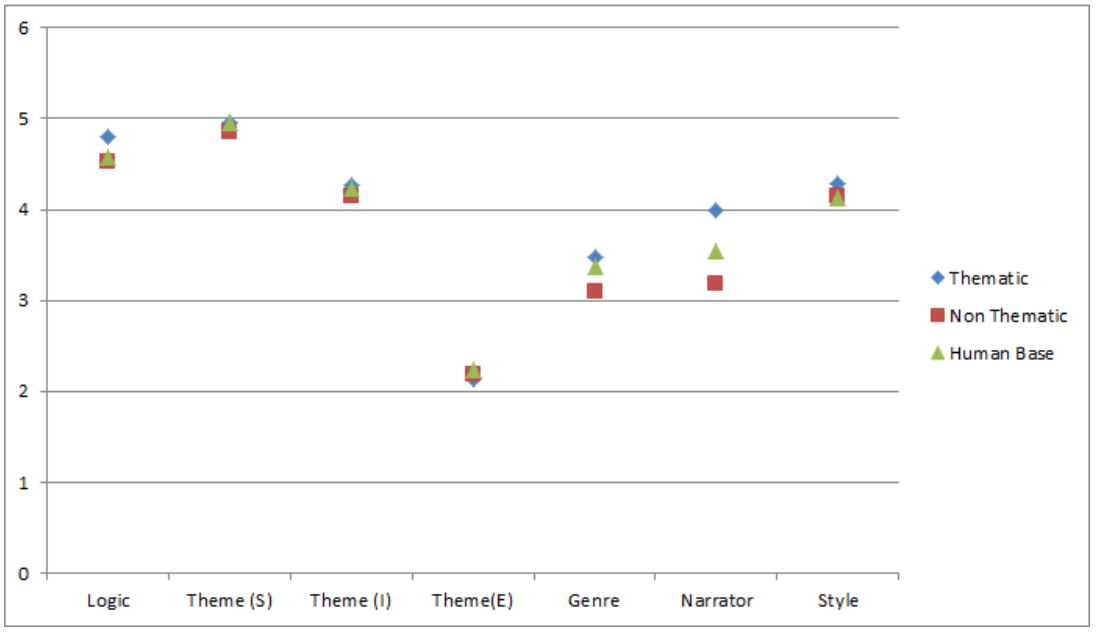


FIGURE 6.6: Cohesion scores for story 3: 'Computer Leon'

the higher values of Theme(S).

There are some other observations relevant to this question that can be made. In figure 6.5 it can be seen that for the thematic approaches improvement of Theme(S) over the non-thematic approach is much more substantial for story 2 ('The Night') then for other stories. This larger improvement may be due to the way in which this particular story is

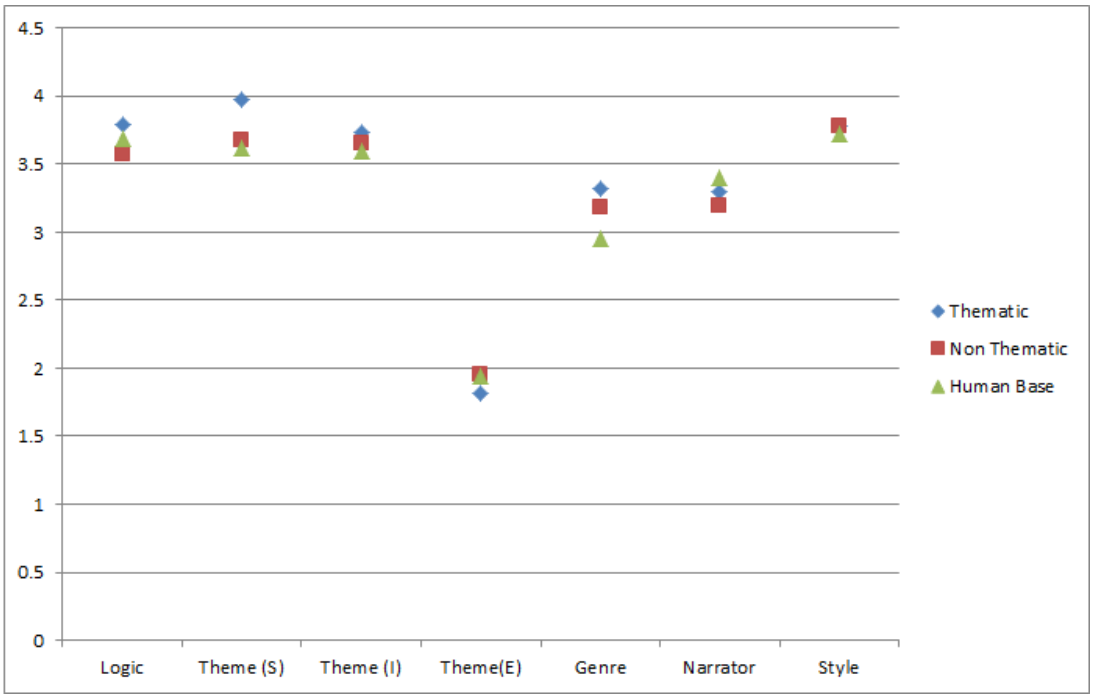


FIGURE 6.7: Cohesion scores for all stories (average)

written with very visual imagery and thematically evocative writing allowing for more significant visual thematic emphasis over less visual stories. Also to be noted is the relatively minor or negative effect on cohesion of thematic emphasis in story 1 ('The Point') as shown in figure 6.4. This could be attributed to the relatively abstract style of story making it difficult to automatically generate relevant or effective illustrations and as such reducing the effect of illustrative emphasis.

With regards to the other research question, 'What effect does this have on the perceived cohesion of the narrative as a whole?' I examined what connection, if any, there was between these different cohesion metrics, so that if thematic cohesion improved, did the cohesion of the narrative as a whole improve with it? To investigate this I performed a Pearsons correlation between Theme(S) and each of the other non-thematic metrics the results of which are presented in table 6.4. What we find is a very strong correlation with Logic ($p = 0.005$), a positive (though not as strong) correlation with Genre ($p = 0.05$), a weak non-significant correlation with Narrator ($p = 0.1$), and almost no correlation at all with Style. What this might suggest is that stories with a strong thematic cohesion binding them together are easier to make sense of and as such more logically cohesive and are easier to contextually relate to a genre. And by the same degree it would suggest there is no connection between thematic cohesion and the narrator presence within the narrative, nor the coherence of the authors style of discourse. However the correlation alone does not provide significant evidence to make a claim that these variables are directly dependent on each other. A new experiment would need to be devised to investigate this where significant shifts in one could be shown to directly

result in a significant shift in the other from the same perspective, as opposed to the indirect comparisons made here (no one participant read the same story with both methods). However this correlation does suggest there is potentially a connection here, that further investigation is warranted, and that in the case of these stories a stronger thematic cohesion did correlate with a stronger logical and genre cohesion.

Metric	Correlation Coefficient
Logic	0.299
Genre	0.189
Narrator	0.131
Style	0.058

TABLE 6.4: Persons Correlation between Theme(S) and other non-thematic metrics

6.4 Cohesion Conclusions

The experiments presented in this chapter support a number of conclusions. First of all, based on existing work we can ascertain that cohesion as a complex factor of narrative may be broken down into Logic, Theme, Genre, Narrator, and Style; five variables that capture different aspects of how the elements of a narrative may be bound together coherently. The fact that these variables also returned expected general results when recorded through participants perception through a questionnaire further suggests that these are a sensible measure of cohesion. There may be other ways of measuring cohesion or ways these variables may be further broken down into more formal specific tests and future work may want to explore this through a qualitative investigation into what aspects of a narrative cause these variables to rise or fall. By recording features that provide evidence for the presence of absence of each of these different variables we can see how they manifest in different stories and how they might be measured. While building an automatic cohesion analysis system using these features is outside the scope of this research, the exercise of exploring what comprises these variables helped us articulate their meaning to experiment participants.

The experiment also began to show evidence to suggest that thematic emphasis leads to stronger thematic cohesion. However while thematic noise was shown to be reduced in a significant way the improvement of core theme strength was not shown to be significant due to a combination of the slightness of the improvement and the low number of data points. It is possible as well that other methods of thematic emphasis should be investigated, while I have shown some results through emphasis using illustrations here it is possible more dramatic improvements in cohesion would be experienced using methods that alter the text of the narrative itself.

Finally, correlations were shown to exist between Thematic Cohesion and Logical and Genre cohesion. This could potentially suggest that a system capable of improving

one could see an improvement in the others. This would provide a strong argument for pursuing methods of thematic emphasis as it might be used to raise the coherence of generated or adaptive narratives. However further work is needed to establish the level to which these variables are dependent on each other. An experiment that directly compared the values for these variables from a fixed perspective using different levels of thematic emphasis would work towards establishing more concrete conclusions.

Within this work I have begun to understand how narrative cohesion may be modelled and captured. The experiments contained have also shown that it is potentially viable to alter the coherence of the narrative through thematic emphasis. While more work is necessary to build a complete understanding of the effect of thematic emphasis, significant steps have been made here to establish metrics, the effect on thematic cohesion (in particular thematic noise), and the relationship between different variables of cohesion.

Chapter 7

Conclusions and Future Work

In this thesis I have detailed a body of work exploring a thematic model for use in narrative systems and documented a set of experiments that investigate the validity of a semiotic approach to themes in narratives. In this final chapter I summarise these findings and cover the major conclusions and contributions of this research. I also identify key areas for future work that can be undertaken to take the results of this work forward.

7.1 Summary

The focus for this thesis has been a machine understandable thematic model for narrative; the model can be used to represent any number of themes and their relationships to other thematic elements and the features within a narrative which may be used to connote them. Narrative is a prevalent representation of information used between people and with an increase in the information handled on the web and the rise of large user generated collections of media, how we represent information and machine understandable models of narrative become more important. While work has been undertaken to create models of narrative these are often concerned with the explicit content and plot and not the subtextual themes; which has been stated to be important in binding the components of a narrative together, giving it meaning and cohesion [113]. The development of a successful method for modelling themes would have a wide range of positive consequences for narrative systems. Narrative generators that currently create often bland and directionless narratives could be given thematic depth and cohesion, search systems could utilise semiotic expansion to find thematically relevant results to queries, and thematic analysis of narratives might be used to categorise documents by subtext or assess their suitability for use in a grand narrative.

7.1.1 The Thematic Model

The thematic model developed was based upon work in structuralism (an aspect of literary theory) such as that by Barthes [13] and Tomashevsky [113] where narrative was broken down into the structures and elements that comprise a form of literary analysis. While structuralist theories were used for inspiration of the elements and structures, work in semiotics by Saussure [102] and again Barthes [12] offered an explanation for how these elements might be related and how their connection to the narrative itself through an explanation of how we interpret signs and how connotations may be formed. The thematic model is comprised of 'Themes' which are connoted by other themes and 'Motifs' which are in turn denoted by 'Features' that may be present within the literal content of the narrative. The rules that controlled the elements and relationships of this model were formally defined and presented.

7.1.2 Experiment 1: Thematic Definition Authoring Process

Each instance of this model forms a collection of thematic definitions for particular themes. These are, as of the time of writing this thesis, authored by hand following a defined and structured process. This process is supported by a guide which has been designed around five steps and with examples so that anyone might contribute thematic definitions. An experiment conducted showed that it was possible to create valid definitions in the terms of this model using the guide but that in some cases it was insufficient and authors would need more sophisticated support probably in the form of an authoring tool. While the guide was not always sufficient for creating fully formed valid definitions, it was observed that, should a community created super-model be formed, parts of incomplete definitions could still be used to contribute towards identifying elements and relationships.

7.1.3 Experiment 2: TMB vs Keyword Search

To evaluate the performance of the model a prototype narrative system was created. The Thematic Montage Builder (TMB) created themed photo montages from images in Flickr utilising definitions made in the terms of the thematic model. Each montage created by the TMB has a subject or content that represented the focus of the montage and then a list of predefined desired themes. The TMB formed a large corpus or 'fabula' of images by performing a search in Flickr for the content keywords and then calculated the relevance of each to the desired themes returning the highest rated as the montage.

An evaluation experiment of my thematic approach was carried out where I compared the ability of the TMB at creating montages for thematically charged titles to that of Flickrs keyword search. If the approach was successfully representing the complex and

subtle structures of each of the themes that had been designed I had expected to find the montages generated by the TMB more relevant to the chosen title than those created by simply searching Flickr using the titles. An experiment was created where users viewed a series of images individually and in montages for a range of different titles, each of which contained both a subject matter and a desired theme, e.g. "London in Winter", and rated their relevance to the given title. The results confirmed my expectations and the TMBs images and montages were consistently rated as more relevant than those of the keyword search. The TMB also demonstrated a greatly improved relevance in the more narrative context of a montage and was more capable of coping with requests involving multiple themes, however in cases where the theme and desired content contradicted each other it struggled to find the relevant motifs it needed.

7.1.4 Experiment 3: TMB vs Co-Occurrence

A follow up experiment used a similar methodology to the previous but sought to compare the TMBs performance to a term expansion system. Term expansion has often been shown to improve the relevance of search results taking a particular relationship between terms and using it as a basis for measuring the semantic similarity between a term and a candidate for expansion. This relationship varies in approach from using synonyms and meronyms in dictionary based approaches such as those investigated by Voorhees [116] and co-occurrence such as in work by Buckley [26]. A review by Mandala [78] identified co-occurrence as the single most successful approach for term expansion in improving the relevance of queries and I used his implementation as a basis for creating my own co-occurrence term expansion system. The thematic definitions used by the TMB are essentially the results of a semiotic term expansion performed at the authoring stage and having shown its improved performance over keyword search and that it could successfully create relevant montages the next step was to compare it to the most successful existing method of term expansion. The thematic definitions currently are very time consuming to create and so if co-occurrence was able to effectively simulate semiotics (even if it isn't directly based upon it) then it would provide a fast automatic solution towards creating a term expansion to represent themes. The experiment used the same methodology as before and both systems expanded the thematic part of their titles using corpuses based on the desired content of the montage, as with before the montages were then rated by users for their relevance to the given titles. As with the prior experiment the TMB demonstrated a consistent higher level of relevance over the other system, even when the titles it has always performed strongest in were removed. This led us to the conclusion that while semiotic expansion may be costly in terms of time when creating the definitions this cost may be justified by the fact that it has led to the most accurate representations of themes.

7.1.5 Experiment 4: Impact on Narrative Cohesion

The final experiment sought to explore the claim that thematic cohesion binds the components of a narrative together leading to greater flow and greater cohesion of the narrative as a whole. To do this a new prototype was created called the thematic illustrator which utilised the TMB to create illustrations for short stories based on their content but also emphasising a key theme. The intention of this system was to emphasise and strengthen the strongest theme in a story to bind all parts of the story together under this theme. In order to evaluate the effect of this on the cohesion of narrative as a whole five variables representing narrative cohesion were identified through a review of literature on cohesion and narratology and an investigation of existing systems and evaluations of narrative cohesion. These five variables (logical sense, theme, genre, narrator, and style) were used to create an experiment where users could view short stories with illustrations and provide perceived cohesion ratings based on each of the variables. I would then compare how the cohesion changed between stories with illustrations based solely on content and those based on content and emphasising the core themes. The results showed that while thematic cohesion was definitely increased by emphasising the core themes the effect on the remainder of narrative cohesion was limited. Some effect was shown on the logical sense and genre variables where an improvement in thematic cohesion also led to an improvement in these variables but this effect was small. It is possible that while illustrations are sufficient for influencing thematic cohesion for the impact to be sufficient enough to effect the cohesion of the narrative textual manipulation based on core themes is necessary.

7.2 Contributions

This thesis has documented a range of research contributions connected to thematics and narrative systems. These contributions lead towards exploring the hypothesis stated at the start of this thesis. Each of the contributions made is detailed below followed by a summary of my conclusions regarding the original hypothesis.

1. *A Thematic Model*
   This thesis has documented a model for representing themes within a narrative. The model is based on structuralist ideas and semiotic relationships where "Themes" are connoted by other themes and "Motifs" which are in turn denoted by "Features" within the narrative. The rules that control structure of this model are formally defined and the definitions created have been shown to be machine readable and capable of creating relevant representations of themes.

2. *An Authoring Method for the Thematic Model*
   A method has been documented along with a guide for supporting authors for

creating definitions for desired themes in the terms of the model. The method is broken down into 5 stages and based on semiotic term expansion, each stage of the method is integrated with the rules of the model to ensure the creation of valid models. A guide has been developed for this method to guide users in creating thematic definitions that uses examples, clear instruction, and forms to guide expression of the definition.

3. *An Evaluation of the Guide for the Authoring Method*
   An experiment which evaluated the guide by having participants use it to create definitions of given themes was undertaken. The experiment shows an evaluation of a process for capturing subjective information from users and the results and conclusions of this experiment are included in this thesis. The results show that, using the established method, building valid definitions is possible. However more support than just the guide is needed.

4. *The TMB*
   The Thematic Montage Builder (TMB) is a prototype that demonstrates how the model may be utilised to evaluate the thematic relevance of content and generate simple thematically relevant narratives in the form of photo montages. The system includes a metric for measuring the thematic quality of content based on tags using definitions created in the terms of the model. This thesis includes the metrics used by this system as well as the methods used to implement it. This contribution represents not just a prototype system, but also a method for using the thematic model in a narrative system.

5. *Evaluation of the TMB against Flickr keyword search*
   An experiment comparing montages generated by the TMB against those generated by a search of Flickr for titles with a thematic content. The experiment takes participants ratings of relevance for the individual images and montages and compares them to investigate if the system with the thematic basis is successfully expressing the desired themes and if its expansion of the theme makes its results more relevant than a simple keyword search for the theme. This contribution comprises of the data and analysis for this experiment which are included in this thesis. The results show the TMB out performs the basic keyword search for both individual image relevance and montage relevance, in most cases.

6. *Evaluation of the TMB against a co-occurrence term expansion system*
   An experiment comparing the TMB to another term expansion system based on co-occurrence that seeks to investigate if co-occurrence can simulate semiotics and whether the cost in time of creating thematic definitions can be justified by an improvement over existing successful term expansion systems. The methodology of this experiment is similar to the previous with the results of the TMB being compared to those of a co-occurrence based system expanding the desired theme. My implementation of the co-occurrence system is explained and my contribution

comprises of the data and analysis for this experiment. The results show the TMB out performs the co-occurence system for both individual image relevance and montage relevance.

7. *Identification of narrative cohesion variables*
   An investigation of the literature and related systems has led us to identify five key variables that can be used to measure narrative cohesion. The description of these variables, how they were identified, and features demonstrating how they might be measured is detailed in this thesis.

8. *Experiment investigating the impact of thematic illustrations on narrative cohesion*
   The creation of a prototype system based on the TMB that generates illustrations for short stories based on their content and themes allows for the emphasis of a key theme within these stories. An experiment was conducted to ascertain whether this had a genuine impact on the thematic cohesion of the story and whether this in turn had an impact on the cohesion of the narrative as a whole. To measure this I used the variables identified above to create a questionnaire that captures the perceived cohesion of participants reading the stories with illustrations that emphasise the key themes and those that don't. This contribution comprises all of the data from this experiment, my implementation of the illustrator, and my analysis of the results. The results show that thematic emphasis can be used successfully to reduce thematic noise, and that there is a correlation between thematic cohesion and logical sense and genre cohesion.

These contributions have been created during my work to investigate my original hypothesis which was:

Thematic Models generated through Semiotic Term Expansion can be used to improve the relevance of search results in the context of a defined theme, and to improve the perceived cohesion of narratives through thematic illustration.

This hypothesis was broken down into three research questions:

1. Can semiotic term expansion be used to generate thematic models?
2. Can thematic models be used to improve the relevance of search results in the context of a defined theme?
3. Can thematic models be used to improve the perceived cohesion of narratives through thematic illustration?

Evidence can be found in this thesis to provide answers to these questions. I have demonstrated that semiotic term expansion can be used to create thematic models through the design of my thematic model and the identified authoring process. The model, based

on structuralist ideas, is a suitable basis for representing themes within a narrative and instances of it are created by capturing the subjective definitions of themes from users by encouraging them to semiotically expand the concept of the theme in a structured process. While the process of creating these themes has been shown to potentially require more support than the current guide my existing process has still shown that is possible to create definitions semiotically. The evidence to answer question one can be found in contributions one, two, and three.

Question two can also be answered with some certainty. Through my experiments with photo montages from Flickr I have demonstrated that the thematic system consistently returns more relevant results then both a keyword search system and a co-occurrence based system. While it is important to remember that this only concerns queries made in a thematic context, and one that has been suitably defined for the system, it none the less shows that a thematic basis leads to an improvement in relevance. The evidence to answer question two can be found in contributions four, five, and six.

The answer for question three is not as clear. While the final experiment demonstrated that thematic emphasis through thematic illustrations can increase the thematic cohesion of a story the resultant effect on the narrative cohesion as a whole is small and only present in some variables. As stated in the conclusions of that experiment it is possible that illustrations are insufficient to have a significant impact and that actual textual manipulation of the narrative would be necessary to have the desired significant effect on narrative cohesion. This third question might have a more positive answer if it was "Can thematic models be used to improve the perceived cohesion of narratives through thematic emphasis within the narrative?" however the work with textual manipulation necessary to undertake this is outside the scope of this thesis. So the answer for the existing question is limited to a yes but only to a slight degree and in certain aspects. The evidence towards question three can be found in contributions seven and eight.

