Annotating Musical Theatre Plots on Narrative Structure and Emotional Content*†

Pablo Gervás[1], Raquel Hervás[2], Carlos León[3], and Catherine V. Gale[4]

1 **Universidad Complutense de Madrid, Spain**
pgervas@ucm.es
2 **Universidad Complutense de Madrid, Spain**
raquelhb@ucm.es
3 **Universidad Complutense de Madrid, Spain**
cleon@ucm.es
4 **Wingspan Productions, London, UK**
cat@wingspanproductions.co.uk

Abstract

Although theoretical models of the structure of narrative arising from systematic analysis of corpora are available for domains such as Russian folk tales, there are no such sources for the plot lines of musical theatre. The present paper reports an effort of knowledge elicitation for features that characterise the narrative structure of plot in the particular domain of musical theatre. The following aspects are covered: identification of a valid vocabulary of abstract units to use in annotating musical theatre plots, development of a procedure for annotation–including a spread-sheet format for annotators to use, and a corresponding set of instructions to guide them through the process – selection of a corpus of musical theatre pieces that would constitute the corpus to be annotated, the annotation process itself and the results of post-processing the annotated corpus in search for insights on the narrative structure of musical theatre plots.

1998 ACM Subject Classification D.2.1 Requirements/Specifications Elicitation methods, I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases Narrative annotation, conceptual representation of narrative, character functions, narrative schemas, musical theatre

Digital Object Identifier 10.4230/OASIcs.CMN.2016.11

1 Introduction

Storytelling systems require the representation and manipulation of large amounts of knowledge. This involves not only the product itself – stories represented at various levels of detail – but also the knowledge resources that are required to inform the construction processes. There are a number of possible sources for extracting such knowledge from existing literature on narrative. However, there is little to be found if plot generation for musical theatre is required.

* This work was partially supported by the project WHIM (611560), funded by the European Commission under FP7, the ICT theme, and the Future Emerging Technologies (FET) program.
† The authors want to acknowledge the fundamental contribution made to this annotation initiative by the following people: Julian Woolford (GSA, University of Surrey), Kat Mace (Researcher, Wingspan Productions), James Sandy (Assistant Producer, Wingspan Productions), Archie Baron (Creative Director, Wingspan Productions). For more information on the associated television and theatrical production see: http://www.wingspanproductions.co.uk/.

The present paper reports an effort of knowledge elicitation for features that characterise the narrative structure of plot in the particular domain of musical theatre.The need for such knowledge arose when researchers at Universidad Complutense de Madrid (UCM) were asked by Wingspan Productions to explore the possibility of producing a software tool that could automatically generate plot lines for musical theatre. This took place as part of a broader effort to explore the extent to which it would be possible to generate the premise, setting, structure and music/lyric content for a new piece of musical theatre, using computational methods. The scientific and theatrical production processes invovled in such an effort were to be documented for Sky Arts (UK) in 2x 1hr documentaries. The researchers at UCM had been working on tools for narrative generation within the domain of Russian folk tales, based on the Vladimir Propp's "Morphology of the Folk Tale" [16]. However, the background research carried out to inform those tools had also identified the limited applicability of that particular formalism beyond the domain of Russian folk tales, which presents significant differences with the domain of musical theatre. For instance, there is no mention of love affairs between characters in Russian folk tales, whereas this is a very important feature in musical theatre. As a result, if the available story generation functionalities were to be applied to the domain of musical theatre, it became imperative to undertake an analysis of the regularities of musical theatre plots much in the line of the work Propp had done for Russian folk tales. Although the present effort cannot be compared to Propp's original work in terms of thoroughness and academic rigour, it constitutes a valuable empirical source for any further analyses that may be attempted.

2 Previous Work

There are multiple dimensions when considering knowledge representation for story generation. Gervás and León [10] provided a list of the most relevant classifications, and proposed their own list of suitable dimensions obtained from the different aspects of a narrative. The dimensions proposed were: discourse, simulation, causality, character intention, theme, emotion, authorial intention, and narrative structure.

The annotation effort carried out to build the case base described in the present paper concerned only two of these aspects: the narrative structure aspect, and the emotional aspect.

2.1 Existing Schemas for Modelling Narrative

Narratives are created and told according to a certain structure (or set of structures). The aristotelian narrative arc, for instance, is one of the first instances of this kind of structure. Vladimir Propp's analysis of Russian folk tales [16] is known to have produced a semi-formal description of the structure of these tales that has acted as inspiration for several story generation systems, both sequential and interactive. Its exhaustive description of the constituent elements of tales of this kind, and the enumeration of the patterns they follow provided a very useful starting point for researchers looking for computational implementations of story generators.

