

A Semiotic Approach for the Generation of Themed Photo Narratives

Charlie Hargood
Learning Societies Lab
ECS, Southampton University
Southampton, United Kingdom
cah07r@ecs.soton.ac.uk

David E Millard
Learning Societies Lab
ECS, Southampton University
Southampton, United Kingdom
dem@ecs.soton.ac.uk

Mark J Weal
Learning Societies Lab
ECS, Southampton University
Southampton, United Kingdom
mjlw@ecs.soton.ac.uk

ABSTRACT

A wide variety of systems could be considered ‘narrative systems’, either directly working towards generating rich narratives or, more frequently, because they present or handle information in a narrative context. These narratives, generated or otherwise handled, may contain themes; an essential part of the subtext of narrative communicating important concepts outside the capabilities of the literal meaning of the content and forming the thematic cohesion that aids the flow of the presented narrative. However despite this very little work has been undertaken to understand of take advantage of these themes, particularly in narrative generation where the presence of well defined themes may improve the richness of those generated narratives. In this paper we evaluate the performance of a system utilising a thematic model in order to generate simple narratives in the form of photo montages compared to a keyword based system that does not. The experiment demonstrates that the system utilising the thematic model is capable of successfully connoting themes within these narratives. It also shows that the relevance of the resulting narratives to the titles used to generate them is higher in the thematic system than those generated by the other system.

Categories and Subject Descriptors

H.1 [Models and Principles]: General

General Terms

Standardization, Human Factors, Experimentation

Keywords

Narrative, Narrative Generation, Thematics, Folksonomies, Semiotics

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HT'10, June 13–16, 2010, Toronto, Ontario, Canada.

Copyright 2010 ACM 978-1-4503-0041-4/10/06 ...\$10.00.

1. INTRODUCTION

Themes play an important part in information. Themes can be representative of the subtext implied beyond the literal meaning of information, and thematic cohesion helps information flow together as successful narratives. However while some work in narratology seeks to deconstruct and understand themes, such as work in thematics [25], there is little understanding of the importance of them in systems that might benefit them.

Narratives are present in a wide variety of systems in different forms, from requested information or media presented to a user to systems that change or author their own narratives. Communication or understanding through stories is something that comes naturally to people and this is why narrative is such a prevalent form of representation of information. However when we say narratives we are not solely referring to novels or movies but indeed any presentation of potential human experience [18], from written stories, to photo montages such as those explored in this paper and Photocopain [26], to search results such as in Topia [1].

Narrative generation is an area of research that seeks to generate its own narratives on demand. This is a powerful prospect, being able to generate narratives on the fly would allow content delivered to be automatically customised for a user or created based entirely on a set of desires. Possibilities for success in this area include systems such as an automatic news reader that knows what the user is interested in or a computer game that constantly regenerates its plot based on the players actions. The existing projects in this space work a number of different approaches and some experience some success but there are also common problems within the field. Stories are often bland or directionless or heavily bound to the plot structure of their chosen genre.

Perhaps more common are systems that work with narratives without generating them, such as some hypertexts. Treating information or media as a story in order to apply narrative principles to customise it or working in the space of adaptive ways of presenting narratives. Such work often falls in the hypertext and adaptive hypermedia field where the structuring and linking of content together is used to form a presentation of a grander narrative providing adaptable entertainment or engaging presentations of content.

We propose the use of the thematic model presented in [12] to improve the thematic cohesion in systems that utilise narrative which will then enrich the resulting narratives, which can make them more relevant thematically, and giving them more subtextual depth. The thematic model offers a way of formally defining specific themes and how they are connoted

from elements within a narrative. Such definitions may be used to inform the way narratives are adapted or generated to improve thematic cohesions towards specific desired themes by evaluating the thematic relevance of content.

In this paper we describe the thematic model and a prototype that utilises it to create simple narratives in the form of photo montages. We also describe the full results and analysis of an experiment that compared the montages generated by the thematic system to ones created by a simple keyword based system and explore the thematic relevance of the resulting montages and what advantages the thematic model offers them.