The work in this thesis has shown a positive outcome for the hypotheses I started from. I have demonstrated a successful semiotic approach to creating thematic definitions in the terms of a thematic model I designed based on existing work in literary theory. I have also shown this model to be successful in representing desired themes and creating simple narratives that are more relevant to desired themes then a simple keyword based system or one using co-occurrence. While the effect of thematic illustrations was not a serious impact on narrative cohesion it has shown to effect the thematic cohesion of these stories (through a reduction in thematic noise) and some small initial effects on other cohesion metrics show promise that further systems might have a more significant effect.

7.3 Publications

Throughout the process of this PhD a series of publications have been created and presented in different workshops, conferences, and a journal. In this section I summarise these publications and achievements that happened as a result of the PhD process.

1. *A Thematic Approach to Emerging Narrative Structure* Web Science Workshop at Hypertext 2008 [50]
   This paper presented at a workshop in hypertext 2008 in Pittsburgh and established my initial approach to thematic modelling for narrative systems. The thematic model was introduced along with some initial analysis of fitting tags to themes as well as the motivation for the research. This research was entered in Hypertext 2008's student research competition, where it won second place.

2. *Investigating a thematic approach to narrative generation* Dynamic Adaptive Hypermedia Workshop at Hypertext 2009 [51]
   This paper was presented at a workshop in hypertext 2009 in Torino. The paper outlines an investigation into how the thematic model may be integrated with existing methods of narrative generation. The main approaches of narrative generation are identified and the process is broken down into stages in order to present an analysis of how integration with different approaches at different stages will have a different effect.

3. *Using a Thematic Model to Enrich Photo Montages* Hypertext 2009 [52]
   This short paper was written for hypertext 2009 in Torino. The paper displays the results of a pilot study completed for the experiment comparing the TMB to Flickrs keyword search. The implementation of the TMB is explained and presented as is the initial data from the experiment.

4. *A semiotic approach for the generation of themed photo narratives* Hypertext 2010 [54]
   This was a full paper presented at hypertext 2010 in Toronto. The paper reports the full results for the experiment comparing the TMB to Flickr. As well as the full results and analysis of this experiment the experiment is contextualised with the work in progress developing a more formal authoring method and a comparison to co-occurrence.

5. *Capturing the semiotic relationship between terms* New Review of Hypermedia and Multimedia volume 16 (2010) [53]
   This was a technical note in the New Review of Hypermedia and Multimedia journal. The article presents the method used to capture the subjective definitions of themes from users through semiotic term expansion. The development of the guide is detailed as is the experiment evaluating the authoring method and guide.

6. *Measuring Narrative Cohesion: A Five Variables Approach* Narrative and Hypertext Workshop at Hypertext 2011 [55]
   This paper was presented as part of a workshop I ran at hypertext 2011 in Eindhoven. The paper presents my analysis of narrative cohesion and identification of five variables for measuring narrative cohesion. The paper explores how these variables could be measured and applied to the results of narrative generation.

All conferences and workshops published in were personally attended and presented at by the author of this thesis. Funding to attend these conferences was awarded by the RAE, SIGWeb, and the University of Southampton.

7.4 Future Work

Although significant conclusions have been reached in this thesis, the work has raised further questions that deserve further research. In this section we explore three main avenues of research that lead directly on from this work.

7.4.1 Authoring Definitions

In chapter 3 of this thesis I introduced my authoring method for creating definitions of themes in the terms of my model. This method was supported by a guide and some forms to allow people to construct valid thematic definitions. While an experiment showed the guide could support the creation of valid definitions there were still a significant proportion of users that produced invalid or incomplete definitions.

The first conclusion from this is that while the method of authoring thematic definitions is suitable it needs to be better supported. While the guide did lead to valid definitions from some there needed to be further support for authors to counter the common problems with the process. A lot of the difficulty with the authoring process was connected to the users overlooking parts of the process for particular elements that lead to incomplete definitions. This included expansion of sub themes and evaluation/removal of associated elements. While the guide provides guidance on how to go about these stages it is possible the users overlooked the relevance of this stage to all the necessary items. Some more persistent guidance could be provided through an authoring tool.

An authoring tool for definitions would provide an interface for creating and connecting thematic elements guiding users through the five stages and alerting them to potentially overlooked steps and elements. Importantly the tool could also contain a validator that could alert users to where their definitions may be invalid on the fly as they progressed through the steps. Furthermore such a tool could export the completed definitions in useable xml rather than be written up by hand after authoring (as is the case with

the current paper system). Support could also be provided for the tool to be tied into other systems such as to support thematic tagging where users could be encouraged to thematically expand relevant tags they use.

As with other author centric systems any implementations utilising the thematic model will only be as successful as the quality and variety of the available definitions allow. To this end I suggest the formation of a 'super-model'; a community created and constantly evolving set of thematic definitions interconnected and available to systems using the thematic model. Users would be able to upload definitions of different themes, perhaps through the authoring tool described above, and these definitions would be merged with other community definitions. The merging would be based on user popularity leading to the unusual subjective definitions being filtered from the established community definitions. As well as providing a route to a large base of useable definitions this work could also mitigate some of the existing troubles with the current authoring process. Incomplete or invalid definitions would now be of use as while the complete definition would not be uploaded to the 'super-model' but valid parts of it; complete elements and their relationships, could be uploaded and made use of in the community definitions.

A community created 'super-model' could also be used to capture subjective thematic definitions from people of different cultures. It has been established that while we can capture a thematic definition in as formal a way as possible it is still subjective to that individual. And although we may merge such definitions with those submitted by others to form a community consensus it is possible that radically different cultures may have very different definitions of themes. In such a case the system gathering definitions would need to be able to detect two very different clusters of definitions and make a division into two different definitions of the same theme but from different perspectives.

7.4.2 Semiotic Expansion

In this thesis I have explored the strength of term expansion using the semiotic relationship between terms. As well as thematic definitions this allows us to explore users implied meaning behind the use of particular terms in queries and tags and use potential connotations of these terms to improve navigation and searching. I have shown in this thesis that semiotic expansion using thematic definitions improves the accuracy of search results for particular thematically charged queries over some existing methods.

Such an approach is limited in that the semiotic relationships on which it depends needs to be authored by hand as opposed to the more normal automatic term expansion approaches. However the cost of this may be mitigated by the creation of an authoring tool or community base of definitions as detailed above. The advantage that semiotic expansion might bring depends to an extent on the level of use of implied language in the context of different systems such as folksonomies or search.

Future work in this area would involve an investigation into the level of implication in the use of language in tags and search on the web, recording the frequency with which there are deliberate connotations. This investigation would also seek to measure the potential impact of semiotic expansion in these systems; evaluating the effect of thematic tagging or query expansion for thematic searching.

7.4.3 Narratives and Subtext

This thesis has demonstrated a thematic model that might be used as a representation of themes within the subtext of a narrative. I have shown this can be used to embed themes within a narrative and that it can be used to reduce thematic noise in narratives as well as potentially improve thematic cohesion, all through thematic illustration.

Further investigation in this area might explore automatic thematic analysis as a part of broader methods in machine understanding of narrative (something we have already begun to initially explore [88]) or perhaps more importantly different methods of thematic imbedding and emphasis for narratives beyond illustration. While I have shown thematic illustration does have some effect this margin is small and a more significant effect might be achieved through alternative methods. Such methods might include direct adaptation of text emphasising or omitting thematic elements of relevance or non-relevance or generating thematically relevant itself using natural language processing. More direct involvement of the thematic system with the body of the narrative might lead to more significant thematic developments in cohesion.

One of the key potential uses of this research is in narrative generation, where its use to embed a thematic subtext might lead to more coherent generated narratives. As well as further investigation into how to embed or emphasise themes further research needs to be done on method of integration with established narrative generation techniques. A first investigation into this was shown in chapter three and a workshop paper [51]. However a more complete evaluation would seek to measure the actual impact of integration with each method of narrative generation. This would involve either, integration with existing systems, or building a prototype system (of a given approach) with which to integrate with. The user evaluations of impact would seek to measure thematic emphasis and the effect on narrative cohesion as previous experiments have but also involve a more qualitative analysis of the effect on the quality of the generated narrative.

7.5 Final Conclusions

In this thesis we have explored the value of a thematic model in narrative systems, and the challenges faced in building better models of narrative and semiotic term expansion.

Narrative is a prevalent form of information communication between people where framing a piece of information from a particular perspective and within a rich discourse helps both the author to explain and the audience to understand. As rich discourses narratives also contain much more complex and subtle features such as subtexts including concepts such as themes and genre. As information is increasingly handled, processed, and presented by systems the value of machine readable narrative increases as it facilitates the semantic understanding of the information concerned. This begins with the building of better machine understandable models of narrative; where we can break down and identify the crucial elements and relationships. In this thesis I acknowledge work already done to model plot and narrative content and give my attention to the subtle, but important, subtextual themes of a narrative.

Our research has shown a range of contributions to do with the thematic models use in different narrative systems. These range from its ability to embed themes in photo montages to reducing thematic noise and potentially increasing narrative cohesion through illustration. These potentially might be utilised in conjunction with other models of narrative to improve narrative generators or other narrative systems by providing machine readable definitions of themes. The success of the thematic model is significantly reliant on the quality of the definitions built for it, and this in turn, due to the subjective semiotic nature of the model, cannot be automated and is reliant on human authors. While this is perfectly valid and can be made more useful through community created super definitions these authors require support and there is further work to be done in creating necessary tools.

The core conclusion of this research is that applyable machine readable definitions of themes can be created using a thematic model, with significant results, but that this requires support in the form of useable (human authored) definitions. In order to move forward research into narrative systems needs workable models of the subtext as well as the content of a narrative as this represents a significant amount of the information and purpose communicated within a narrative. Semiotic approaches provide important opportunities for us to understand the implication behind content and language used within such systems. This work represents a contribution towards that goal, but requires further development both in how it can be applied and how successful implementations can be supported. The ultimate goal is that systems will understand and utilise the subtler aspects of narrative such as themes in as meaningful a way as we do, and that their ability to analyse, generate, or present narratives becomes subsequently more powerful.

Appendix A

Thematic Definition Authoring Guide

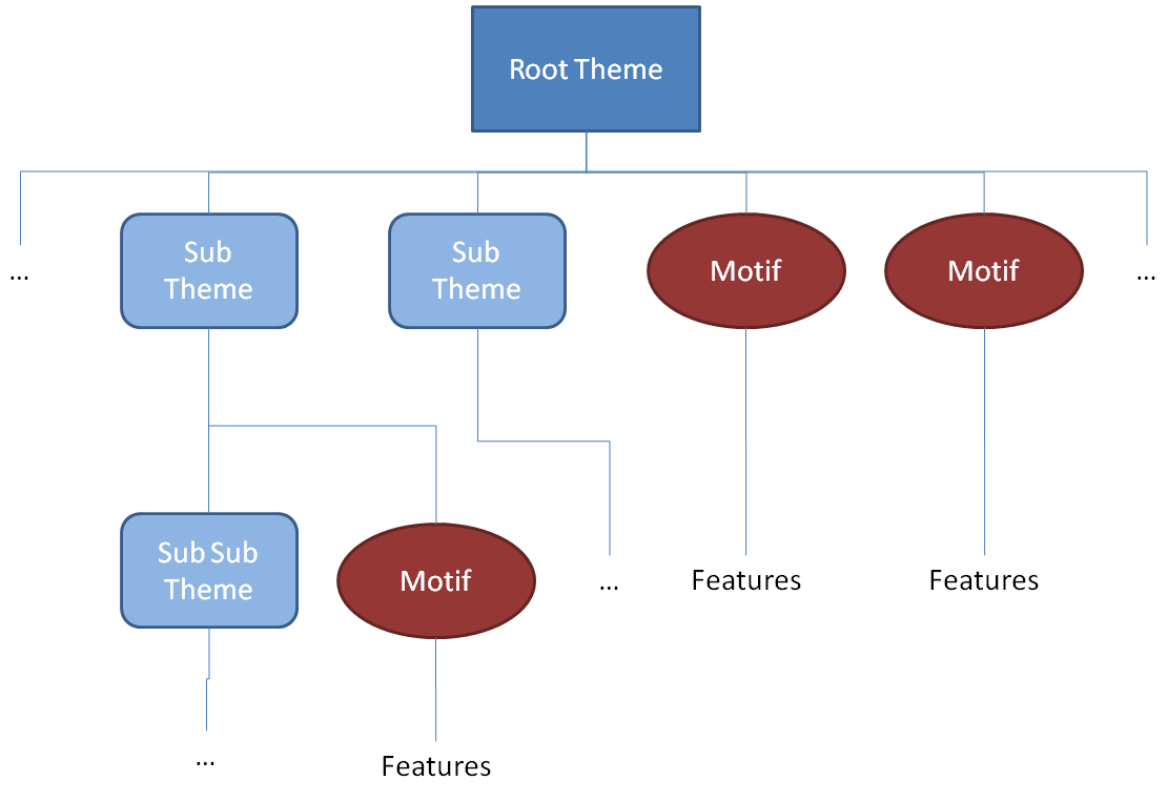
The following appendix includes the authoring guide and forms given to participants of the authoring experiment as detailed in chapter 3 section 3.2.

Guide begins on next page, and is displayed as it was presented to users.

Authoring Themes: An Exercise

Thank you for participating in an exercise of thematic authoring. The focus of this exercise is to explore whether themes, as an important part of the subtext of narrative, may be formally defined in a specific structure, and how to go about doing this. Work has been undertaken to explore a thematic model as a structure for representing themes in a machine readable way, for this exercise we would like you to follow a set of instructions that will lead to a theme being defined in the terms of this model. We will then analyse the effect of the process of defining the theme and the quality of the thematic definition to help us refine the thematic model and the method of authoring instances of it.

The thematic model allows us to expand a theme into its component subthemes and motifs. This expansion ultimately builds a network of concepts where individual features present in a narrative might show the existence of motifs which in turn show the existence of themes. This network of concepts, if represented as a tree, might look similar to this:



In this experiment we will be expanding a root theme into its component motifs and sub themes and in turn expanding those sub themes into further themes and motifs and so on.

You will be given a theme to define, this may later be referred to as the “root theme”, follow the instructions below and to complete a definition of this theme, you may want to use the supplied forms at the end of the instructions as a basis for completing the instructions but you do not have to. This whole process should in practice take approximately an hour.

Notes on using attached forms: There are two types of form attached to help you with this process should you want to use them. The first type is for following the first 3 stages for a theme, allowing you to expand and classify your themes and motifs. Use the first of these forms for expanding your root theme, then 5 additional copies are attached for expanding additional sub themes. If you have a lot of sub themes and require more there should be extras available or just use note paper in a similar format. The final forms offered are a guide for writing up the definition as a tree network. Use this to help you identify associated elements and remove or refactor them before coming up with the finished result.

Guide

Stage	Task	Example
1	**List Associated Words:** *The objective of this stage is to build a base of elements that imply the root theme.* Spend no more then 10 minutes listing words associated with your given theme. As a guide think of films, books, adverts, or other media you know that have incorporated the theme and list the elements from that which have implied the theme in question.	For the theme of 'Competition' we might get: Prize, Trophy, Contest, Sport, Game, Rules, Hard work, Opponent, Conflict, Challenge, Training, Uniform, Team, Patriotism
2	**Divide Themes and Motifs:** *The objective of this stage is to distinguish which elements are Themes and which are Motifs.* A definition of a theme is made up of motifs and other themes (sometimes referred to as sub themes). Themes and Motifs are also sometimes collectively referred to as elements. A Theme is an intangible concept, a high level idea. A Motif is a generalisation of something tangible in the story such as a place, person, object, or specific event. Divide your listed words into themes and motifs. Themes and motifs may later on be collectively referred to as elements.	Themes: Sport, Hard work, Conflict, Patriotism Motifs: Prize, Trophy, Contest, Game, Rules, Opponent, Challenge, Training, Uniform, Team
3a	**Determine the context of relationship with root theme:** *The objective of this stage is to consider how given elements are related to the root theme.* For each element consider how it is related to the theme it is a part of. Do this for both themes and motifs. How does this element connote and substantiate this theme? What role does it serve to help the theme become apparent? For both your list of motifs and your list of themes go through and write next to each what its relationship is with the root theme.	Some examples: Conflict is related because it is 'caused by competition'. Game is related because it's a 'type of contest'. Team is related because it's 'a part of a contest's structure'. Opponent is related because they are 'competed with'.

3b	**Group Elements:** *The objective of this stage is to consolidate elements that serve the same purpose in connoting the root theme.* It is important to consolidate the wide range of elements we have now into the core components of the theme in question. Elements that represent the same or similar sets of entities within a narrative should be grouped together as a single element, if you find 2 motifs for example that cover very similar things group them together into a new motif with a more generalised name. Carefully consider elements that are components of already listed themes or are aspects or attributes of another already listed more important element and remove them by incorporating them as a part of the other theme or motif. Also be sure to consolidate elements that share the same role. This is where we use the relationships constructed in 3a. Any two elements that have the same or extremely similar relationship with the root them should be grouped together in a new theme or motif based around that relationship. Similar to above this may mean they are grouped together under the name of one of the elements being grouped rather then creating a new separate element but often it is more appropriate to make a new element with a name based on the relationship they share, choose which ever you feel is more appropriate. Ensure that what ever name used to group the elements together under is relevant to the contents of all consolidated elements.	Prize and trophy can both be generalized under 'Prize'. Uniform is an aspect of Team and can be consolidated together under 'Team'. Team, Challenge, Training, and Rules, all have the relationship 'aspect of contests structure' and as such can first be consolidated into 'Structure of Contest' and then generalised together with Contest into 'Contest' Game is 'a type of Contest' and as such is generalised with Contest.
4	**Expand Themes:** *The objective of this stage is to fill out all sub themes by expanding them.* The associated words that you have designated themes now need to be expanded in the same way the root theme was. Repeat stages 1-4 for all words designated a theme in stages 2 and 3 (this includes expanding any new themes that arise as well). This might take quite some time, in order to save time you may want to read the instructions for stage 5 now and use them in order to prune particular themes away before spending the time to fully expand them.	It would be possible to fully expand the listed themes as we have currently in this example in the same way we have for Competition thus far. Conflict for example might have themes of 'War' and motifs of 'Contest', 'Argument', and 'Fight'.

5	**Remove or Refactor Elements with Irrelevant Associations:** *The objective of this stage is to remove elements that contain features irrelevant to the root theme to prevent thematic drift.* For each element in turn we now need to check that it is 100% relevant to the root theme. It is likely that there will be elements which you have as a part of your definition that are in some way related to the root theme and may often occur along side it but are not completely relevant. For each sub-theme you have consider each of its component elements; its sub-sub-subthemes if you will, and motifs. Consider if these elements are in turn relevant to the root theme it self, if not then their parent; the sub theme they were expanded from, is considered an "associated" theme, rather then a component theme. Likewise for motifs in your model, if it is possible that they might be implied in a context that is irrelevant to your root theme then they to should be labelled associated. All associated elements must now either be refactored or removed. If the element can be changed and retitled in a way that excludes the irrelevant elements then do so, this is called refactoring. However if this is impossible without damaging the point of the element entirely then remove the element from your definition. If any elements are refactored be sure to consider steps 2-5 on the new element to ensure it is correct, you may find that a new element would need to be grouped with other elements or may change from a motif to a theme or vica versa.	Hard work, Conflict, Patriotism, and Sport all would contain motifs or themes not directly relevant to 'Competition', such as 'War' or 'Politics' or 'Money' and as such are labelled as associated. To refactor Conflict, Hard work, or Patriotism would damage the underlying themes. Sport could be refactored into a motif of 'Sporting Contest' which would be relevant but then going through the previous checks we would find it consolidated with 'Contest'.
6	**Write up finished definition:** *Turn the notes and lists for the theme into a neat finished definition.* Write a list titled with your root theme that lists each of the final remaining elements, list themes first under a sub heading "themes" then follow with motifs. After this make additional lists for any subthemes your root theme may have. Alternatively use the supplied forms to write up your final definition.	**Competition** *Themes* None *Motifs* Prize Contest Opponent

Root Theme:

Stage 1 – Associated words:

Stage 2, 3 a), and 3b) - Themes and Motifs, relationships with root theme, and Consolidation.

Allocate terms from stage 1 to either theme or motif, as per stage 2. Then list their relationship with the root theme, as per stage 3a). In stage 3b), when consolidating elements together, neatly cross out elements consolidated into a new or existing element and list newly formed elements if necessary.

Themes	*Relationship with root theme*	**Motifs**	*Relationship with root theme*

Stage 4

Sub theme expanded:

Stage 1 – Associated words:

Stage 2, 3 a), and 3b) - Themes and Motifs, relationships with root theme, and Consolidation.

Allocate terms from stage 1 to either theme or motif, as per stage 2. Then list their relationship with the root theme, as per stage 3a). In stage 3b), when consolidating elements together, neatly cross out elements consolidated into a new or existing element and list newly formed elements if necessary.