Propp identified a set of regularities across a corpus of Russian folk tales in terms of *character functions*, understood as acts of the character, defined from the point of view of their significance for the course of the action. These character functions have been used as building blocks for story generators in many instances [7, 8].

Gervás et al. [11] describe an effort to mine existing literature on narrative to identify descriptions of the structure of narrative that might serve as instances of the schemata that might be used as skeletons over which to build more elaborate instantiations of narrative

during narrative generation processes. This approach led to the identification of a particular way of abstracting units of description that carry meaning with respect to the overall plot (following the work of Vladimir Propp on character functions [16]) and a number of descriptions of plot lines (based on Booker [1], Tobias [18], and Polti [15]).

2.2 Existing Approaches to the Elicitation of Knowledge Pertinent to Narrative Modelling

Storytelling systems are known to require large amounts of explicit knowledge to operate successfully. Generated stories are only as good as the knowledge they have been derived from. Recent attempts have been made to address this problem via crowdsourcing [13]. In this work, a number of human authored narratives are mined to construct a *plot graph*, which models the author-intended logical flow of events in the virtual world as a set of precedence constraints between plot events [19]. When required to develop stories on a given topic that it has no knowledge of, the system is designed to send a query to Amazon Mechanical Turk (AMT) soliciting typical narratives in natural language on that topic. Such narratives are parsed and merged into a combined representation in terms of plot graph for the domain being explored, which is later used to inform the process of constructing narratives.

To inform the development of the Dramatis system for modelling suspense [14], O'Neill carried out an effort of knowledge engineering driven by methods adapted from qualitative research. The goal was to collect typical reader genre knowledge while simultaneously limiting engineer bias. The process was to acquire a corpus of natural language text and the conversion of that corpus into the knowledge structures required by Dramatis.

The Story WorkBench [6] is a free, open-source, cross-platform framework for text annotation that provides a number of common text annotation operations, including representations (e.g., tokens, sentences, parts of speech), functions (e.g., generation of initial annotations by algorithm, checking annotation validity by rule, fully manual manipulation of annotations) and tools (e.g., distributing texts to annotators via version control, merging doubly-annotated texts into a single file). It has been used extensively to build resources that capture particular aspects of narrative, such as narrative structure [5] or use of referring expressions [12], or that focused on particular topics such as Islamic terrorism [4].

3 Development of an Annotated Case-Base of Musical Theatre Plots

Very little analysis of the structure of musical theatre in terms of narrative has been carried out in the past. Talor and Simmonds [17] provides a brief description of the narrative structure of musicals in relation to the classic view of Aristotle and Vladimir Propp [16] and mentioning concepts from Booker's work [1]. However the description is sketchy and a very limited number of musicals is analysed in a manner more didactic than systematic. Woolford [20] provides a more methodical approach, based on Joseph Campbell [2], and includes detailed analysis of a small number of musicals. In both cases the size of the analised samples and the level of detail where insufficient to support the type of computational learning effort envisaged for the current project. In view of this – after initial consultation with the UCM researchers – the Wingspan Productions team organised a mass narrative annotation of musical theatre plots, the intention being that the data resulting from this could be used by the UCM researchers to further explore musical theatre narratives. Successful annotation of 42 musicals provided material for the UCM researchers to develop their narrative generation system, specifically with musicals in mind.

The research effort carried out on this initiative involved several tasks:

1. identification of a valid vocabulary of abstract units to use in annotating musical theatre plots
2. development of a spreadsheet format for annotators to use, and a corresponding set of instructions to guide them through the process
3. selection of a corpus of musical theatre pieces that would constitute the corpus to be annotated
4. carrying out the annotation process itself
5. post-processing the results of the annotation

3.1 Annotation Vocabulary

The work carried out by the UCM researchers on narrative generation [7, 8] had identified a number of valuable abstractions for the representation of narrative that were then employed in the case base of schemas described in [11] and the case base reasoning procedure described in [9]. These abstractions were based on Propp's concept of a character function as an abstract contribution by a particular character to the overall plot. However, the background research carried out to inform those processes had also identified the need for a wider range of actual character functions to consider. The set of character functions described by Propp was originally designed to capture the abstract descriptions of character's contributions to plot in the context of Russian folk tales. This is a limited domain and the set of abstractions that resulted is constrained to the basic actions that are considered valid for that genre. This presented significant differences with the domain of musical theatre. For instance, there is no mention of love affairs between characters in Russian folk tales, whereas this is a very important feature in musical theatre.