2. BACKGROUND

2.1 Narratology

Narratology is the analytical study of literature and as such represents a large body of work exploring the effect of narrative. Although heavily based in literature it is often applied to a range of media, where a narrative can be defined as any presented series of human experiences [18]. It can be difficult to draw solid conclusions from much literary criticism making it difficult to use it as a solid basis to contribute towards narrative systems. However a structuralist approach to narratology looks at deconstructing narrative in a system of associated parts and elements and analysing the system in which they fit together, this abstraction of the structure of narrative provides something much more tangible.

A prominent structuralist Barthes asserts in [4] that narrative may be deconstructed into a story and a discourse where the story (or fabula) represents a chronology of all the information to be communicated and the discourse (or *sjuzhet*) represents what parts of the story are told and how those parts are presented (shown in Figure 1). Other layered approaches to narrative structure analysis have since been made as well such as Bal [2] which similarly makes a division between the sum of all events (Fabula) and the story that is told but divides the selection of story elements and its presentation (collectively the discourse according to Barthes) into separate layers of story and narrative.

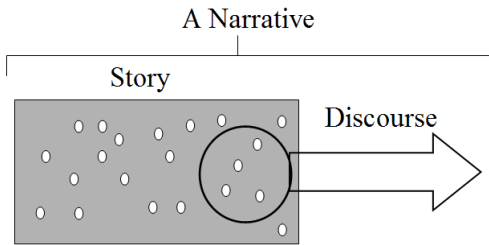


Figure 1: A narrative can be deconstructed into story and discourse

While in terms of generating a narrative the ‘story’ could be considered to set of all resources the system has in order to create it (characters, places, events, or even individual narrative items) the discourse is composed of a multitude of different mechanics including how the story is presented, what medium is used, the style, the genre, the themes of the

narrative, and any bias present, it is the telling of the story which forms the complete narrative. In [5] Booth explored the idea of an authorial voice, how authors position them self within a narrative, perhaps as a narrator or more often as a bias in the discourse, this to is a part of discourse. It is possible that the way a narrative is presented, its discourse, is even more important than its content or story. How these structures apply to different narratives is however not always clear, in [10] Calvi comments on how the concepts of Fabula and Plot (with similar definitions to Barthes Story and Discourse) fit within traditional texts and hypertexts. Calvi asserts that the fabula and plot are intertwined and that in some cases, particular where a narrative is composed of other narratives, there is no clear fabula that the plot has arisen from. Calvi also asserts how these structures can be used to guide the reader through analepsis and prolepsis the selection of fabula elements at key times in the plot guiding the users experience.

The study of thematics approaches themes with a structuralist method of deconstruction and attempts to identify the narrative elements that communicate themes. Themes are an important part of the discourse; a subtle subtext of the narrative communicating an objectivity aside from the literal events in the plot and add a personal influence to the story. 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’) [25]. He describes a structure of themes being built out of sub-themes and motifs. 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.

For an explanation of how themes become apparent from parts of a narrative we can look to semiotics. Semiotics or semiology is the study of signs and how we extract meaning from them. Saussure wrote that all signs are built of two parts [23], a signifier (the physical signal from the sign such as the appearance of an apple) and a signified (the denotation of that sign such as the concept of ‘apple-ness’ or ‘fruit’). Barthes made a distinction between denotative signs (signifiers that lead directly to their signified, such as a word having a literal meaning) and connotative signs (signifiers that lead indirectly to some contextual or culturally important signified, such as the red light implying Stop to a driver) [3]. Barthes goes on to point out that should a sign connote something then the signifier of such a sign would itself be built out of a denotative sign (a picture of a red light denotes a red light, red light connotes Stop). In such a way we can draw contextual cultural concepts from static basic objects that in a particular context have a greater meaning.

2.2 Narrative Systems

Narrative generation has a variety of applications in systems that deal with different information, as a narrative can be any collection of human experience it is not limited to written prose but to any representation of human experience [18]. Some systems use narrative as a lens through which to view a larger collection, for example Photocopain [26] which presents narrative photo montages. While many systems seek to generate full narratives for entertainment such as the virtual storyteller [24] and AConf [22] some sys-

tems use narrative generation to add additional meaning to information by representing it as a narrative using narratological devices like sequencing, emphasis and omission. This can be seen in systems such as Topia [1] and some adaptive hypermedia systems such as AHA! [6], where narrative generation allows for the adaptive presentation of information based on initial objectives which has been highlighted as fundamental to adaptive hypermedia [7].

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 in narration and presentation [24]). 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 illustrates this process.