Themes	*Relationship with root theme*	**Motifs**	*Relationship with root theme*

Stage 4

Sub theme expanded:

Stage 1 – Associated words:

Stage 2, 3 a), and 3b) - Themes and Motifs, relationships with root theme, and Consolidation.

Allocate terms from stage 1 to either theme or motif, as per stage 2. Then list their relationship with the root theme, as per stage 3a). In stage 3b), when consolidating elements together, neatly cross out elements consolidated into a new or existing element and list newly formed elements if necessary.

Themes	*Relationship with root theme*	**Motifs**	*Relationship with root theme*

Stage 4

Sub theme expanded:

Stage 1 – Associated words:

Stage 2, 3 a), and 3b) - Themes and Motifs, relationships with root theme, and Consolidation.

Allocate terms from stage 1 to either theme or motif, as per stage 2. Then list their relationship with the root theme, as per stage 3a). In stage 3b), when consolidating elements together, neatly cross out elements consolidated into a new or existing element and list newly formed elements if necessary.

Themes	*Relationship with root theme*	**Motifs**	*Relationship with root theme*

Stage 4

Sub theme expanded:

Stage 1 – Associated words:

Stage 2, 3 a), and 3b) - Themes and Motifs, relationships with root theme, and Consolidation.

Allocate terms from stage 1 to either theme or motif, as per stage 2. Then list their relationship with the root theme, as per stage 3a). In stage 3b), when consolidating elements together, neatly cross out elements consolidated into a new or existing element and list newly formed elements if necessary.

Themes	*Relationship with root theme*	**Motifs**	*Relationship with root theme*

Stage 4

Sub theme expanded:

Stage 1 – Associated words:

Stage 2, 3 a), and 3b) - Themes and Motifs, relationships with root theme, and Consolidation.

Allocate terms from stage 1 to either theme or motif, as per stage 2. Then list their relationship with the root theme, as per stage 3a). In stage 3b), when consolidating elements together, neatly cross out elements consolidated into a new or existing element and list newly formed elements if necessary.

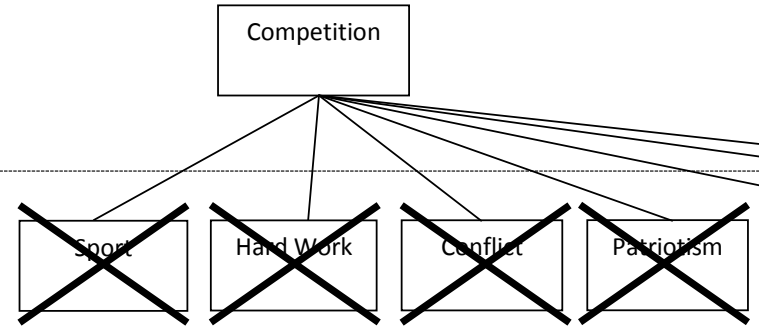
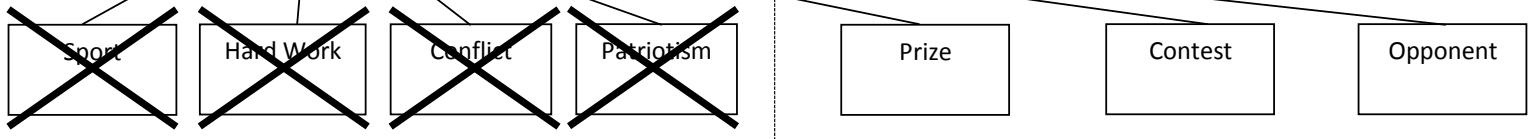
Themes	*Relationship with root theme*	**Motifs**	*Relationship with root theme*

Stage 5 and 6)

Having a completed list of themes and motifs sketch out the network they form on here, similar to the diagram presented in the introduction. Remove or refactor elements as per stage 5. On the following page the example from the guide (competition) is displayed upon such a grid.

Tier	Themes	Motifs
1		
2		
3		
4		
5		

EXAMPLE: Competition

Tier	Themes	Motifs
1		
2		
3		
4		
5		

Appendix B

Thematic Definitions Used in Experiments

In this appendix I include the thematic definitions that were used by the TMB in the experiments in chapters 4 and 5, and by the thematic illustrator for the experiment in chapter 6. The definitions for simplicity have been represented diagramatically without feature lists.

B.1 Thematic Definitions used in Experiments for Chapters 4 and 5

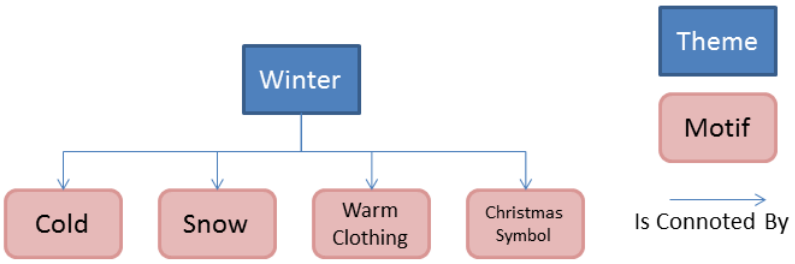


FIGURE B.1: Thematic definition for winter

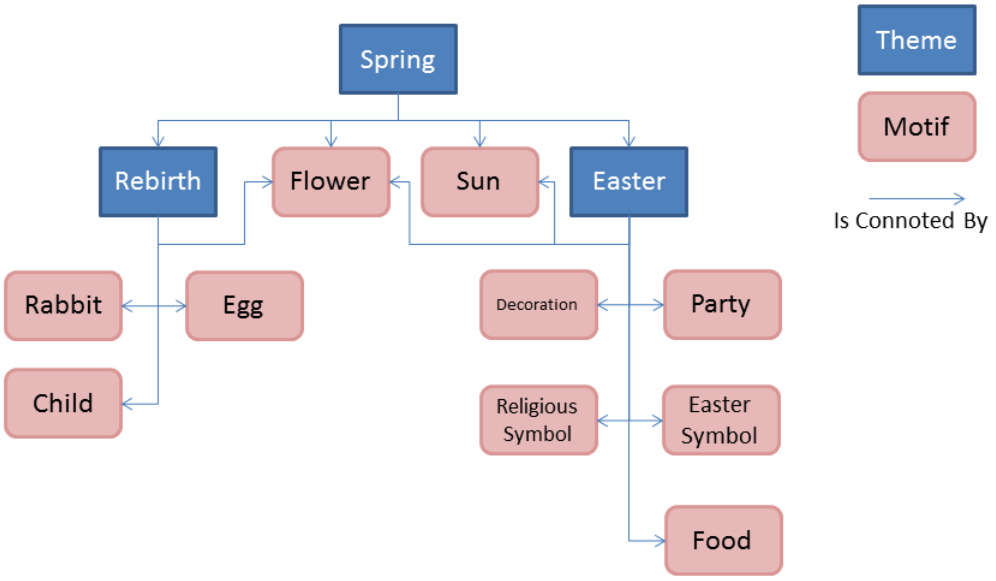


FIGURE B.2: Thematic definition for spring

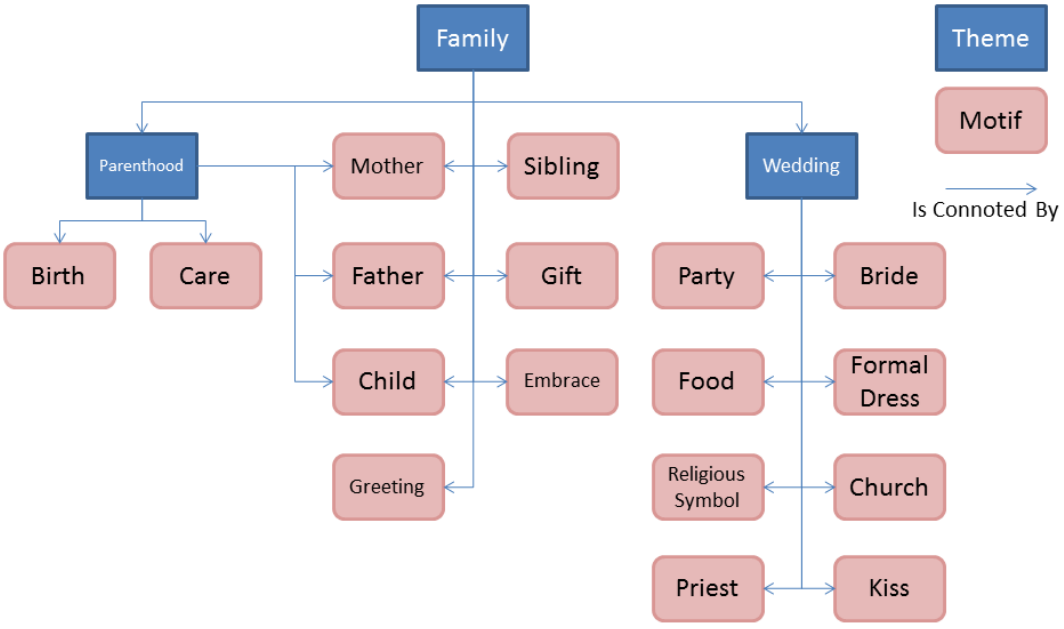


FIGURE B.3: Thematic definition for family

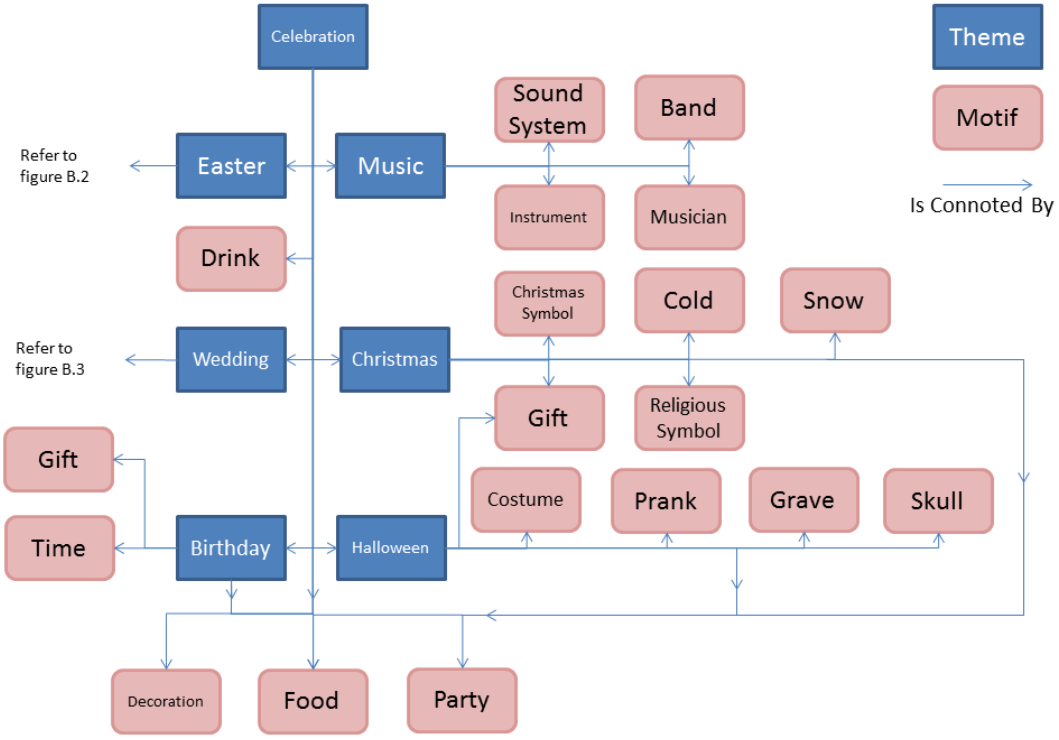


FIGURE B.4: Thematic definition for celebration

B.2 Thematic Definitions used in Experiment for Chapter 6

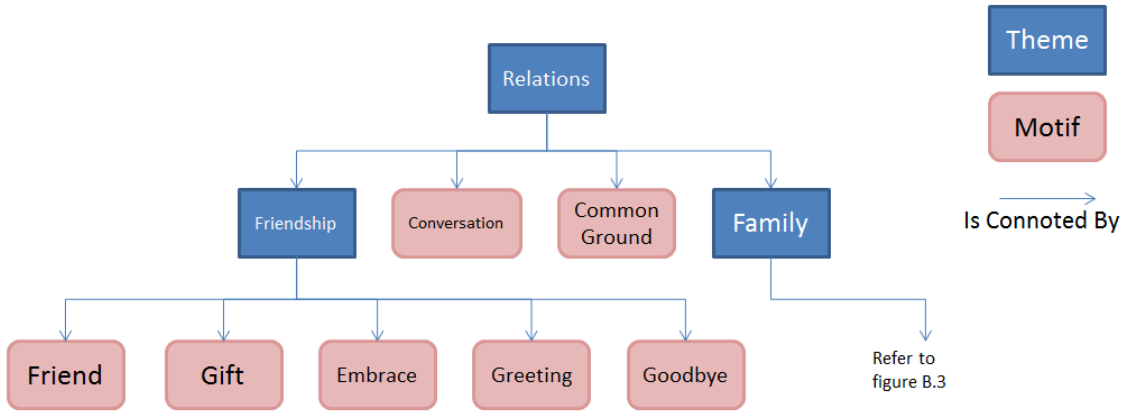


FIGURE B.5: Thematic definition for relations

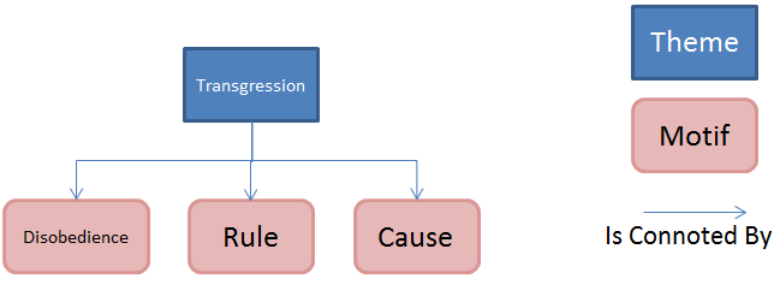


FIGURE B.6: Thematic definition for transgression

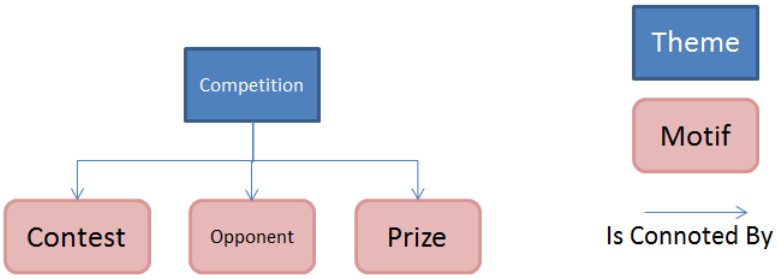


FIGURE B.7: Thematic definition for competition

Appendix C

Stories used in Cohesion Experiment for Chapter 6

The following appendix includes the short stories, in full, by Steve Ersinghaus used in the narrative cohesion experiment as detailed in chapter 6 section 6.3.

C.1 The Point

Point was a point. He had no mass, thickness, or real measurement. He was, as the definition goes, a position in space and time.

In the parlance of the day, Point or Max could be identified as existing at the intersection of two or more lines, vectors, or intentions, such as two peopleone named Jane, the other Jam.

Jane leaves the house at nine on a normal morning, 72 degrees. Jam puts her coffee into the holder, turns on the car. Its five after nine. The sun cuts her head into two halves. The bottom half of her face is submerged in shadow.

Or two nations. One of these nations drives its population in one direction, the other into many, like marbles dropped from a box onto the sidewalk. There are good days in both places when people identify falling stars. Night clouds pass over the city roofs like dramatizations of uneasy slumber. Leaves shake like hard to read words.

Jane stops at Point or Max or some other arbitrary point along the way. Jam sees her and she also sees the street light. Jam waits. Jane considers Jams hair through the window, how she would look or feel with hair like Jams. Jam smiles. She wonders what it would be like to have grown up like Jane, with that hair and those lips and that orgasmic car.

Moments later Max reaches out, touched by his implicit emergence, for he is moded by complexity, an immanent matrix. He is not height until another point appears, unwidthed until inspired by volume. A name is an intertwingle of vectors.

War is collision. One people wants this way, it wants clarity; the other is hungry for calendars and wealth and can do okay with ambiguity. Jane powers down. Jam reaches for her sunglasses and departs the car, and in the vortex of honking in the street, they meet and kiss.

C.2 The Night

A young boy no more than ten woke to the dark with the need to pee. In this world, peeing at night was not so easy a thing. His mother and father slept in the room at the far end of the hall, the bathroom in between, two picture windows that looked out onto the lawn and the lawns of the other houses and a park beyond. His sister, Betty, slept in the bed beside him.

He got out of bed and paused with amazement at the door to the hall: Mother or Father had forgotten to shutter the windows at dusk. He should, he told himself, go back to bed but the pee threatened to explode. He thought about asking Betty. But something in him said no and to look. To test what always had been said and warned against. The windows must remain shut up when the sun goes down. The doors must be locked at night and every night, no exceptions.

Hed seen photos. They were only photos: stars, the moon, eclipses. Ill lock you out at night, Betty would threaten. Stop, Father would say. It isnt funny.

But could he just look? Would he? A quick look. Just take a quick glance, then pee, then back to bed, and shuffle through the whys in the morning when it was safe to do so, safe to ask questions, or safe to say I risked and it wasnt true. None of it has been true, Father.

Last month he and Betty and snuck into the hallway and sat beneath the shutters, their hearts patting in their chests like crows wings. No, the boy had said finally. No way.

We just shouldnt, Betty whispered. We just cant.

So, why not just look now? He watched down the hall. The moon shown bright through the big windows. It was that Blue Time. Four or so. Hed seen the blue light only once before, early morning high in the mountains, where everyone sat on an enormous balcony, the hotel rising behind spatter-lit yellow from the random windows in the higher rooms. It had been so rare, Mother standing by Father at a stone lookout. They had drinks. Even together then they looked lonely and helpless and somehow guilty of a quiet collaboration in solitude and failure.

That time had been about this same time, so late and early, and everyone happy, a wakeful holiday, refusing in unison to sleep. He remembered how his mothers head had slowly fallen to Fathers shoulder, his head to her head. So high they were, the clouds swirled below silver and humped, high and safe above a world where night never really came.

He felt the energy or thrill of impulse. He felt drawn from the door, drawn to the window. He moved his eye to the left of the frame, looked out for the first time at an alien night and encountered moon glow, the shadows of the nearer trees soft and silent on the bright white ground. Nothing, he thought. There was nothing, nothing to fear out there, all the rules untrue or just false.

The windows on the block were shuttered, blank spots against the gray sky. In the park, it must have been very far, he saw a dark shape drift rapidly out from the pines into a clear space on the grass, something that suggested spider but with only two or three legs, small so far away, then it was gone, like something painted on the ground by a breeze. He felt a pressure in his chest, something of an illicit sweetness at the bottom of his throat.

They sense vision, he remembered hearing. They ignore us if we dont see . . . Grandfather said. They had been in Wisconsin with relatives. They had been eating in a windowless cafeteria. Grandfather told stories. He explained, while Mother and Father ate, ready to silence the speaker if he strayed. . . . or look. I dont know when precisely they came, but it was just after your mother was born, and we had to change. All of us, even governments. We lost the night everywhere to their teeth, their nails. But were safe if we dont look. Safe if we shut the night out. We no longer have the moon to ourselves.

Theyre drawn to the human eye. They only take he or she who sees them. And so we board our windows. Its safest that way. Sleep through the night. Keep the windows shut. Love the daytime.

Father, Father said. Please.

The boy watched the night. Something dark slashed across a nearer street. Then the boy heard his Father.

No, Father said. He was standing at the bedroom door. My god, son, he said.

Father, the boy whispered, his voice shaking. I just had to pee.

Oh my god, the Father said. Oh my god.

C.3 Computer Leon

They called Computer Leon. Soon, wire and card lawn signs went up saying We called Computer Leon.

If it werent for Computer Leon, Id have nothing but a thousand dollar paper weight on my desk, a man said.

Computer Leons a real pro. He had my PC up and running in little more than an hour. He didnt even break a sweat.

Hes like the dog whisperer but with computers.

More signs went up. We called Computer Leon, the lawn signs said. Even people who didnt call Computer Leon put up signs that said We called Computer Leon because they wanted everyone to believe that they were the kind of people who would call Computer Leon, if they had computers. Computer Leon got busy with so many calls he had the income to make custom signs, such as I called Computer Leon and Computer Leon for President.

Sure, hes a bit odd. But he gets the job done.

I got a virus and I called Computer Leon. He came right over. He said he was on his lunch break but he came anyway. He found the problem and sent that virus packing to hell where it fucking belongs.

Computer Leon installed my new operating system. While we waited, I served him cookies. I didnt know he was allergic to nuts though. Sure, his tongue got all swoled up and everything but he stayed till he had my machine purring like a new engine. Computer Leons the best.

Computer Leon drove the neighborhoods computing the number of signs. Im on the constant go, he told me. I dont have time to worry about politics and the worlds problems. Im constantly on the go. You wouldnt believe the things people do to their computers. The problems they get themselves into.

For instance, I said.