To account for these differences, a new specific set of abstractions in the spirit of character functions was defined. This set was constructed from a number of sources:

- the original set of Propp's character functions [16] was mined for abstractions general enough to be applicable across different domains
- instances of Propp's character functions that were specific to the domain of Russian folk tales were generalised to produce a set of options (for instance, Propp only considered the alternative of the hero succeeding in his struggles with the villain, whereas a character function for defeat of the hero was considered necessary to achieve the required generality)
- additional abstractions that had been identified during the exploration of the set of schemas reported in [11] were introduced as additional terms for the vocabulary (including functions like **Repentance** and **Repentance Rewarded** and a large set of abstractions to describe plot-relevant contributions related with love between characters)
- a number of abstractions specific to the description of the plot of musical theatre, such as abstractions for scenes where the plot so far is summarised (**Summary**), or where characters describe themselves and/or their point of view (**I Am What I Am**)

A more detailed description of how this process of identifying abstractions was carried out may be found in [11].

An initial set of abstractions was proposed by the UCM researchers. This initial set was reviewed, revised and extended by a group of experts in musical theatre convened by Wingspan Productions. The revision and extension was carried out keeping in mind that the set was to be used to annotate plot relevant events in pieces of musical theatre. The final set of abstractions is described in Appendix A.

For ease of reference, these labels were grouped into subsets that fell under one of 6 generic labels. This classification is described in Appendix B.

Additionally, it was decided that the term "character function" would be difficult for untrained annotators to relate to, and the term "plot element" was chosen instead to describe these labelled abstractions of character's contributions to plot.

3.2 Annotation Tool and Annotation Instructions

The set of abstractions (from now on referred to as *plot elements*) so obtained was used to build an annotational tool in the form of a spreadsheet that annotators could use to produce a standard register of an annotated musical.

This involved making available the set of plot elements paired with brief descriptions of their meaning. It also allowed for inclusion of the synopsis of the plot of the musical that was to be annotated, so that annotations could be made with respect to a canonical reference (specific abstractions are to be assigned to spans of text in the synopsis).

The tool provided a specific spreadsheet that allowed for the following operations:

- identifying the span of the synopsis covered by a particular plot element (by indicating the start and end of the text span in question)
- selecting which label out of the vocabulary to assign to that plot element (this was provided as two drop down lists, one to select the subset of the labels and to select which label out of that subset to assign)
- annotating whether that plot element corresponded to a musical number, and, if so, what was its title
- declaring an additional label for plot elements not covered by the vocabulary provided
- the possibility of indicating, for each member of the cast, which emotions that character experienced in the context of the annotated plot element, and whether the character sang or danced
- indicating which emotions where prevalent in the context of the annotated plot element for the set of characters involved overall

The set of emotions that were chosen to annotate where: Love, Hate, Danger, Violence, Happiness, Sadness, Humour, Surprise, and Fear.

Additionally, a plot element called OTHER was introduced for annotators to use whenever they reckoned that a new label was needed, not covered by the ones in the list, but fundamental for annotating a particular plot.

A screenshot of the basic spreadsheet for annotation is shown in Figure 1.

3.2.1 Selection of the Corpus to Annotate

The corpus to annotate was compiled by the team at Wingspan Productions based on a list produced by an earlier analysis carried out at the University of Cambridge (categorising musicals by length of run and numbers of awards won).

The musicals listed in Table 1 were annotated.

3.3 Annotation Process

The annotation process was carried out at the University of Surrey, supervised by Julian Woolford, Programme Leader for the MA in Musical Theatre and the team from Wingspan Productions.