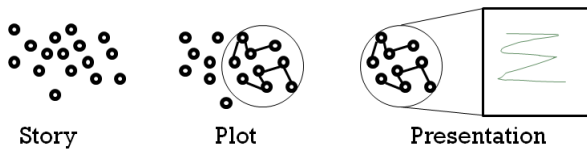


Figure 2: Narrative generation can be broken down into three stages

According to Riedl and Young [22] narrative systems take either a character 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. Riedl and Young [22] also identify a third approach in the form of story centric approaches these are however less common and due to their more linguistic focus are less relevant to this research.

Character centric narrative generation centers, as the name suggests, around the narratives characters. The idea being that narrative is emergent from the choices these intelligent characters make. As such character centric approaches are often centered around creating an intelligence, often an agent, that simulates the characters behaviour and personality in an environment, the system then reports on these actions taken by the variety of characters in the chosen environment. Examples of character centric approaches are the very early TaleSpin [19] and the more recent agent based approaches by work such as that by Cavazza [11].

Contrastingly author centric approaches are more detached from the stories content and instead aim to simulate the process of creating the narrative, that of the author. Author centric approaches utilise a set of rules and structures that define how a narrative is written and use them to compile available resources into a narrative. Author centric approaches are often grammar based but can also be governed by rules and structures presented in other ways such as a content selection process followed by emphasis. Examples of author centric approaches are systems such as ArtEquAKT [28], Topia [1], and to some extent Universe [15].

However often the most successful narrative generation systems are based on a compromise solution where both

the content of the narrative and the creation process are utilised. These compromise approaches often have more in common whether character centric approaches, relying on an emergent narrative from simulated content, but control the intelligence driving the agents with author based rules. Systems such as the Virtual Storyteller [24] and Facade [17] use an extra 'director' agent which vets actions of characters and encourages certain directions in order to make the story better. Where as AConf [22] uses plot planning to construct a structure within which the characters may take actions.

Ultimately existing narrative generation systems experience at best mixed success. Character centric approaches can seem bland and directionless, the reporting of the content not giving a rich enough discourse. While author centric approaches are often too strongly tied to the rules of their given genre, and inevitably seem stale and formulaic. Compromise solutions work towards mitigating some of these problems but still lack the richness of human authored narrative. Potentially this could be attributed to the discourse, in that the systems concentrate on the creation and selection of content rather than the telling of the narrative.

2.3 Term Expansion

Improving the relevance of content by expanding a given term used in its creation is not a new concept. Query expansion to improve the relevance of search results has been used for some time. The principle of term expansion is to understand relationships between terms so that when a term is used in a query it can be expanded to other terms related to it in order to increase the chance of finding positive matches.

The different methods of terms expansion are often distinguished by what relationship is used to expand the given term. Systems might at use a generic thesaurus to connect terms through synonyms, antonyms, or other similar relationships. However Voorhees work [27] exploring this method using WordNet [20] has demonstrated that this offers little advantage. Although Buscaldi experienced more success in his work [9] it is only in a very specific area and cannot be said to demonstrate broad success.

Co-Occurrence term expansion expands a given term with terms that are likely to exist along side it. Such systems take a given corpus of items and use them calculate the probability of a given term co-occurring with another, high probability terms are then chosen for expansion. Numerous systems have reported success and improvement with co-occurrence such as in the work done by Buckley [8], and further work has been done to improve the way co-occurrence is used such as how the corpus is selected in work by Xu [29] and how the probability of co-occurrence is calculated by Peat [21].

Mandalas performed an in depth comparison between a variety of term expansion methods and combination approaches in [16] and found co-occurrence to be the most effective. There seems to be little doubt that generally speaking co-occurrence is the most effective way of increasing query relevance through term expansion. However, it is not without its own drawbacks; co-occurrence is reliant on the corpus used to train it where misinterpreted co-occurrence of terms can influence results. Strings of frequently co-occurring terms can lead to query drift as identified by Zhou in [30], where through a list of connections from the expansion a query is tainted with concepts irrelevant to the users desire.