Like desktop clutter or using their CD trays as cup holders. But theyre all nice people. They show me their houses. Plans for additions. I know some of those houses better than the owners do. Some Ive had to rewire and when you rewire you learn things.

It was a Saturday when Computer Leon was on calls and he saw a sign on the lawn of a big blue Colonial that said, We called Computer Geek Woman.

He read the sign as an accusatory blow. Computer Leon asked the man at the door. The man said, Because I called you and I got a message. I called and she came right

over. Sorry.

Competitions good for the soul, his wife said.

The hell you say, Computer Leon said, hatching schemes.

More signs that said We called Computer Geek Woman went up. Soon he saw We called Computer Geek Woman signs next door to his own. Happy green clover decorated the borders. The bold brown script gave Computer Leon chills.

He called her. My computer. I think its the hard drive.

Tell me exactly what happened, a woman said, cordial, competent. He said, You cant guess? I thought you were the great Computer Geek Woman.

Sorry, she said. But if you tell me what happened Id be glad to help as best I can.

He told a story. He tried to sound convincingly naive. Then the thingy started to go bloop bloop and it just seized up he said. She sped right over. He brought her downstairs. Hed rigged his virtual memory as a trap. She caught the problem straight away and reset to default.

Shes good, he told his wife.

Leave her alone, Leon. Shes just trying to make a living. Its hard times out there.

He knew peoples schedules. He understood the neighborhoods pulse. He slipped in through a sliding door and went deep into a mans PC, and on his drive away he grinned at the We called Computer Geek Woman sign and said, Fix that.

In between calls he fouled two more Geek Woman clients. On one he broke open the case and took a needle to the ATX pins.

These power connector pins are mutilated, he told the woman when she called him. No problem. When he drove up the We called Computer Geek Woman sign had been removed and a fresh We called Computer Leon sign had been tamped in its place.

Computer Leons my man when it comes to equipment, the woman told me. I run a home business. My life depends on my computer. I should never have trusted Computer Geek Woman.

A few days later, nearby the Catholic Church, he saw a sign that said, Computer Leon is a Fraud and another that said Its Computer Leon: Lock Up Your Goat!

On his answering machine: This message is for you, Computer Leon. Two can play this game.

What the hells this about, Leon? his wife wanted to know. She had a wooden spoon in her hand. She slapped its bowl hard into her palm. Who is that? Whatever it is youre up to, quit it.

The hell you say, Computer Leon said.

Whats this about, Leon? asked Cruz, a man with a bank of Mac Pros in his office.

Its nothing, Computer Leon swore, opening a case of CDs. I dont know anyone with goats.

In his driveway, someone had put a wire-frame sign that said Computer Leon called Computer Geek Woman. He stuffed it in the trash.

He pulled signs up from the church, library, and bank lawn at midnight. His truck bed was soon full of signs that said Computer Leon called Computer Geek Woman. He dreamed about illicit signs. They checkered the landscape. He pushed a sign up hill over and over again that said Computer Leon is Losing His Mind and up hill still because the hill went up hill still.

He grew dreary with sleep loss; he swore while brushing his teeth. The skin under his left eye began to sag. But the calls kept him running: Im getting that fatal error message, Leon. My CD drive wont read disks, Leon. My computer wont turn on, Leon. It keeps freezing, Leon. What was that registry thingy again, Leon? What does a question mark on gray folder mean at bootup, Leon? My computer keeps telling me Im an asshole, Leon. I spilled apple cider on my keyboard, Leon. I got that virus again, Leon. How do I start up a Facebook account, Leon. Are those commercials really true, Leon? Whats the matter with your eye, Leon. You look like crap, Leon. Well, Leon, if thats the way youre going to be, Ill just call Computer Geek Woman.

In the morning, Computer Leon practically had to throw himself behind the wheel. His hands were raw from pulling signs. He bore troubled dreams on his back like a load of hard drives.

On a lawn at the end of the block was a sign that said, We called Dan the Computer Man. Happy lady bugs danced across the bottom of the sign. The bold blue typography gave him chills.

The woman who opened the door said, You were busy; she was busy. But this Dan fellow. Well, I called, and he came right over.

Appendix D

Experiment Results Image Selections

Within this appendix I include the images selected for montages by different methods for the TMB experiments in chapters 4 and 5, as well as the images used in the illustration experiment in chapter 6.

D.1 TMB Photo Montage Experiment

Images selected as part of the TMB and Flickr comparisson experiment in chapter 4.

D.1.1 London in Winter

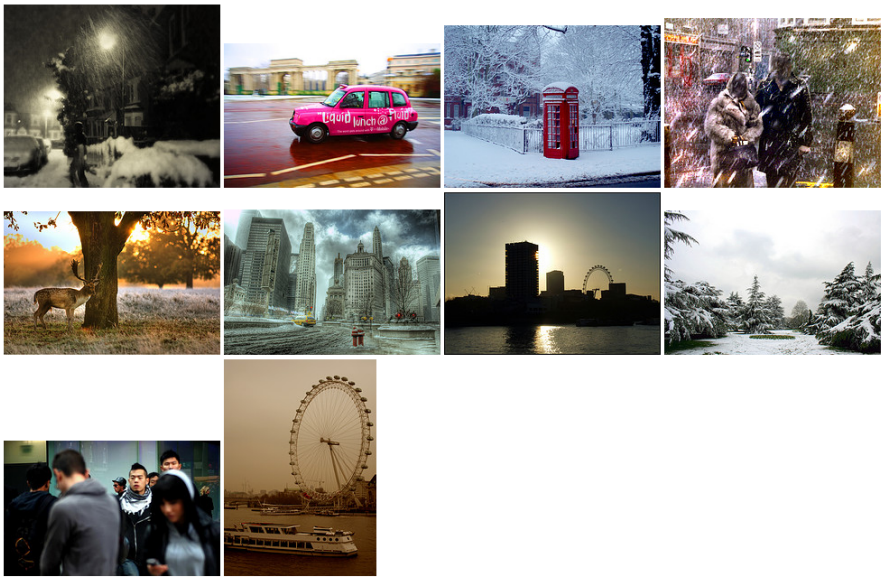


FIGURE D.1: TMB: selection for London in Winter as part of chapter 4 experiment

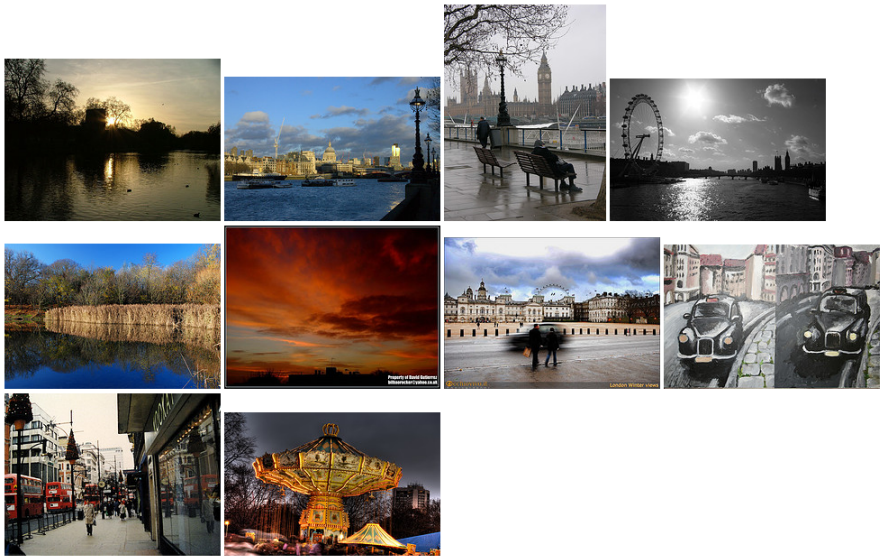


FIGURE D.2: Flickr: selection for London in Winter as part of chapter 4 experiment

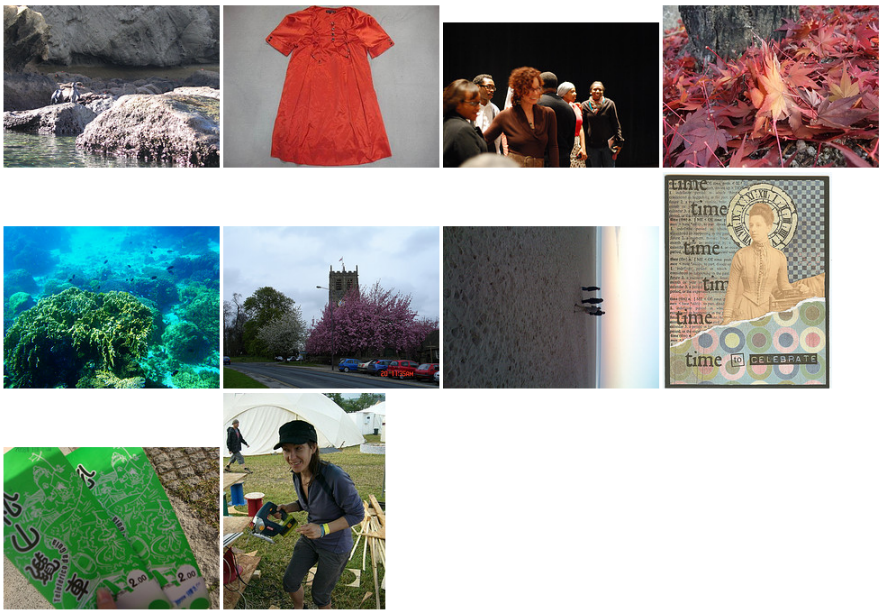


FIGURE D.3: BaseL: selection for London in Winter as part of chapter 4 experiment

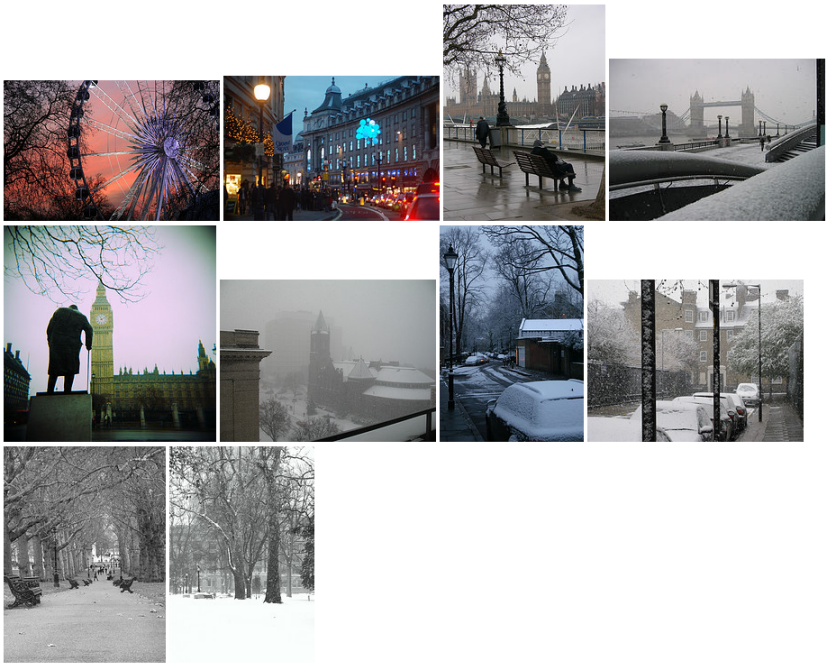


FIGURE D.4: BaseH: selection for London in Winter as part of chapter 4 experiment

D.1.2 Earthquake and Celebration

FIGURE D.5: TMB: selection for Earthquake and Celebration as part of chapter 4 experiment

FIGURE D.6: Flickr: selection for Earthquake and Celebration as part of chapter 4 experiment

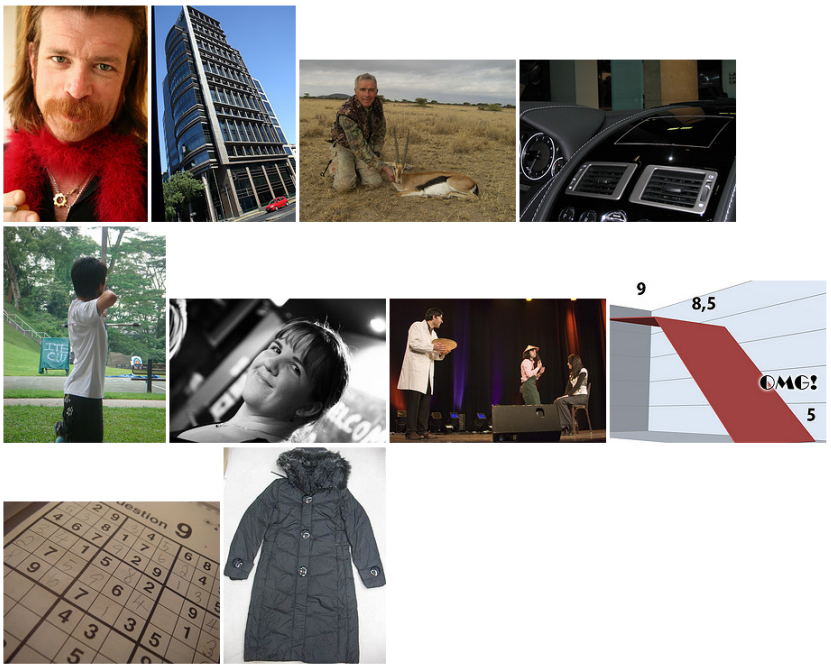


FIGURE D.7: BaseL: selection for Earthquake and Celebration as part of chapter 4 experiment

FIGURE D.8: BaseH: selection for Earthquake and Celebration as part of chapter 4 experiment

D.1.3 Family Factory

FIGURE D.9: TMB: selection for Family Factory as part of chapter 4 experiment

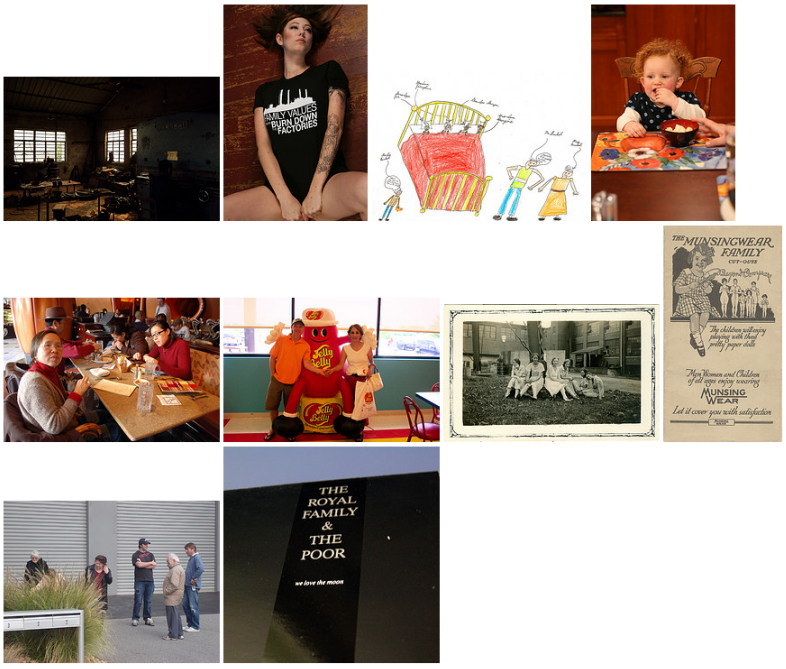


FIGURE D.10: Flickr: selection for Family Factory as part of chapter 4 experiment