3. THE THEMATIC MODEL

3.1 The Model

In order to give systems an understanding of the thematic content of narratives, and as such better handle or generate them, we have developed a thematic model. The model is a representation of how themes are present within a narrative, a structure with a set of rules controlling how a theme is defined and how it is connoted from a narrative. This model was first presented in [12], and then further explored how it could be utilised by narrative systems in [13]. By creating a model that controls the structure of thematic definitions these definitions may be used by systems to detect the presence of themes, imbed themes in generated content, or ensure the thematic coherence of a set of items.

The Model is based upon two key thematic elements; *Themes*, and *Motifs*, and how they become apparent through *Features* within the narrative. Features within the narrative denote Motifs and from these Themes can be connoted. We use the term Narrative-Atoms or *Natoms* to describe the segments of content that make up a narrative; small atomic pieces of narrative that cannot be further broken down, for example a single photo or paragraph. The content of these natoms is rich with information, however only some of it visible to a machine (such as generated meta data and authored tags), we call these visible computable elements *Features*. 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 *food* in the context of a gathering may connote *feasting*. These themes, when combined with other themes or motifs could in turn be used to further connote other themes, for example feasting might connote celebration. This forms the foundation of our thematic model of a narrative:

- Natoms contain tagged features
- Features denote motif's
- Themes are connoted by other themes and motifs

This model is displayed diagrammatically in figure 3. Figure 3 also demonstrates how the semiotic relationships between the thematic elements form signs similar to those described by Barthes [3] through the features and motifs forming a denotative sign that further connotes the themes.

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

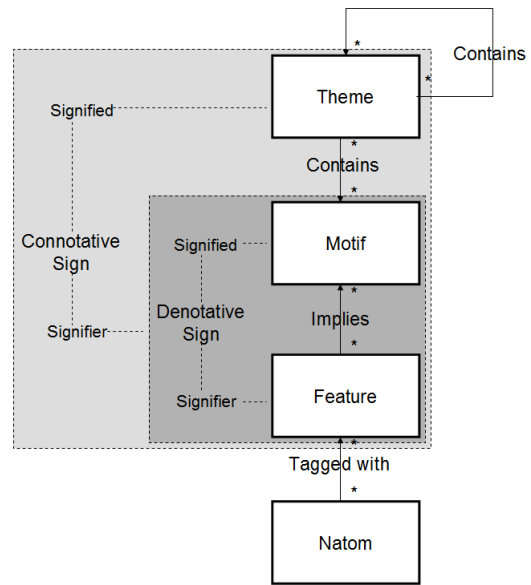


Figure 3: The Thematic Model

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.

These rules could be articulated more formally. We have been experimenting with a variety of constraint based languages for making a formal expression of the rules and have settled on DLV¹, a language that extends datalog. 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 4.

A formal process through which definitions of themes within the terms of this model are authored has been developed. Based on a decomposition of the process of expanding a theme with relevant terms and classifying them as related to either particular themes or motifs before further consolidating them based on their justification for connoting the

¹<http://www.dbai.tuwien.ac.at/proj/dlv/>

```

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 4: Rules of the thematic model expressed in DLV

root theme. This is further followed by recursively expanding any sub themes and removing elements that, while associated with the root theme, also contain elements that are irrelevant to it. The performance of this formal process is current still under analysis and experimentation.

3.2 An Example

Figure 5 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 text², and two photographs 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.

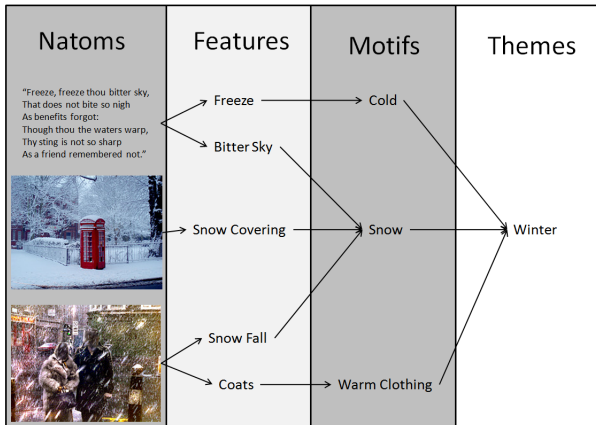


Figure 5: A Worked Example

3.3 The TMB

To evaluate the effectiveness of the model a prototype was built which utilised it. The prototype would build simple photo montages (as a simple form of narrative) around titles which contained a designed subject and desired themes. In doing this we would be able to ascertain firstly if definitions made in terms of the model could be used by a system to successfully imbed that theme and secondly to compare it to other systems in order to measure any improvement in relevance on thematic queries. As a photo montage builder that utilised the thematic model this initial prototype was called the Thematic Model Builder or TMB.

²text from William Shakespears Blow, Blow, Thou Winter Wind

The prototype itself was written in Java with a simple JSP front end. As a source of natoms (photos in this case) Flickr³ was used due to its large amount of readily available and well tagged items. The definitions of themes that the prototype would use were written in XML, each file representing an element. 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 and then fully recursively expanded (all their subthemes and motifs were defined as well). Using this method it would be impractical to expand a full list of features for every motif but a large list of likely denoting features was made for each. The themes selected for the prototype were *Winter*, *Spring*, *Family*, and *Celebration*.

The prototype generated montages by taking a desired length, a desired content, and a desired list of comma separated themes. The TMB then makes a search in Flickr for the desired content and forms a fabula using the top 30,000 images. The thematic quality 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 images *component coverage* and *thematic coverage* is calculated. How these are calculated and how thematic quality is calculated from them is presented in the equations 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 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 in question.

$$\begin{aligned}
 TC &= (t * 100)/T \\
 CC &= (c * 100)/C \\
 TQ &= (TC + CC)/2
 \end{aligned}
 \tag{1}$$

This means the final thematic quality is expressible as a percentage and is based on how many of the desired themes the image is relevant to and how relevant it is to each theme.

4. THE PHOTO NARRATIVE EXPERIMENT

4.1 Methodology

For this experiment we sought to evaluate how effective a system utilising the thematic model would be. The experiment would need to evaluate whether the system was

³<http://www.flickr.com>