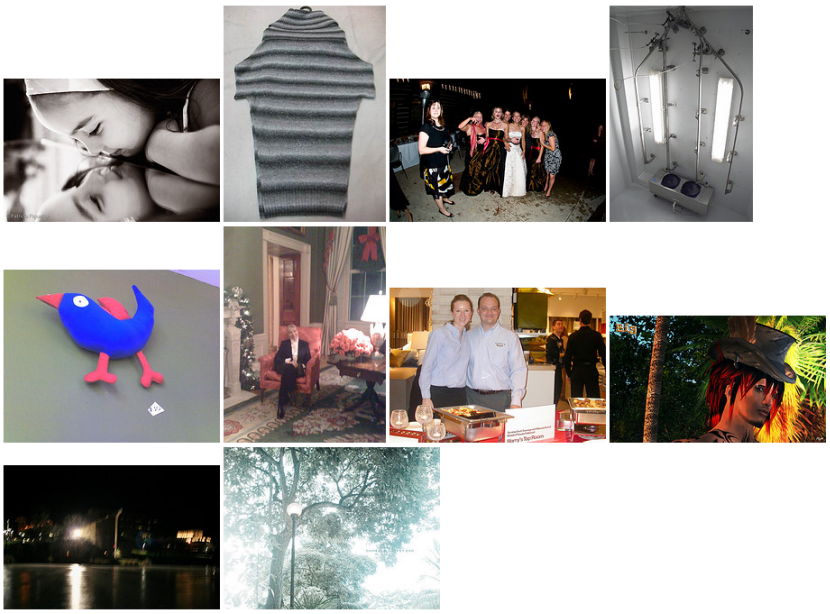


FIGURE D.11: BaseL: selection for Family Factory as part of chapter 4 experiment

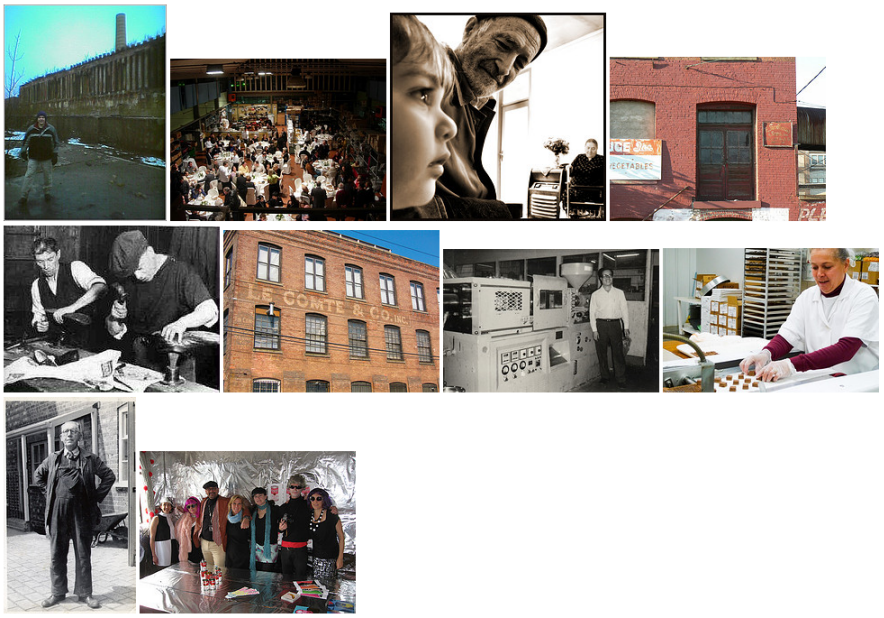


FIGURE D.12: BaseH: selection for Family Factory as part of chapter 4 experiment

D.1.4 Spring Picnic

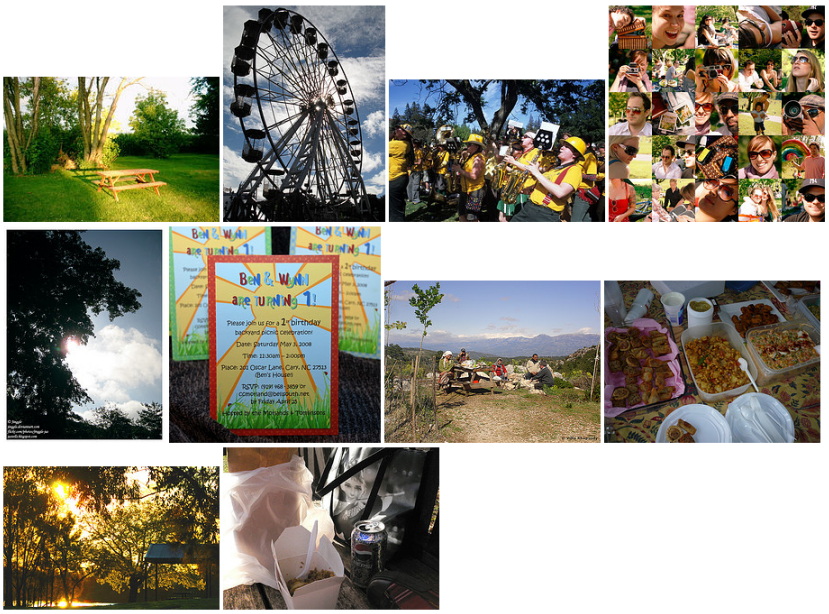


FIGURE D.13: TMB: selection for Spring Picnic as part of chapter 4 experiment

FIGURE D.14: Flickr: selection for Spring Picnic as part of chapter 4 experiment

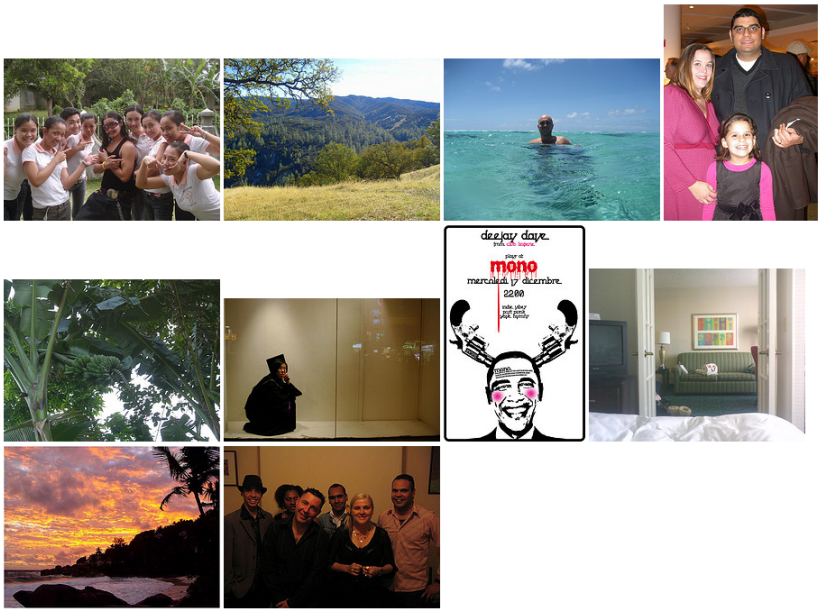


FIGURE D.15: BaseL: selection for Spring Picnic as part of chapter 4 experiment

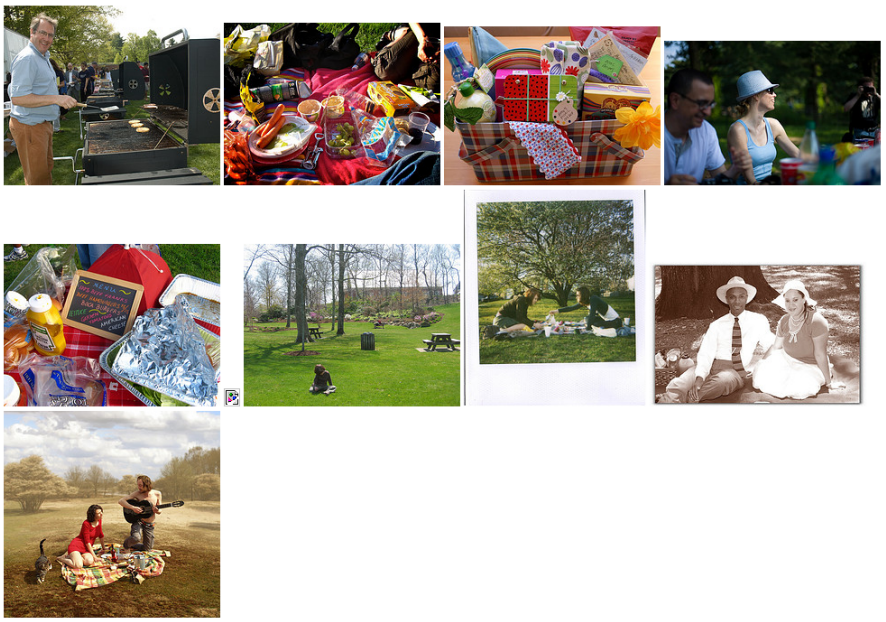


FIGURE D.16: BaseH: selection for Spring Picnic as part of chapter 4 experiment

D.1.5 Family in New York at Winter

FIGURE D.17: TMB: selection for Family in New York at Winter as part of chapter 4 experiment

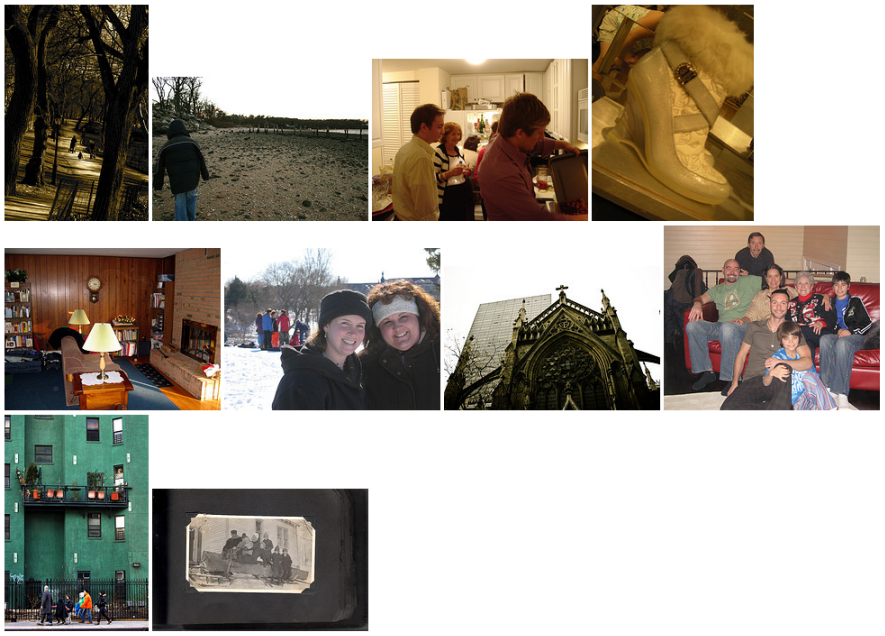


FIGURE D.18: Flickr: selection for Family in New York at Winter as part of chapter 4 experiment

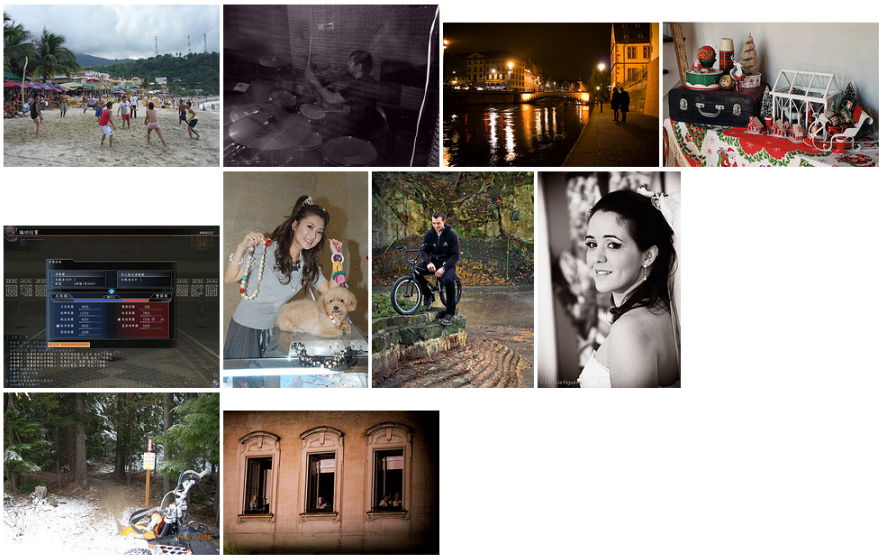


FIGURE D.19: BaseL: selection for Family in New York at Winter as part of chapter 4 experiment

FIGURE D.20: BaseH: selection for Family in New York at Winter as part of chapter 4 experiment

D.1.6 Celebration of New House in Spring

FIGURE D.21: TMB: selection for Celebration of New House in Spring as part of chapter 4 experiment

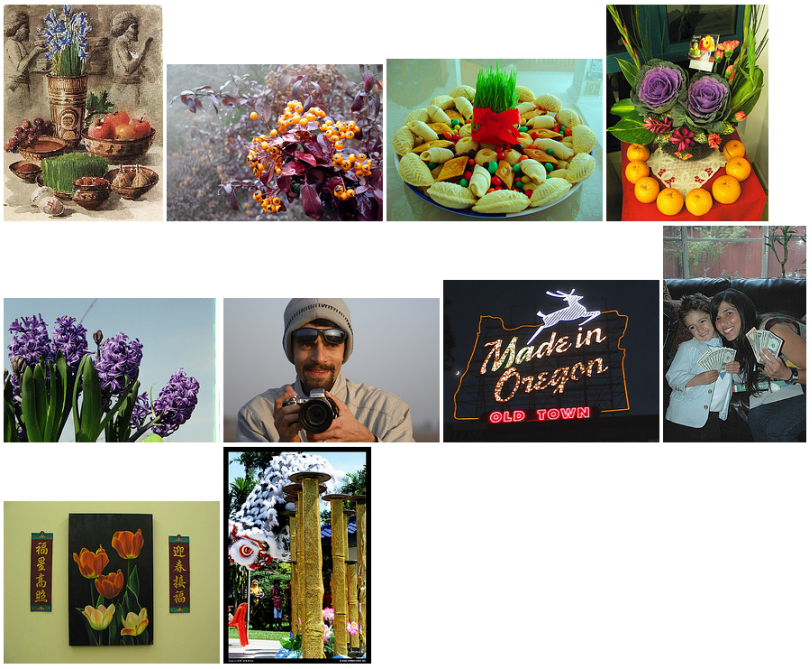


FIGURE D.22: Flickr: selection for Celebration of New House in Spring as part of chapter 4 experiment

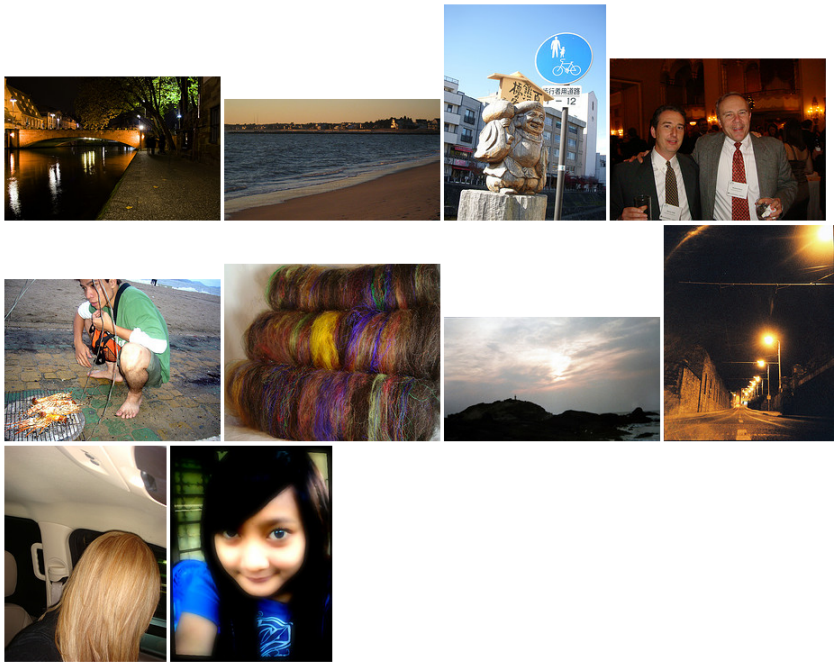


FIGURE D.23: BaseL: selection for Celebration of New House in Spring as part of chapter 4 experiment

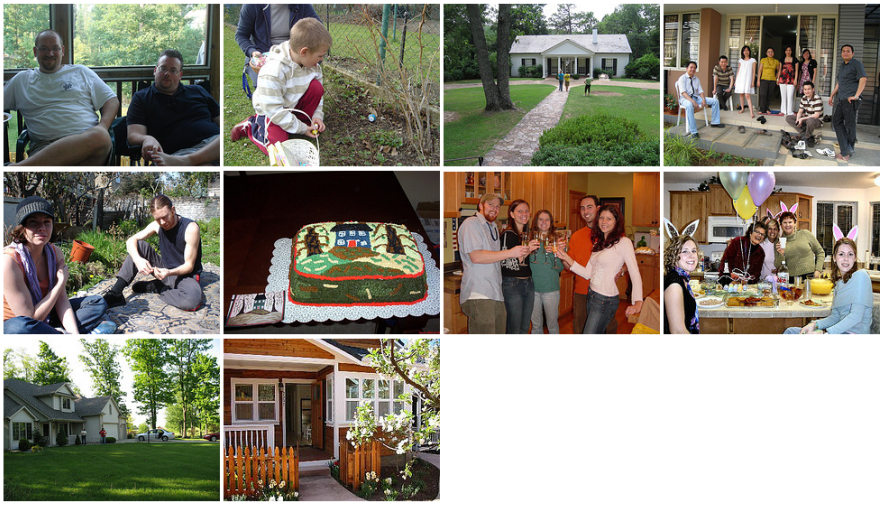


FIGURE D.24: BaseH: selection for Celebration of New House in Spring as part of chapter 4 experiment

D.2 TMB and Co-occurence Experiment

Images selected as part of the TMB and Co-occurence comparisson experiment in chapter 5.

D.2.1 London in Winter

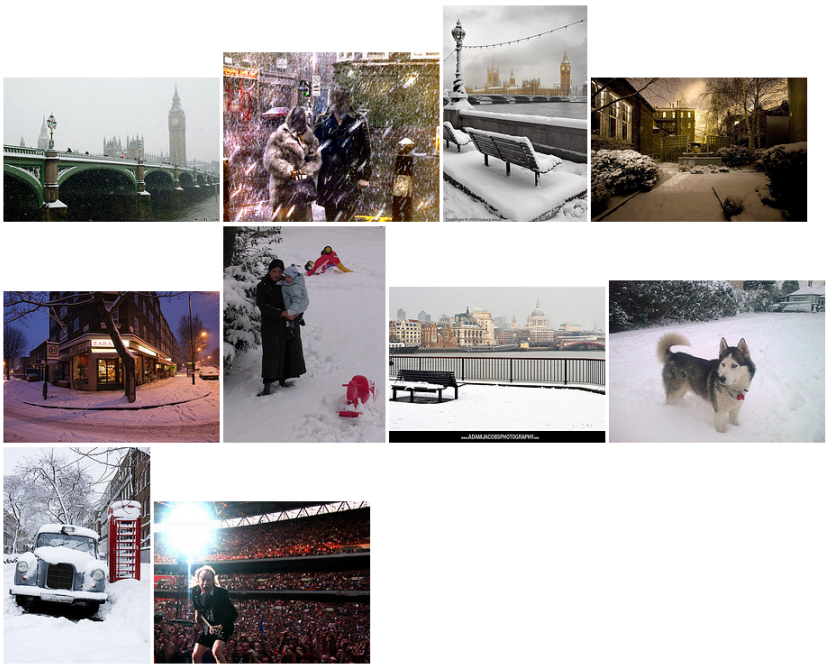


FIGURE D.25: TMB: selection for London in Winter as part of chapter 5 experiment

FIGURE D.26: Co-occurence: selection for London in Winter as part of chapter 5 experiment

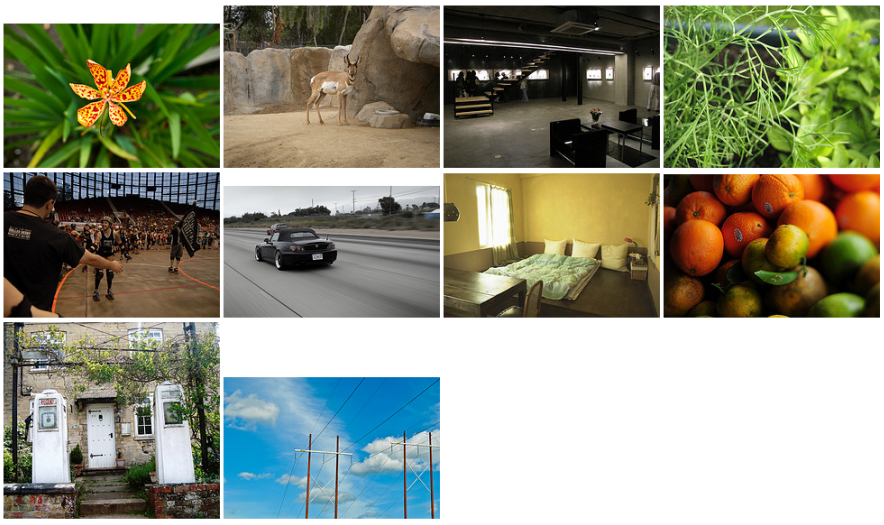


FIGURE D.27: BaseL: selection for London in Winter as part of chapter 5 experiment

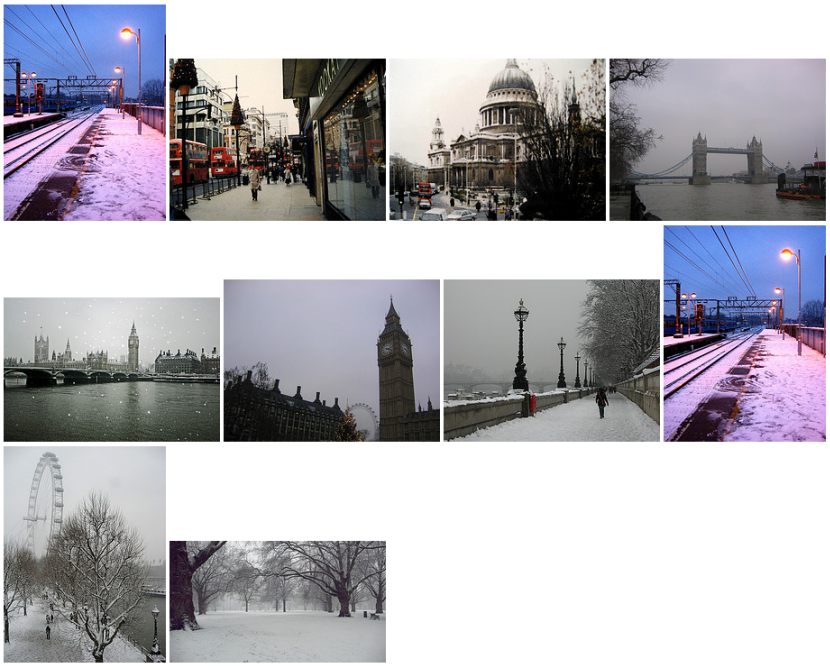


FIGURE D.28: BaseH: selection for London in Winter as part of chapter 5 experiment

D.2.2 Earthquake and Celebration

FIGURE D.29: TMB: selection for Earthquake and Celebration as part of chapter 5 experiment

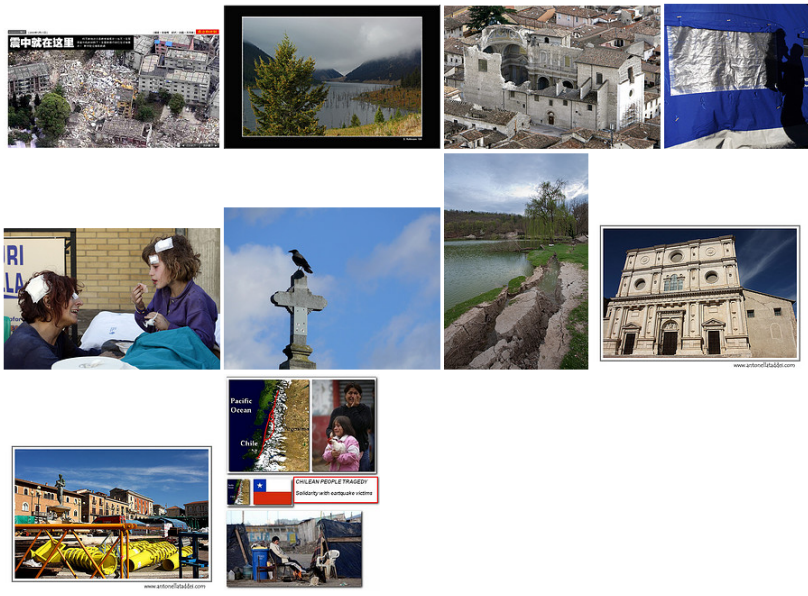


FIGURE D.30: Co-occurence: selection for Earthquake and Celebration as part of chapter 5 experiment

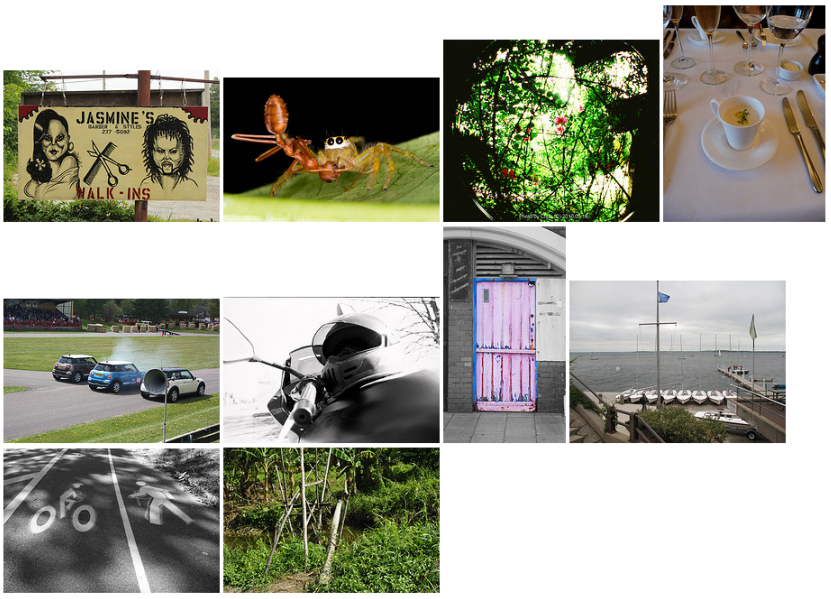


FIGURE D.31: BaseL: selection for Earthquake and Celebration as part of chapter 5 experiment

FIGURE D.32: BaseH: selection for Earthquake and Celebration as part of chapter 5 experiment

D.2.3 Family Factory

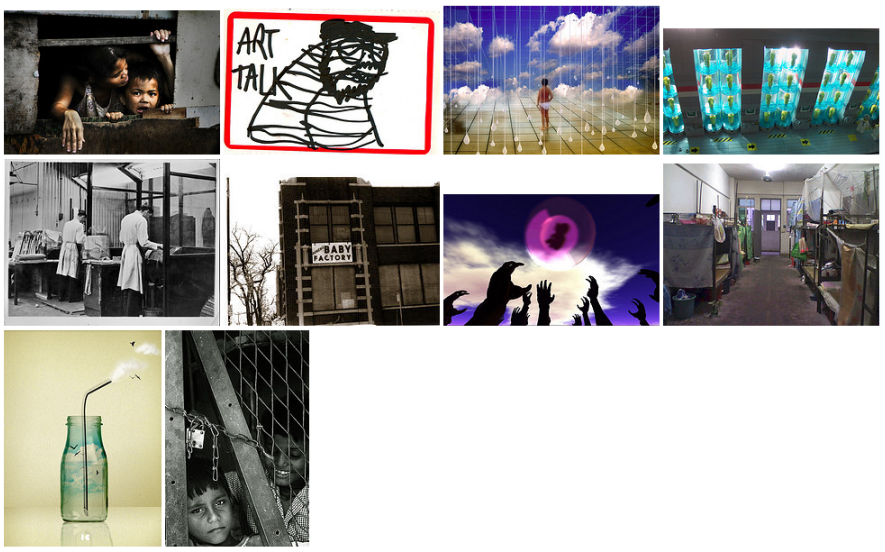


FIGURE D.33: TMB: selection for Family Factory as part of chapter 5 experiment

FIGURE D.34: Co-occurence: selection for Family Factory as part of chapter 5 experiment

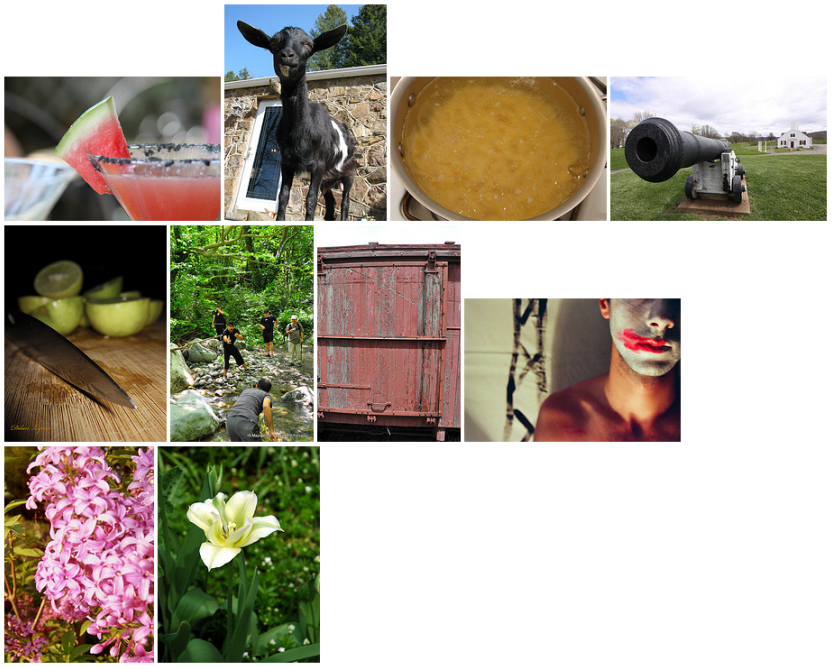


FIGURE D.35: BaseL: selection for Family Factory as part of chapter 5 experiment

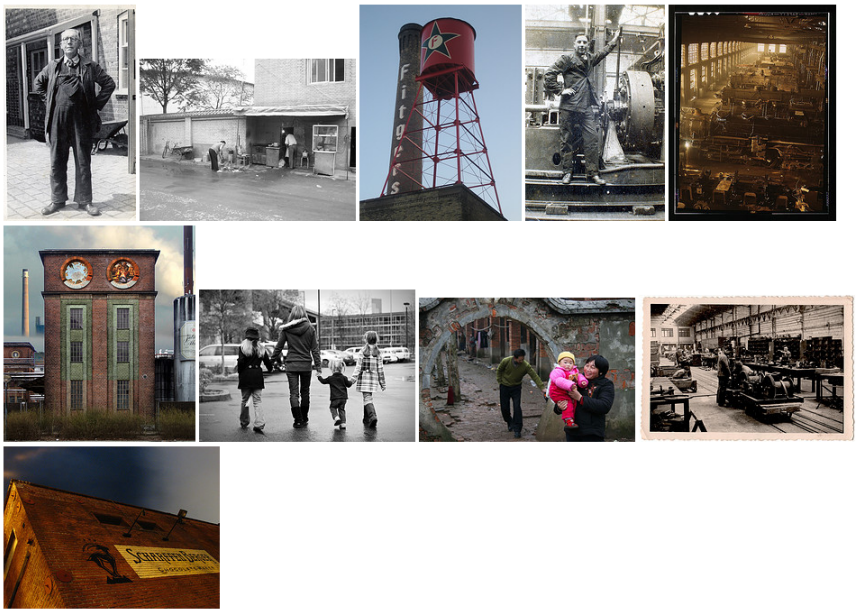


FIGURE D.36: BaseH: selection for Family Factory as part of chapter 5 experiment

D.2.4 Spring Picnic

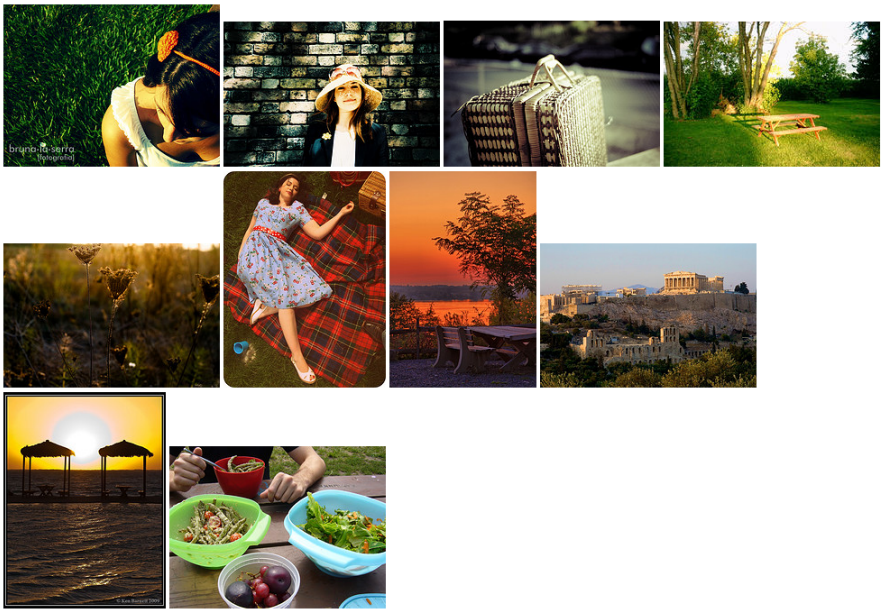


FIGURE D.37: TMB: selection for Spring Picnic as part of chapter 5 experiment

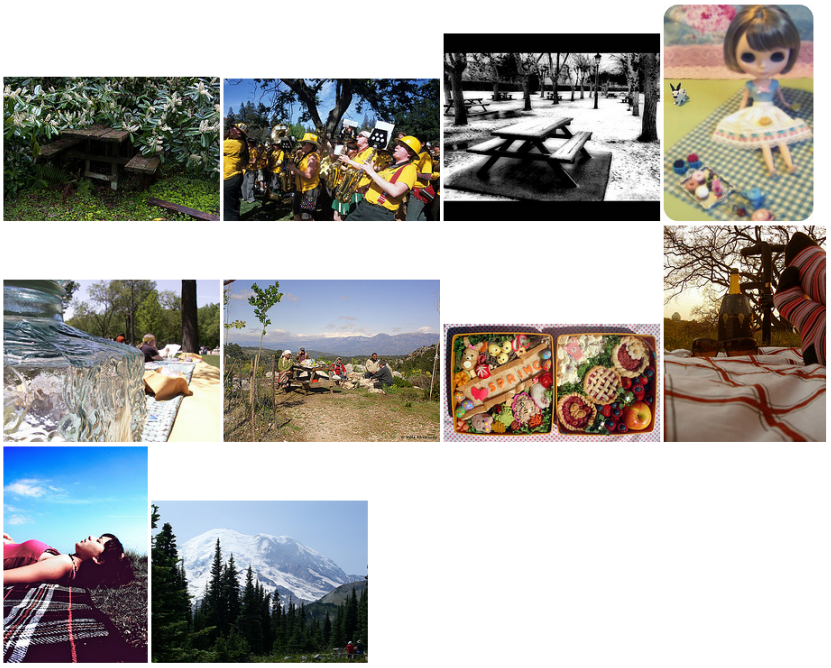


FIGURE D.38: Co-occurence: selection for Spring Picnic as part of chapter 5 experiment

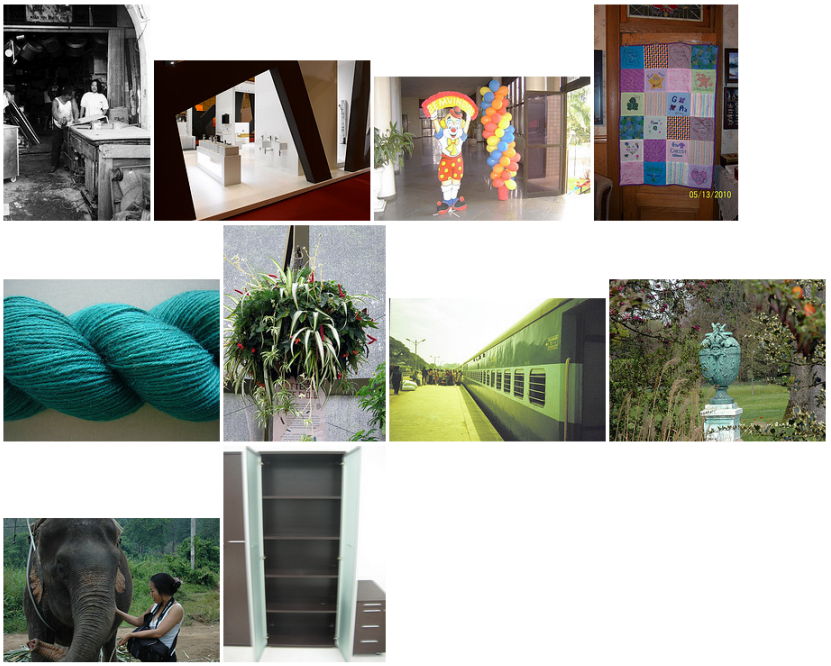


FIGURE D.39: BaseL: selection for Spring Picnic as part of chapter 5 experiment

FIGURE D.40: BaseH: selection for Spring Picnic as part of chapter 5 experiment

D.2.5 Family in New York at Winter

FIGURE D.41: TMB: selection for Family in New York at Winter as part of chapter 5 experiment

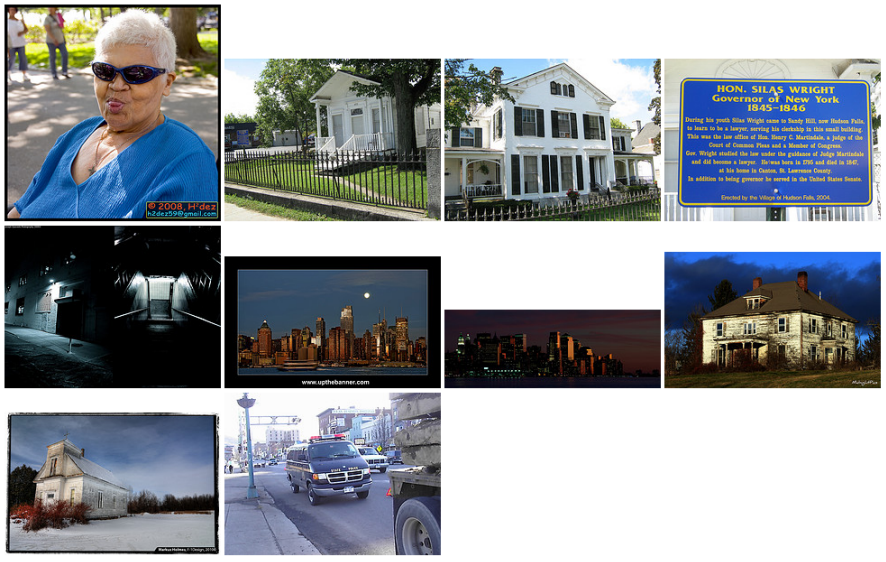


FIGURE D.42: Co-occurence: selection for Family in New York at Winter as part of chapter 5 experiment

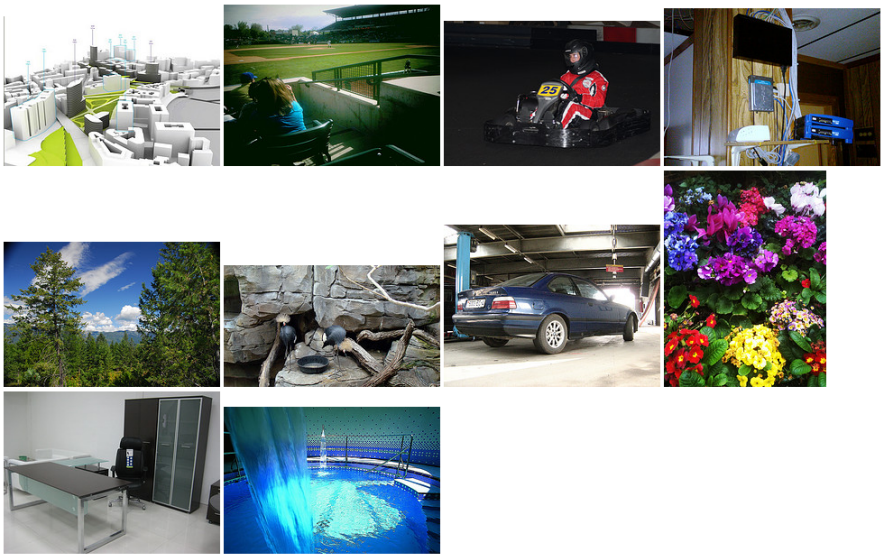


FIGURE D.43: BaseL: selection for Family in New York at Winter as part of chapter 5 experiment

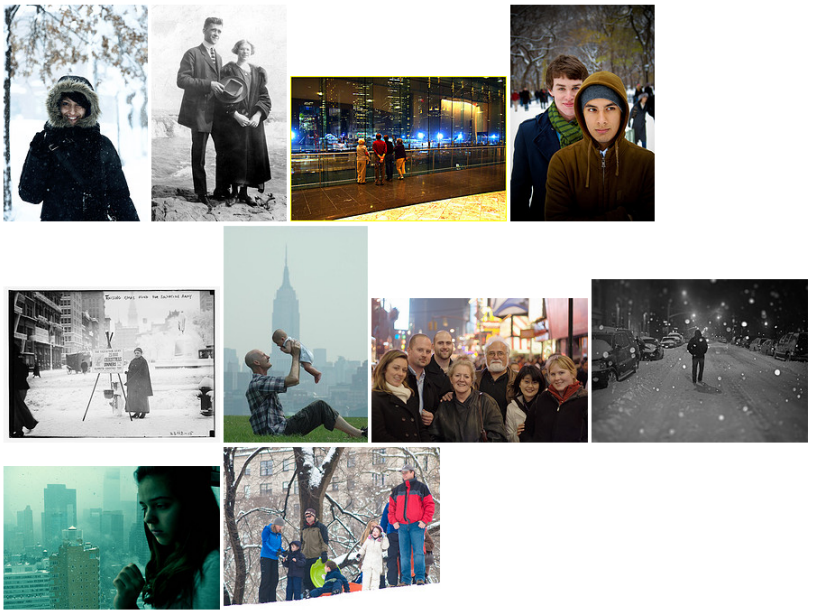


FIGURE D.44: BaseH: selection for Family in New York at Winter as part of chapter 5 experiment

D.2.6 Celebration of New House in Spring

FIGURE D.45: TMB: selection for Celebration of New House in Spring as part of chapter 5 experiment

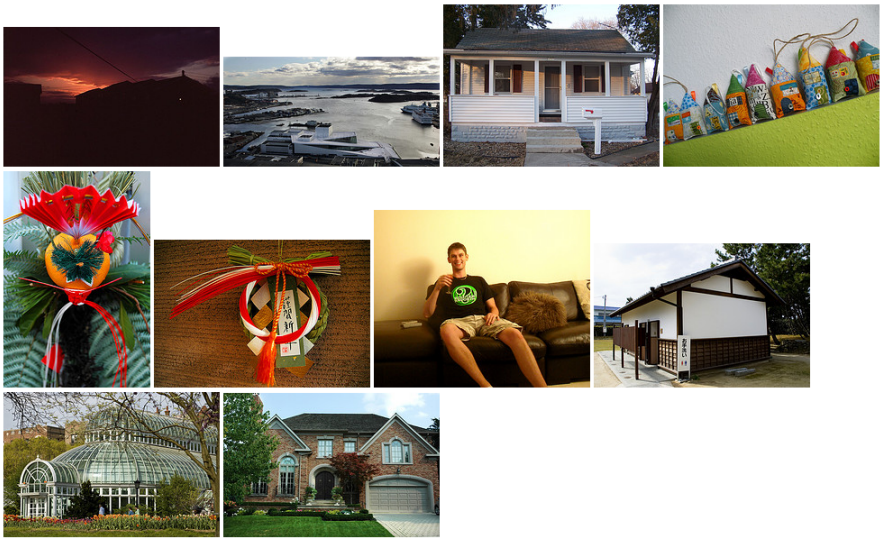


FIGURE D.46: Co-occurence: selection for Celebration of New House in Spring as part of chapter 5 experiment

FIGURE D.47: BaseL: selection for Celebration of New House in Spring as part of chapter 5 experiment

FIGURE D.48: BaseH: selection for Celebration of New House in Spring as part of chapter 5 experiment

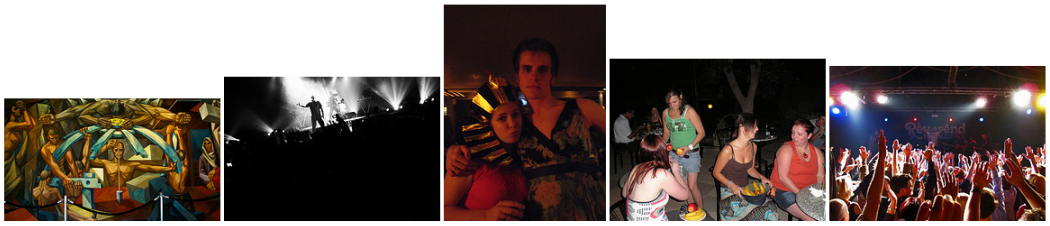
D.3 Narrative Cohesion Experiment

Images selected as part of the narrative cohesion experiment in chapter 6. In some cases images used in this experiment have subsequently been removed by Flickr, in these cases the image is not present in this appendix.