able to generate results that successfully connoted the desired themes. It would also be important to compare these results to those produced by a simple existing system to see what value, if any, the thematic approach afforded. As the TMB is based in part on Flickr we elected to compare it to Flickr’s keyword search. As well as comparing the thematic relevance of both approaches for individual images we were keen to see how well the thematic system performed in the more narrative context of many ‘natoms’; in this case a photo montage. From this we see two key objectives emerge:

- Evaluate the effectiveness of the TMB in selecting images connoting desired themes in comparison to simple keyword search.
- Evaluate the effectiveness of the TMB in the 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 contains a content subject and a desired theme (for example, ‘London in Winter’). The images and sets were generated for four different methods to be compared:

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

In this way we 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 under each 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 we 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 well regular or complimentary theme 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 systems 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 we 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 we randomised the order in which montages appeared. We 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 naturally flow and would artificially seem to be stronger montages. Finally users were only allowed to take the evaluation once, a unique evaluation link for each user was given out per email address. The methodology for this experiment was granted ethics approval from the universities ethics committee.

Having finalised the methodology we 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.

Our 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 we 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 we found the test to be stable and only minor rewording of instructions to make the test easier to understand. The results of the pilot study are present in a published paper [14].

4.2 Results and Analysis

The full evaluation achieved 107 test subjects and some significant results. The mean rating of natoms from the TMB is higher than from a keyword search (Flickr) in both single and group images. Figure 6 and Tables 1 and 2 show the data and t-tests for single images. Figure 7 and Tables 3 and 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 percent 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 1: Single Images Rating Frequency

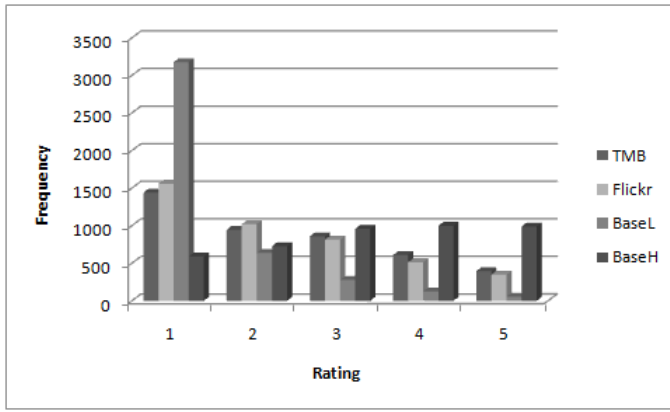


Figure 6: Single Image Rating Frequency

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 2: Single Images Rating Statistics

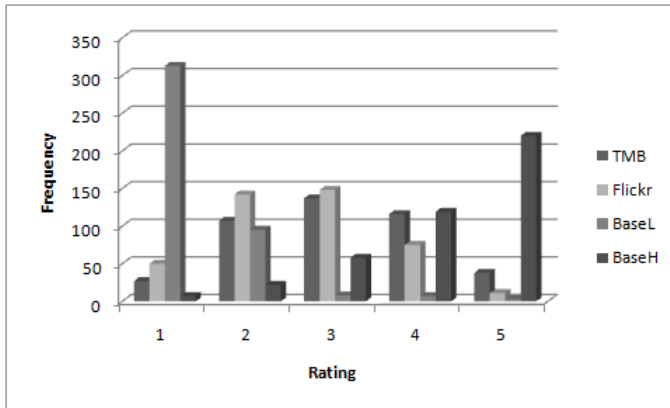


Figure 7: Grouped Image Rating Frequency

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 3: Grouped Images Rating Frequency

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: Grouped Images Rating Statistics

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). Figures 8 and 9 show the relevant means in a way that they can be compared. These ranges are rather smaller than we might expect, and in this context the improvement given by TMB is rather more impressive.

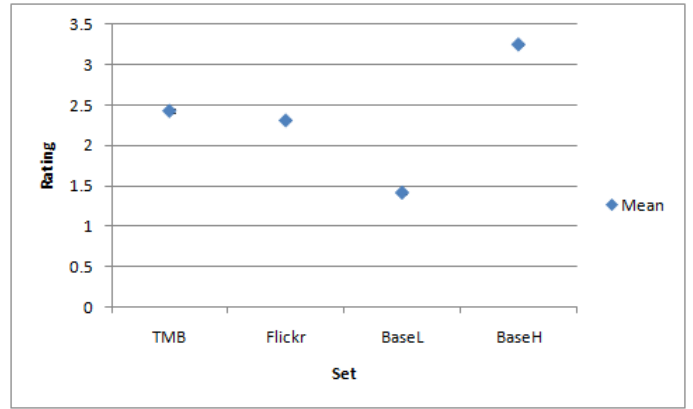


Figure 8: Single Image Mean and Std. Dev.

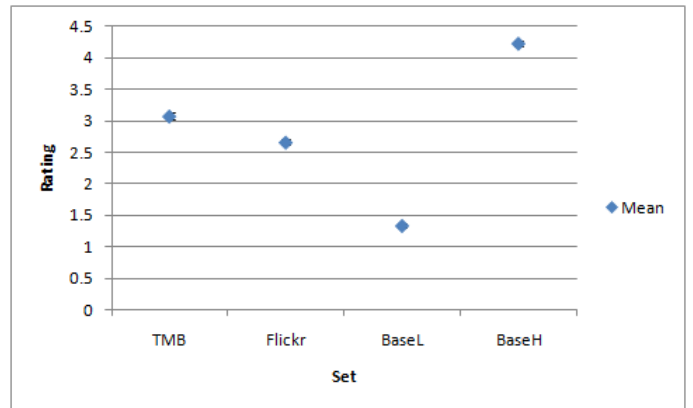


Figure 9: Grouped Image Mean and Std. Dev.

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 percent probability of error. The data shown in table 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 percent probability of error.

As explained we 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 6 and 7 and figures 10 and 11 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

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 5: Grouped Images Improvement Statistics

table in figure 12 and figures 13 and 14 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 6: Single Images Single/Multiple Themes in Title Contrast Statistics

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 7: Grouped Images Single/Multiple Themes in Title Contrast Statistics

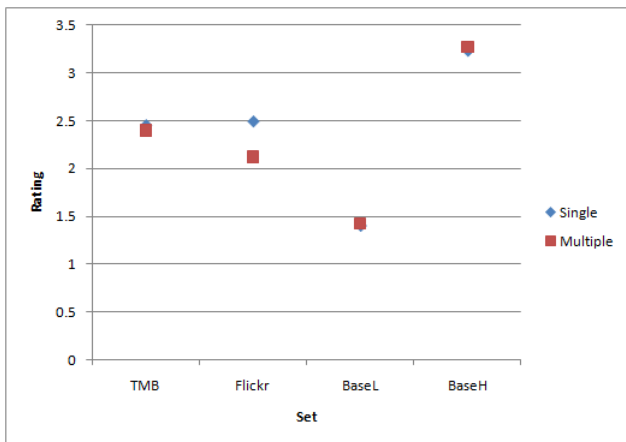


Figure 10: Single Images Single/Multiple Themes in Title Contrast

The results show that in both single and grouped images Flickr's keyword search always performed worse in titles with multiple themes where as the TMB only performed worse on

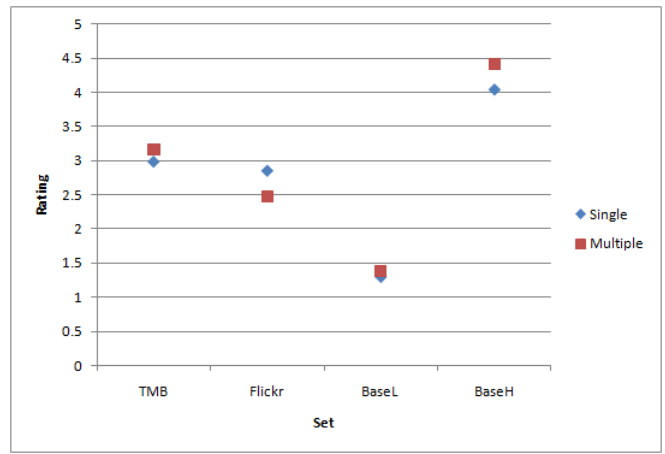


Figure 11: Grouped Images Single/Multiple Themes in Title Contrast

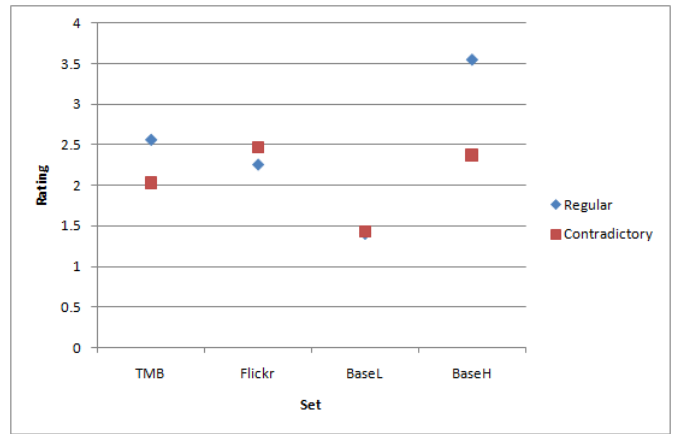


Figure 13: Single Images Contradictory/Regular Theme Fabula Pairing in Title Contrast

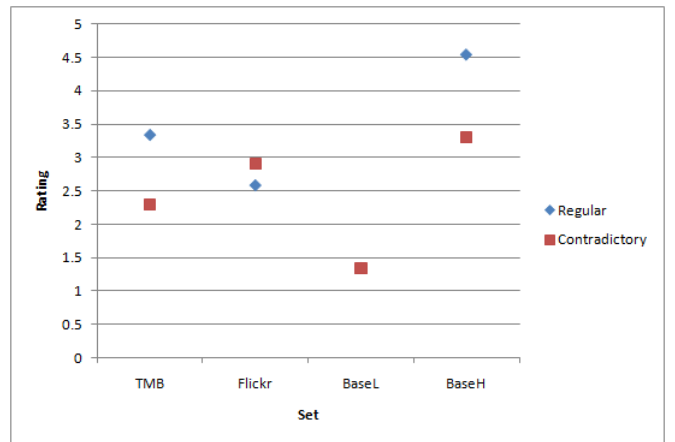


Figure 14: Grouped Images Contradictory/Regular Theme Fabula Pairing in Title Contrast

single images and that TMB performed better than Flickr in all other cases, multiple themes included. The results also show that in both single and grouped images Flickr performs

Set	Grouped			Single		
	Mean	SD	Variance	Mean	SD	Variance
TMB Regular	3.332	0.988	0.977	2.563	1.336	1.785
Flickr Regular	2.578	0.966	0.934	2.258	1.282	1.644
BaseL Regular	1.337	0.656	0.431	1.409	0.822	0.675
BaseH Regular	4.534	0.737	0.544	3.546	1.265	1.601
TMB Contradictory	2.292	0.894	0.799	2.028	1.224	1.499
Flickr Contradictory	2.905	1.001	1.001	2.464	1.322	1.749
BaseL Contradictory	1.339	0.702	0.493	1.429	0.864	0.746
BaseH Contradictory	3.301	1.034	1.069	2.370	1.204	1.451

Figure 12: Grouped and Single Images Contradictory/Regular Theme Fabula Pairing in Title Contrast Statistics

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 score nouns with both themes present highly. 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 its 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 further more it seems the TMB is very strong within a group context, this could lead us to believe it could perform similarly strongly within a narrative context. 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.

5. CONCLUSIONS AND FUTURE WORK

The work presented in this paper shows a promising beginning for investigation into the importance of thematic in narrative systems. The 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 in a narrative context 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 TMBs improved performance over keyword search demonstrates both of these; the higher relevance of its results than 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. Work is currently underway investigating a way to formalise the way themes are defined as well as supporting the capture of a wide variety of themes. There are however other issues that need to be explored; currently the thematic approach has only been shown to be more relevant than simple keyword approaches. It is possible 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. A future experiment might seek to explore the difference between utilising term expansion on a semiotic basis, such as the thematic model, and more automatic statistical expansion, and which produces more thematically cohesive and relevant results.