D.3.1 Story: The Point

Section 1

Section 2



Section 3

FIGURE D.49: Content and Theme: illustration selections for The Point

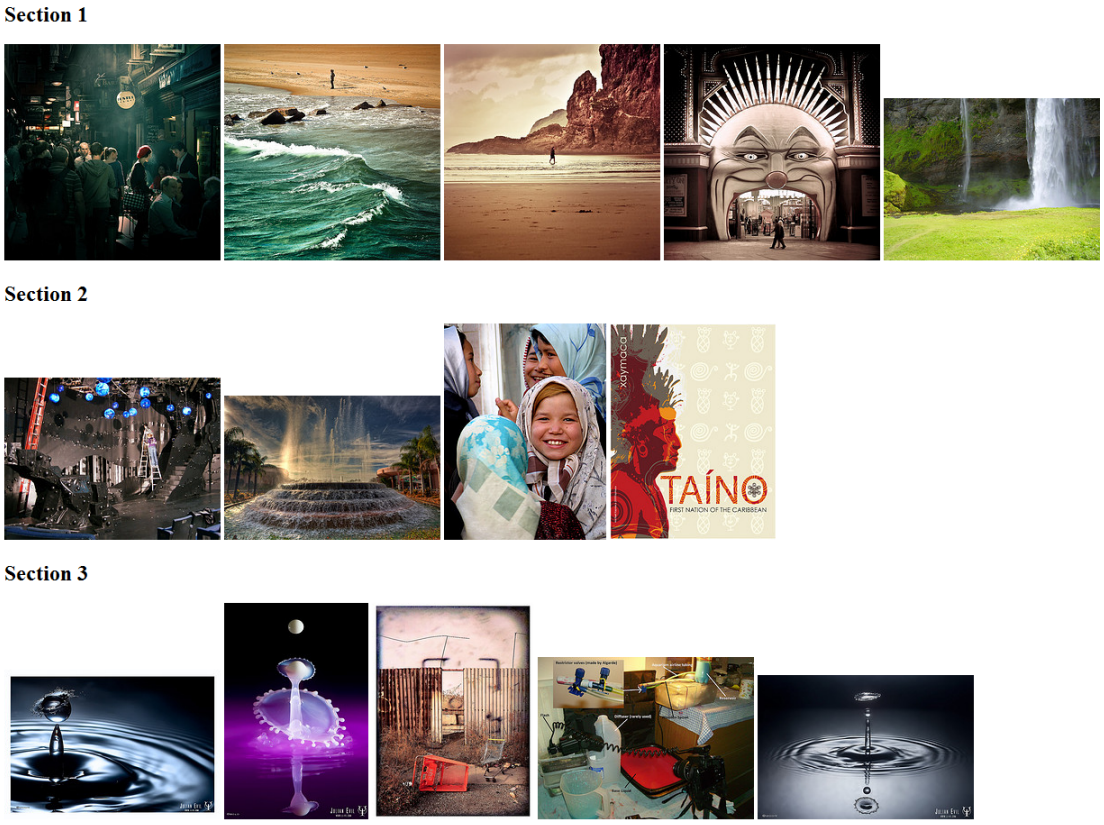


FIGURE D.50: Content only: illustration selections for The Point

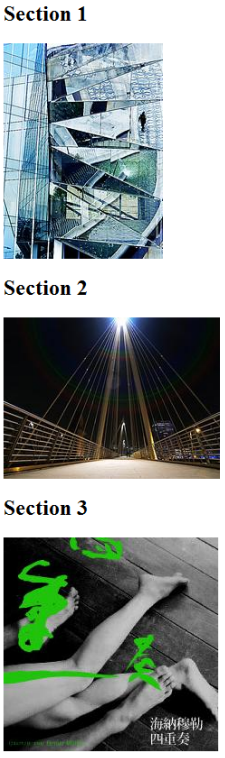


FIGURE D.51: Human selected: illustration selections for The Point

D.3.2 Story: The Night

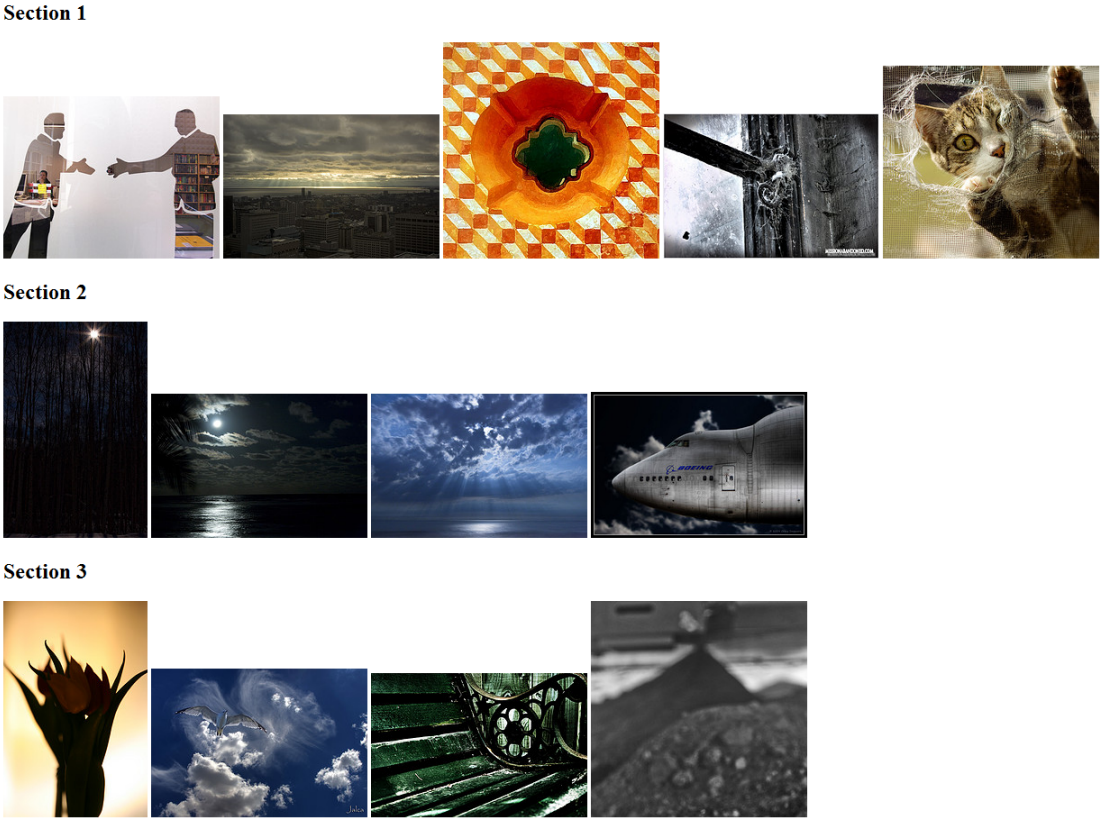


FIGURE D.52: Content and Theme: illustration selections for The Night

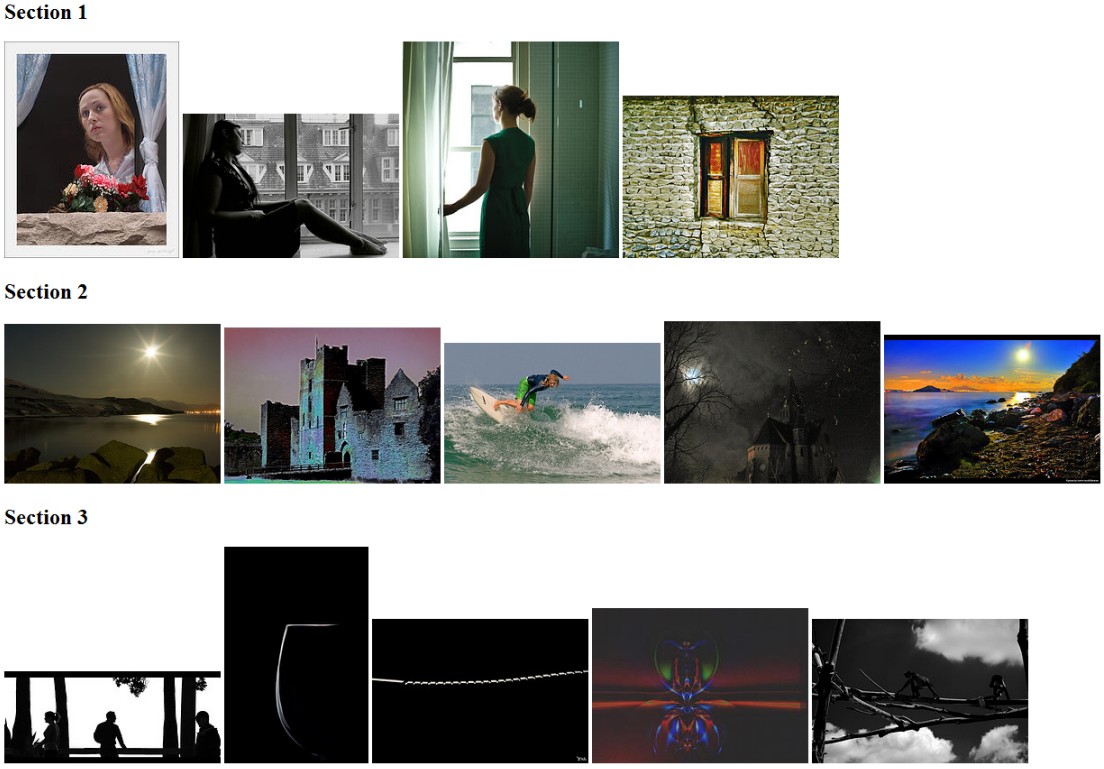


FIGURE D.53: Content only: illustration selections for The Night

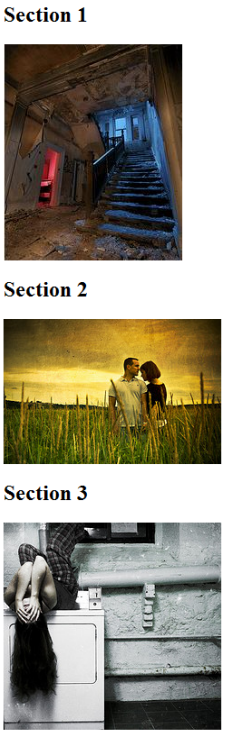


FIGURE D.54: Human selected: illustration selections for The Night

D.3.3 Story: Computer Leon

FIGURE D.55: Content and Theme: illustration selections for Computer Leon

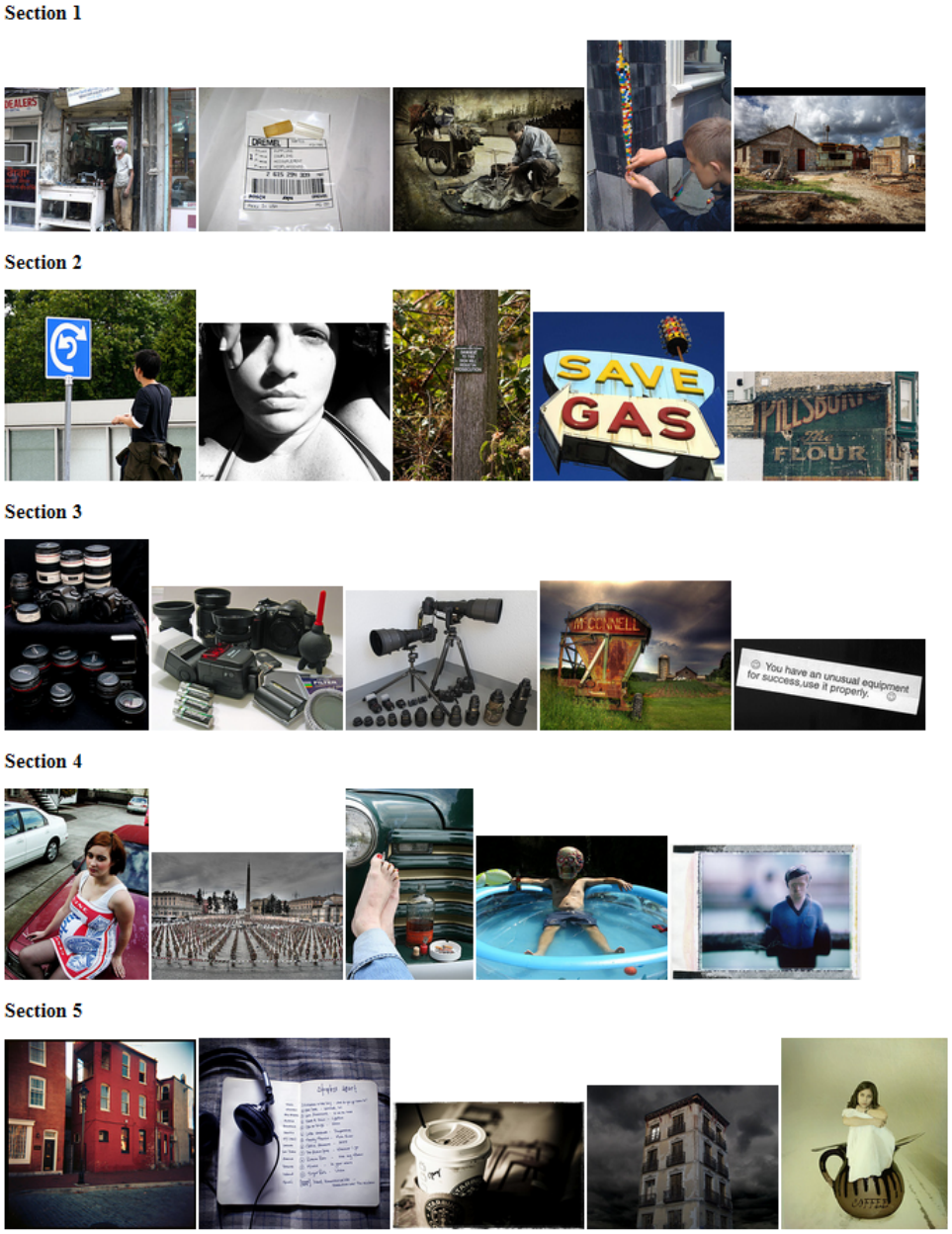


FIGURE D.56: Content only: illustration selections for Computer Leon

Section 1

Section 2

Section 3

Section 4

Section 5

FIGURE D.57: Human selected: illustration selections for Computer Leon

Appendix E

Ethics Approval Granted For Experimentation

Experiments at conducted at the university of Southampton which include human participants are required to seek approval from an internal ethics committee before being carried out. This approval is granted based on submitted details of methodology and relevant materials. I include in this appendix the methodology submissions to the ethics committee and for reference the code under which each was approved.

E.1 Thematic Definition Authoring Experiment

Ethics submission form thematic definition authoring experiment detailed in Chapter 3 Section 3.2.

E.1.1 Methodology

I am investigating the benefit of the use of a thematic model to capture and use semiotic understanding of terms to enrich narrative systems and other presentations of information. One of the draw backs to the use of my thematic model is that themes must be defined for it by hand in the terms of the model. So that anyone might define themes I have developing a guide which can be used to follow a set of rules to author a valid instance of the model for any given theme. In order to explore whether anyone could use such a guide (not just computer scientists) to generate an instance of the model. I am looking to run an experiment where 30 participants use the guide to write an instance for one of 5 themes (Danger, Winter, Celebration, Spring, or Family ... 6 participants per theme, distributed randomly) using either supplied forms or note paper. The participants will hopefully be volunteers from the English department but should there be

a shortfall the numbers will be made up of other non computer scientist volunteers. One a list of volunteers has been established I will book 3 rooms on different times and places where I will run 3 focus sessions, the volunteers will be invited to attend any one of these sessions (their choice) where we will conduct the experiment. Early tests have shown the authoring process using the guide to take approximately an hour but I shall book the rooms for 2 hours to allow for late comers. On completion of the experiment the participants will be rewarded with a 10 Amazon gift voucher soon afterwards. No personal data will be collected, only the completed instances of the model from the participants as a result of following the instructions and using the guide. I will then use these results to perform an analysis to measure the similarity with the resulting instances to an established instance of the model for each theme and also an analysis of the similarity of its performance.

E.1.2 Attached Materials

- Standard student project ethics form
- Guide and forms (As included in Appendix A)

E.1.3 Approval Code

ES/09/11/012

E.2 TMB Photo Montage Experiment

Ethics submission for TMB photo montage as detailed in chapter 4 section 4.2.

Approval for this experiment was extended from original approval given for the pilot study. The application included below is the original application for the pilot study.

E.2.1 Methodology

The project is a pilot study for a larger evaluation evaluating the effectiveness of a model for describing themes in a narrative, as well as the effectiveness of specific features of the model and metrics used by the model. The evaluation takes the form of a website questionnaire. A prototype generates a photo montage for a specific title and users are asked to evaluate the relevance of the photos to the title for which they were generated, for example users might be asked to consider 10 photos generated for the title 'london in winter', rating either individual photos or groups of photos from 1-5 on their relevance.

The effectiveness of the prototype using the model is compared with two different base cases and a keyword method of montage generation as well in the same way.

The image montages are a result of either running my prototype which selects images based on my thematic model, doing a image search on flickr, selecting images manually, or selecting images randomly from flickr - depending on which set you are looking at. However the result image sets for each method of generation were produced in advance and checked manually before putting them into the questionnaire, the titles were also manually decided by me as well, and there is nothing that I would consider offensive.

E.2.2 Attached Materials

- Standard student project ethics form
- Link to demo version of experiment website[1]

E.2.3 Approval Code

ES/09/01/005

E.3 TMB and Co-occurence Experiment

Approval for the TMB and co-ocurence experiment detailed in chapter 5 section 5.2, due to its similarity, was extended from the earlier TMB experiment with the application detailed above in D.2. Approval was again granted under code ES/09/01/005.

E.4 Narrative Cohesion Experiment

Ethics submission made for narrative cohesion experiment as described in chapter 6 section 6.3.

E.4.1 Methodology

I am investigating the effect of using a thematic system to emphasise themes through illustrations in a short story on how a reader perceives the cohesion of that story. Readers will each read 3 short stories each with illustrations and after each will answer a series of short questions on how they perceived a number of variables we have ascertained as

[1]Hosted at http://droopy.ecs.soton.ac.uk:8080/ThematicModelEval/intro.jsp?id=34&type=2 as of 12/07/2011

related to the narrative cohesion of the story (Logical sense, themes, genre, style, and narrator). The illustrations for the stories will be generated using 3 different methods and each story each participant views will have illustrations from a different method to the others, the illustration method and story pairings will cycle with each participant ensuring the collection of data for all illustration methods on all stories.

The stories are split up into a number of logical sections and each section has an illustration selected by the relevant method. The 3 methods are:

1. Illustrations chosen randomly from the top 5 results for a flickr search for a keyword describing the content of the section.
2. Illustrations chosen randomly from the top 5 results returned by the TMB system which uses both the keyword describing the content of the section and a model of the stories strongest theme.
3. Human selected illustrations from flickr, purposefully selected to be the best illustrations possible, as a base case to put the others in context.

The reason the first two choose an image randomly from a set of 5 candidate images is to minimise the impact of individual anomalous images on the results. The keywords describing the content and themes have been selected by an independent expert who agreed to assist the experiment. The models representing the themes are also constructed by independent assistants who are experienced in using the model and have volunteered to help. Also it is worth noting we have received permission from the author of the short stories to use these stories in the context of this methodology for the experiment.

The experiment will be run through a website, on reaching the initial page users will enter an email address and a link to their experiment page will be sent to them. This is only to ensure that each participant only does the experiment once and the email is hashed to prevent it from being identifiable in any way. The user then reads the instructions and is presented with each story in turn and asked to complete the questionnaire after they have completed reading them. We estimate this process to take approximately 30 minutes, based on the size of the stories and questionnaires.

The website is not yet complete but I have attached documents containing the proposed questionnaire(Thematiccohesion2.docx) and stories (ersinghausstories.docx) so that their suitability may be assessed. It is to be noted that the document I am sending includes 4 stories despite the fact we will be using 3, this is because we have allowed the independent expert providing illustrations to avoid one story if they find it too challenging to provide illustrations for it. He has not completed this process yet and as such we don't know exactly which 3 of those 4 we will be using. Despite this hopefully these documents will be sufficient to ethically asses this experiment in light of the methodology above.

E.4.2 Attached Materials

- Standard student project ethics form
- Questionnaire to be used in the experiment (as detailed in Chapter 6)
- Short stories used (As included in Appendix C)

E.4.3 Approval Code

ES/10/08/001

Bibliography

[1] E. J. Aarseth. *Cybertext: Perspectives on Ergodic Literature*. The Johns Hopkins University Press, 1997.

[2] H. Al-Khalifa and H. Davis. Folksonomies versus automatic keyword extraction: An empirical study. *IADIS International Journal On Computer Science And Information Systems (IJCSIS)*, 1:132–143, 2006.

[3] H. Alani, S. Kim, D. Millard, M. Weal, W. Hall, P. Lewis, and N. Shadbolt. Automatic ontology-based knowledge extraction and tailored biography generation from the web. *IEEE Intelligent Systems*, 18:14–21, 2003.

[4] M. Alberink, L. Rutledge, and M. Veenstra. Sequence and emphasis in automated domain-independent discourse generation. In *Information systems*, pages 1–10, 2003.

[5] R. Alvarez and J. Urla. Tell me a good story: using narrative analysis to examine information requirements interviews during an erp implementation. *SIGMIS Database*, 33:38–52, February 2002.

[6] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, J. T. Eppig, M. A. Harris, D. P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J. C. Matese, J. E. Richardson, M. Ringwald, G. M. Rubin, and G. Sherlock. Gene ontology: tool for the unification of biology. *Nature Genetics*, 25(1):25–29, May 2000.

[7] R. Aylett. Narrative in virtual environments-towards emergent narrative. Technical report, Working notes of the Narrative Intelligence Symposium, 1999.

[8] R. Aylett, S. Louchart, J. Dias, A. Paiva, and M. Vala. Fearnot! - an experiment in emergent narrative. In T. Panayiotopoulos, J. Gratch, R. Aylett, D. Ballin, P. Olivier, and T. Rist, editors, *Intelligent Virtual Agents*, volume 3661 of *Lecture Notes in Computer Science*, pages 305–316. Springer Berlin / Heidelberg, 2005.

[9] R. Aylett, S. Louchart, A. Tychsen, M. Hitchens, R. Figueiredo, and C. D. Mata. Managing emergent character-based narrative. In *Proceedings of the 2nd international conference on INtelligent TEchnologies for interactive enterTAINment*,

INTETAIN '08, pages 5:1–5:8, ICST, Brussels, Belgium, Belgium, 2007. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

[10] P. Bailey. Searching for storiness: story generation from a readers perspective. In *AAAI Fall Symposium on Narrative Intelligence*, 1999.