Specifically with regards to the narrative perspective there remains to be further investigation on the importance of thematic cohesion in more complicated narratives. While photo montages are a good starting point the next step for this research is to investigate whether a narrative generation system could benefit from the use of the thematic model or whether a more complex narrative such as a short story could be enriched and have its thematic cohesion improved by being represented through a system with a thematic understanding of its content, using emphasis and omission to highlight desired themes.

6. REFERENCES

- [1] M. Alberink, L. Rutledge, and M. Veenstra. Sequence and emphasis in automated domain-independent discourse generation. In *Information systems*, pages 1–10, 2003.
- [2] M. Bal. *Narratology: Introduction to the Theory of Narrative*. University of Toronto Press, January 1998.
- [3] R. Barthes. *Mythologies*. Editions du Seuil, 1957.
- [4] R. Barthes and L. Duisit. An introduction to the structural analysis of narrative. *New Literary History*, 6:237–272, 1975.
- [5] W. Booth. *The Rhetoric of Fiction*, chapter Types of Narration, pages 69–74. University of Chicago Press, 1974.

- [6] 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.
- [7] P. D. Bra, P. Brusilovsky, and G.-J. Houben. Adaptive hypermedia: from systems to framework. In *ACM Computing Surveys (CSUR)*, volume 31, 1999.
- [8] C. Buckley. Automatic query expansion using smart : Trec 3. In *In Proceedings of The third Text REtrieval Conference (TREC-3)*, pages 69–80, 1995.
- [9] 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.).
- [10] 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.
- [11] 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.
- [12] C. Hargood, D. Millard, and M. Weal. A thematic approach to emerging narrative structure. In *Web Science at Hypertext08*, 2008.
- [13] C. Hargood, D. Millard, and M. Weal. Investigating a thematic approach to narrative generation. In *DAH at Hypertext 09*, 2009.
- [14] C. Hargood, D. Millard, and M. Weal. Using a thematic model to enrich photo montages. In *Proceedings of Hypertext 09*, 2009.
- [15] 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 Society Conference.
- [16] R. Mandala, R. M. T. Tokunaga, and H. Tanaka. Combining multiple evidence from different types of thesaurus for query expansion, 1999.
- [17] M. Mateas and A. Stern. Façade: An experiment in building a fully-realized interactive drama. In *Game Developers Conference*, 2003.
- [18] M. McQuillan. *The Narrative Reader*. Routledge, London, 2000.
- [19] 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.
- [20] G. Miller. Wordnet: An on-line lexical database. *International Journal of Lexicography*, 3(4), 1990.
- [21] 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.
- [22] 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.
- [23] F. Saussure, C. Bally, A. Sechehaye, and A. Riedlinger. *Course in General Linguistics*. McGraw-Hill, 1966.
- [24] 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.
- [25] B. Tomashevsky. *Russian Formalist Criticism: Four Essays*, chapter Thematics, pages 66–68. University of Nebraska Press, 1965.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] 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.
- [30] 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.