[11] M. Bal. *Narratology: Introduction to the Theory of Narrative*. University of Toronto Press, January 1998.

[12] R. Barthes. *Mythologies*. Editions du Seuil, 1957.

[13] R. Barthes and L. Duisit. An introduction to the structural analysis of narrative. *New Literary History*, 6:237–272, 1975.

[14] H. Berghel. Cyberspace 2000: dealing with information overload. *Commun. ACM*, 40(2):19–24, 1997.

[15] T. Berners-Lee, J. Hendler, O. Lassila, et al. The semantic web. *Scientific American*, 284(5):34–43, 2001.

[16] M. Bernstein. Card shark and thespis: exotic tools for hypertext narrative. In *Proceedings of the twelfth ACM conference on Hypertext and Hypermedia*, 2001.

[17] K. Bischoff, C. S. Firan, W. Nejdl, and R. Paiu. How do you feel about "dancing queen"?: deriving mood & theme annotations from user tags. In *Proceedings of the 9th ACM/IEEE-CS joint conference on Digital libraries*, JCDL '09, pages 285–294, New York, NY, USA, 2009. ACM.

[18] W. Booth. *The Rhetoric of Fiction*, chapter Types of Narration, pages 69–74. University of Chicago Press, 1974.

[19] P. D. Bra, A. Aerts, B. Berden, B. de Lange, B. Rousseau, T. Santic, D. Smits, and N. Stash. Aha! the adaptive hypermedia architecture. In *Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*, pages 81–84, 2003.

[20] P. D. Bra, P. Brusilovsky, and G.-J. Houben. Adaptive hypermedia: from systems to framework. In *ACM Computing Surveys (CSUR)*, volume 31, 1999.

[21] C. Bremond. *The Logic of Narrative Possibilities*. The Johns Hopkins University Press, 1980.

[22] A. D. Brown, M. Humphreys, and P. M. Gurney. Narrative, identity and change: a case study of laskarina holidays. *Journal of Organizational Change Management*, 18:312–326, 2005.

[23] J. Bruner. The narrative construction of reality. *Critical Inquiry*, 18:1–21, 1991.

[24] J. Brusilovsky, P. Eklund. Web-based education for all: A tool for developing adaptive courseware. In *Computer Networks and ISDN Systems (Proceedings of Seventh International World Wide Web Conference)*, 1998.

[25] P. Brusilovsky. Adaptive hypermedia. *User Modelling and User Adaptive Interactive*, 11:1–2, 2001.

[26] C. Buckley. Automatic query expansion using smart : Trec 3. In *In Proceedings of The third Text REtrieval Conference (TREC-3)*, pages 69–80, 1995.

[27] S. E. Buscaldi D., Rosso P. A wordnet-based query expansion method for geographical information retrieval. In: CLEF 2005 Working Notes, 21-23 September 2005. Vienna, Austria C. Peters (Ed.).

[28] C. B. Callaway and J. C. Lester. Narrative prose generation. In *Proceedings of the 17th international joint conference on Artificial intelligence - Volume 2*, pages 1241–1248, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.

[29] L. Calvi. "lector in rebus": the role of the reader and the characteristics of hyperreading. In *HYPERTEXT '99: Proceedings of the tenth ACM Conference on Hypertext and hypermedia : returning to our diverse roots*, pages 101–109, New York, NY, USA, 1999. ACM.

[30] I. Calvino. *Una pietra sopra*, chapter La sfida al labirinto, pages 82–97. 1980.

[31] C. Carpineto, R. de Mori, G. Romano, and B. Bigi. An information-theoretic approach to automatic query expansion. *ACM Trans. Inf. Syst.*, 19(1):1–27, 2001. Scoring and weighting specific terms for expansion.

[32] M. Cavazza and F. Charles. Dialogue generation in character-based interactive storytelling. In *AAAI First Annual Artificial Intelligence and Interactive Digital Entertainment Conference*, California, USA, 2005.

[33] M. Cavazza, F. Charles, and S. J. Mead. Agents interaction in virtual storytelling. In *Intelligent Virtual Agents, Springer Lecture Notes in Artificial Intelligence 2190*, pages 156–170. Springer-Verlag, 2001.

[34] F. Charles, S. Lemercier, T. Vogt, N. Bee, M. Mancini, J. Urbain, M. Price, E. André, C. Pélachaud, and M. Cavazza. Affective interactive narrative in the callas project. In *Proceedings of the 4th international conference on Virtual storytelling: using virtual reality technologies for storytelling*, ICVS'07, pages 210–213, Berlin, Heidelberg, 2007. Springer-Verlag.

[35] M. Chau, J. J. Xu, and H. Chen. Extracting meaningful entities from police narrative reports. In *Proceedings of the 2002 annual national conference on Digital government research*, dg.o '02, pages 1–5. Digital Government Society of North America, 2002.

[36] Y. Cheong and R. Young. A computational model of narrative generation for suspense. In *AAAI 2006 Computational Aesthetic Workshop, Boston, MA, USA*, 2006.

[37] C. Coupland and A. D. Brown. Constructing organizational identities on the web: A case study of royal dutch/shell. *Journal of Management Studies*, 41:1325–1347, 2004.

[38] J. Derrida. *Limited inc.* Northwestern Univ Pr, 1977.

[39] U. Eco. *Lector in fabula.* Bompiani, Milano, 1979.

[40] U. Eco. *Six Walks in the in the Fictional Woods.* Harvard, 1994.

[41] M. El-Nasr. Interactive narrative architecture based on filmmaking theory. *International Journal of Intelligent Games & Simulation*, 3(1):96–10, 2004.

[42] M. El-Nasr. Interaction, narrative, and drama: Creating an adaptive interactive narrative using performance arts theories. *Interaction Studies*, 8(2):209–240, 2007.

[43] A. Farquhar, R. Fikes, and J. Rice. The ontolingua server: A tool for collaborative ontology construction, 1996.

[44] O. Ferret. How to thematically segment texts by using lexical cohesion? In *ACL-36: Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, pages 1481–1483, Morristown, NJ, USA, 1998. Association for Computational Linguistics.

[45] D. M. Frohlich, D. Rachovides, K. Riga, R. Bhat, M. Frank, E. Edirisinghe, D. Wickramanayaka, M. Jones, and W. Harwood. Storybank: mobile digital storytelling in a development context. In *CHI '09: Proceedings of the 27th international conference on Human factors in computing systems*, pages 1761–1770, New York, NY, USA, 2009. ACM.

[46] G. Fu, C. B. Jones, and A. I. Abdelmoty. Ontology-based spatial query expansion in information retrieval. In *In Lecture Notes in Computer Science, Volume 3761, On the Move to Meaningful Internet Systems: ODBASE 2005*, pages 1466–1482, 2005. Geographical ontology for expanding location queries. types of query expansion (external knowledge based).

[47] K. Gee. The ergonomics of hypertext narative: usability testing as a tool for evaluation and redesign. *ACM J. Comput. Doc.*, 25(1):3–16, 2001. Making hypertext more useable - usability study on hypertext interfaces.

[48] J. Goguen and D. Harrell. Style as a choice of blending principles. *Style and Meaning in Language, Art Music and Design*, pages 49–56, 2004.

[49] A. Graesser, D. McNamara, M. Louwerse, and Z. Cai. Coh-metrix: Analysis of text on cohesion and language. *Behavior Research Methods*, 36:193–202, 2004.

[50] C. Hargood, D. Millard, and M. Weal. A thematic approach to emerging narrative structure. In *Web Science at Hypertext '08*, 2008.

[51] C. Hargood, D. Millard, and M. Weal. Investigating a thematic approach to narrative generation. In *DAH at Hypertext 09*, 2009.

[52] C. Hargood, D. Millard, and M. Weal. Using a thematic model to enrich photo montages. In *Proceedings of the 20th ACM conference on Hypertext and hypermedia*, pages 135–140. ACM, 2009.

[53] C. Hargood, D. Millard, and M. Weal. Capturing the semiotic relationship between terms. *New Rev. Hypermedia Multimedia*, 16(1-2):71–84, 2010.

[54] C. Hargood, D. Millard, and M. Weal. A semiotic approach for the generation of themed photo narratives. In *HT '10: Proceedings of the 21st ACM conference on Hypertext and hypermedia*, pages 19–28, New York, NY, USA, 2010. ACM.

[55] C. Hargood, D. Millard, and M. Weal. Measuring narrative cohesion: A five variables approach. In *Narrative and Hypertext at Hypertext '11*, 2011.

[56] D. Harrell. Shades of computational evocation and meaning: The griot system and improvisational poetry generation. In *Proceedings, Sixth Digital Arts and Culture Conference*, pages 133–143. Citeseer, 2005.

[57] D. Harrell. Walking blues changes undersea: Imaginative narrative in interactive poetry generation with the griot system. In *AAAI 2006 Workshop in Computational Aesthetics: Artificial Intelligence Approaches to Happiness and Beauty*, 2006.

[58] D. Harrell and J. Zhu. Agency play: Dimensions of agency for interactive narrative design. In *Proceedings of the AAAI 2009 Spring Symposium on Narrative Intelligence II. Menlo Park: AAAI Press*, pages 44–52, 2009.

[59] M. A. Hearst. Texttiling: Segmenting text into multi-paragraph subtopic passages. *Computational Linguistics*, 23:33–64, 1997.

[60] M. A. Hearst and C. Plaunt. Subtopic structuring for full-length document access. In *SIGIR '93: Proceedings of the 16th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 59–68, New York, NY, USA, 1993. ACM.

[61] D. Hindle. Noun classification from predicate argument structures. In *In Proceedings of 28th Annual Meeting of the ACL*, pages 268–275, 1990.

[62] P. Hoffmann and M. Herczeg. Distributed storytelling for narrative in spacious areas. In S. Gbel, N. Braun, U. Spierling, D. J, and H. Diener, editors, *Proceedings of Technologies for Interactive Digital Storytelling an Entertainment*, Computer Graphik Edition, pages 346–350, Band 9.Stuttgart: Fraunhofer IRB Verlag, 2003.

[63] E. Hovy. Automated discourse generation using discourse structure relations. *Artificial Inteligence*, 63:341–385, 1993.

[64] J. A. Hudson and L. Shapiro. *Developing Narrative Structure*, chapter From Knowing to Telling: The Development of Children's Scripts, Stories, and Personal Narratives, pages 89–109. American Psychological Association, 1991.

[65] M. Jewell, F. Lawrence, and M. Tufield. Ontomedia: An ontology for the representation of hetrogeneous media. In *Multimedia Information Retrieval Workshop*, 2005.

[66] V. John-Steiner and C. Panofsky. The development of children's retold narratives. Paper presented at the meetings of the society for childrens research, Baltimore, April 1987.

[67] H. Kozima. Text segmentation based on similarity between words. In *Proceedings of the 31st annual meeting on Association for Computational Linguistics*, pages 286–288, Morristown, NJ, USA, 1993. Association for Computational Linguistics.

[68] A. Kuchinsky, C. Pering, M. L. Creech, D. Freeze, B. Serra, and J. Gwizdka. Fotofile: a consumer multimedia organization and retrieval system. In *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit*, CHI '99, pages 496–503, New York, NY, USA, 1999. ACM.

[69] W. Labov and J. Waletzky. *Narrative analysis: Oral versions of personal experience.* Blackwell, 2003.

[70] F. Lawrence, M. Tuffield, M. Jewell, A. Prugal-bennett, D. Millard, M. Nixon, monica schraefel, and N. Shadbolt. Ontomedia - creating an ontology for marking up the contents of heterogeneous media. In *Ontology Patterns for the Semantic Web ISWC-05 Workshop*, Galway, Ireland, 2005.

[71] M. Lebowitz. Creating characters in a story-telling universe. *Poetics*, 13:171–194, 1984.

[72] M. Lebowitz. Story-telling as planning and learning. *Poetics*, 14:483–502, 1985.

[73] M. Lebowitz. Planning stories. In *Program of the Ninth Annual Conference of the Cognitive Science Society*, volume 9, pages 234–241, Seattle, Washington, 1987. Cognitive Science Scoiety Conference.

[74] C. Levi-Strauss. *Structural Anthropology.* Penguin Books, 1972.

[75] C. Linde. Narrative and social tacit knowledge. *Journal of Knowledge Management*, 5:160–171, 2001.

[76] S. Louchart and R. Aylett. Solving the narrative paradox in ves lessons from rpgs. In *Intelligent virtual agents: 4th International Workshop*, 2003.

[77] B. Magerko. A comparative analysis of story representations for interactive narrative systems. *Proceedings of AIIDE 2007*, pages 91–94, 2007.

[78] R. Mandala, R. M, T. Tokunaga, and H. Tanaka. Combining multiple evidence from different types of thesaurus for query expansion, 1999.

[79] N. Maria and M. J. Silva. Theme-based retrieval of web news. In *Selected papers from the Third International Workshop WebDB 2000 on The World Wide Web and Databases*, pages 26–37, London, UK, 2001. Springer-Verlag.

[80] M. Mateas and A. Stern. Faade: An experiment in building a fully-realized interactive drama. In *Game Developers Conference*, 2003.

[81] Y. Matsuo and M. Ishizuka. Keyword extraction from a single document using word co-occurrence statistical information, 2003.

[82] D. P. McAdams. The problem of narrative coherence. *Journal of Constructivist Psychology*, 19(2):109–125, April 2006.

[83] D. McNamara, M. Louwerse, and A. Graesser. Coh-metrix: Automated cohesion and coherence scores to predict text readability and facilitate comprehension. Technical report, Institute for Intligent Systems, University of Memphis, TN, 2002.

[84] M. McQuillan. *The Narrative Reader*. Routledge, London, 2000.

[85] J. R. Meehan. Tale-spin, an interactive program that writes stories. In *In Proceedings of the Fifth International Joint Conference on Artificial Intelligence*, pages 91–98, 1977.

[86] D. Michaelides, D. Millard, M. Weal, and D. D. Roure. Auld leaky: A contextual open hypermedia link server. In *Revised Papers from the nternational Workshops OHS-7, SC-3, and AH-3 on Hypermedia: Openness, Structural Awareness, and Adaptivity*, volume 2266, pages 59–70, 2001.

[87] D. Millard, C. Bailey, T. Brody, D. Dupplaw, W. Hall, S. Harris, K. Page, G. Power, and M. Weal. Hyperdoc: And adaptive narrative system for dynamic multimedia presentations. Technical report, ECS, University of Southampton, 2003.

[88] D. Millard, C. Hargood, M. Weal, L. Warren, and L. Harris. Social media and emergent organizational narratives. In *Narrative and Hypertext at Hypertext '11*, 2011.

[89] G. Miller. Wordnet: An on-line lexical database. *International Journal of Lexicography*, 3(4), 1990.

[90] M. Mitra, A. Singhal, and C. Buckley. Improving automatic query expansion. In *SIGIR '98: Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*, pages 206–214, New York, NY, USA, 1998. ACM.

[91] T. Murray, T. Shen, J. Piemonte, C. Condit, and J. Thibedeau. Adaptivity for conceptual and narrative flow in hyperbooks: The metalinks system. *Adaptive Hypermedia and Adaptive Web-Based Systems*, 1892/2000:155–166, 2000.

[92] U. Musarra-Schroeder. *Il labirinto e la rete: Percorsi moderni e postmoderni nellopera di Italo Calvino*. Bulzoni editore, 1996.

[93] F. Nack and L. Hardman. Denotative and connotative semantics in hypermedia: proposal for a semiotic-aware architecture. *New review of hypermedia and multimedia*, 7(1):7–37, 2001.

[94] F. Nack and A. Parkes. The application of video semantics and theme representation in automated video editing. *Multimedia tools and applications*, 4(1):57–83, 1997.

[95] F. Paradis and C. Berrut. Experiments with theme extraction in explanatory texts. In *In: Second International Conference on Conceptions of Library and Information Science, CoLIS*, pages 433–437, 1996.

[96] H. J. Peat and P. Willett. The limitations of term co-occurrence data for query expansion in document retrieval systems. *Journal of the American Society for Information Science*, 42:378–383, 1991.

[97] V. Propp. *Morphology of the folktale*. Austin, University of Texas Press, 1928.

[98] M. O. Riedl. A comparison of interactive narrative system approaches using human improvisational actors. In *Proceedings of the Intelligent Narrative Technologies III Workshop*, INT3 '10, pages 16:1–16:8, New York, NY, USA, 2010. ACM.

[99] M. O. Riedl and R. M. Young. Character-focused narrative generation for execution in virtual worlds. In *Proceedings of the International Conference on Virtual Storytelling*, pages 47–56, 2003.

[100] M. Roussos, A. Johnson, J. Leigh, C. Barnes, C. Vasilakis, and T. Moher. The nice project: Narrative, immersive, constructionist/collaborative environments for learning in virtual reality. In *In EDMEDIA/ED-TELECOM*, pages 917–922. AACE, 1997.

[101] G. Salton and M. J. McGill. *Introduction to modern information retrieval,*. McGraw-Hill., New York, 1983.

[102] F. Saussure, C. Bally, A. Sechehaye, and A. Riedlinger. *Course in General Linguistics*. McGraw-Hill, 1966.

[103] N. Sawhney, D. Balcom, and I. Smith. Hypercafe: narrative and aesthetic properties of hypervideo. In *Proceedings of the the seventh ACM conference on Hypertext*, pages 1–10, 1996.

[104] N. Shadbolt, W. Hall, and T. Berners-Lee. The semantic web revisited. *Intelligent Systems, IEEE*, 21(3):96–101, 2006.

[105] R. Shaw, R. Troncy, and L. Hardman. Lode: Linking open descriptions of events. *The Semantic Web*, pages 153–167, 2009.

[106] M. B. Skov and J. Stage. Designing interactive narrative systems: is object-orientation useful? *Computers & Graphics*, 26(1):57 – 66, 2002.

[107] J. Sturrock. *Structuralism and Since: from Levi Strasuss to Derrida*. Oxford University Press, 1979.

[108] I. Swartjes and M. Theune. The virtual storyteller: Story generation by simulation. In *In: Proceedings of the 20th Belgian-Netherlands Conference on Artificial Intelligence (BNAIC)*, 2008.

[109] N. Szilas. Idtension: a narrative engine for interactice drama. In *Proceedings of TIDSE 03*, 2003.

[110] M. Theune, S. Faas, A. Nijholt, and D. Heylen. The virtual storyteller: Story creation by intelligent agents. In *TIDSE 2003: Technologies for Interactive Digital Storytelling and Entertainment*, 2003.

[111] Q. T. Tho, S. C. Hui, A. C. M. Fong, and T. H. Cao. Automatic fuzzy ontology generation for semantic web. *Knowledge and Data Engineering, IEEE Transactions on*, 18(6):842–856, 2006.

[112] T. Todorov, R. Howard, and J. Culler. *The Poetics of Prose*. Oxford: Basil Blackwell, 1977.

[113] B. Tomashevsky. *Russian Formalist Criticism: Four Essays*, chapter Thematics, pages 66–68. University of Nebraska Press, 1965.

[114] R. Troncy, B. Malocha, and A. Fialho. Linking events with media. In *Proceedings of the 6th International Conference on Semantic Systems*, pages 1–4. ACM, 2010.

[115] M. Tuffield, S. Harris, D. P. Dupplaw, A. Chakravarthy, C. Brewster, N. Gibbins, K. O'Hara, F. Ciravegna, D. Sleeman, Y. Wilks, and N. R. Shadbolt. Image annotation with photocopain. In *First International Workshop on Semantic Web Annotations for Multimedia (SWAMM 2006) at WWW2006*, Edinburgh, United Kingdom., 2006.

[116] E. M. Voorhees. Query expansion using lexical-semantic relations. In *SIGIR '94: Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 61–69, New York, NY, USA, 1994. Springer-Verlag New York, Inc.

[117] M. Weal, H. Alani, S. Kim, P. Lewis, D. Millard, P. Sinclair, D. D. Roure, and N. Shadbolt. Ontologies as facilitators for repurposing web documents. *International Journal of Human-Computer Studies*, 65:537–562, 2007.

[118] J. Xu and W. B. Croft. Query expansion using local and global document analysis. In *SIGIR '96: Proceedings of the 19th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 4–11, New York, NY, USA, 1996. ACM.

[119] R. M. Young. Story and discourse: A bipartite model of narrative generation in virtual worlds. *Virtual Reality*, 2003.

[120] R. M. Young, M. E. Pollack, and J. D. Moore. Decomposition and causality in partial-order planning, 1994. original longbow stuff. AI action decision planners.

[121] R. M. Young and M. Riedl. Towards an architecture for intelligent control of narrative in interactive virtual worlds. In *IUI '03: Proceedings of the 8th international conference on Intelligent user interfaces*, pages 310–312, New York, NY, USA, 2003. ACM.

[122] R. M. Young, M. O. Riedl, M. Branly, A. Jhala, R. Martin, and C. Saretto. An architecture for integrating plan-based behavior generation with interactive game environments. *Journal of Game Development*, 1(1):51–70, 2004.

[123] X. S. Zhou and T. S. Huang. Unifying keywords and visual contents in image retrieval. *IEEE MultiMedia*, 9(2):23–33, 2002. Key word paper, keyword searching for images, appending relevant keywords to query expansion, effectiveness of keyword searching.