The 35 annotators were predominantly Musical Theatre MA (Masters) students, ranging in age from 20–30; male/female 40/60%. All participants will have completed secondary

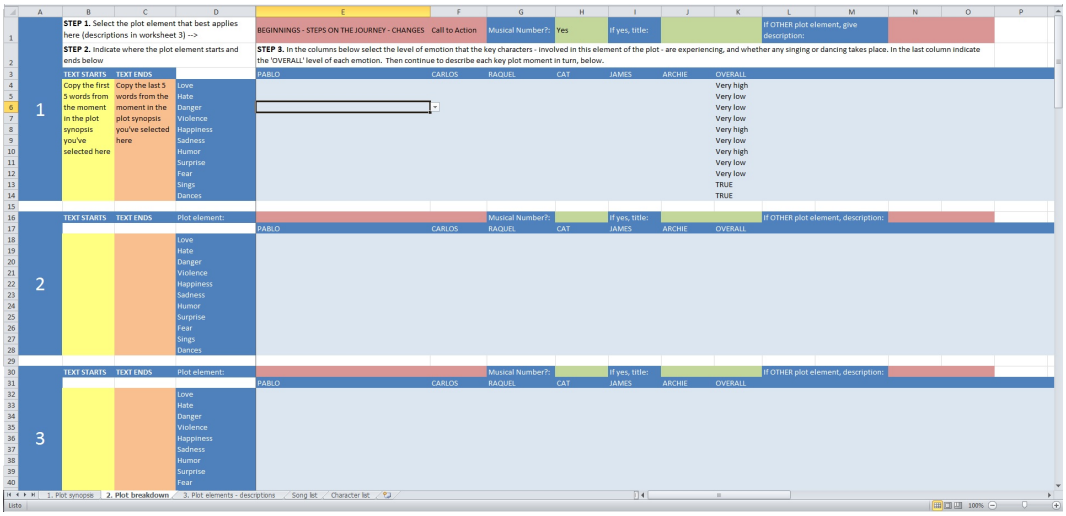

Figure 1 Screenshot of the annotation spreadsheet.

education and the majority also a first degree in a related subject. Two members of academic staff from University of Surrey also participated.

Annotators were offered a long list of musicals to choose from, prior to the session, to ensure they were familiar with the musical they annotated. If they enjoyed the first session they were then invited to choose a second musical to annotate (that they were familiar with) which was completed at home.

Annotators were provided with the instruction sheet prepared initially by Raquel Hervás, modified slightly for the group annotation scenario. They were emailed this the day before, to read as preparation, and were then provided with a hard copy of the instructions on the day. All participants were briefed verbally at the start of the session, and three people were on hand to answer questions and provide support as necessary.

An additional small set of annotations were completed by members of the Wingspan Productions team, completed during development of the methodology.

4 Discussion of Results

Results are discussed from two different points of view: observations on how the annotation process took place, and observations arising from the postprocessing of the annotated data.

4.1 Discussion of the Annotation Process

The majority of people took around 90 minutes to complete the annotation of a single musical (ranging from ca. 45 mins to 2 hours). Those that completed the process quickly did not turn in incomplete annotations/use fewer plot elements than others – they simply seemed to have genuine aptitude for the task, and were very comfortable thinking about the shows they annotated in the terms the process required.

From the annotations session, it was clear that familiarity with the musical being annotated is an enormous help. Having a keen knowledge of the structure of a musical (particularly how it is performed, in addition to how it appears on the page) makes it much more straightforward to break it down into coherent units, expressed as plot elements.

Table 1 Musicals annotated.

A Little Night Music	Oliver
Amour	Miss Saigon
Annie	Oklahoma!
Avenue Q	Parade
Barnum	Passion
Beauty and the Beast	Rent
Billy Elliot	Pippin
Brass	South Pacific
Cabaret	Spring Awakening
Carousel	State Fair
Chicago	Sunday In The Park With George
The Fields of Ambrosia	Sweeney Todd
Grease	The Book of Mormon
Gypsy	The Lion King
Jekyll and Hyde	The Pirate Queen
Jesus Christ Superstar	The Rocky Horror Show
Joseph and the Amazing Technicolor Dreamcoat	The Sound Of Music
La Cage Aux Folles	Thoroughly Modern Millie
Mary Poppins	Urinetown
Matilda	West Side Story
Memphis	Wicked

The process most people adopted was to read the synopsis first (as advised), then proceed with the annotation alternating between the main spreadsheet and the list of plot element definitions. Most annotators found the given instructions to be clear, with only a small number requiring significant help. Help tended to be required where people were unfamiliar with using Microsoft Excel.

The chosen platform used in the annotation raised some minor technical problems. For instance, there were some problems with compatibility of the spreadsheet across different operating systems, although these could be resolved on the day. Having predicted that all user would have Microsoft Office installed in their devices turned out to be false, and several ad-hoc solutions had to be taken in order to have every subject complete their annotation(s). With this in mind, an online form with cross-browser compatibility could be much more straightforward both for annotation and data collection.

Annotators frequently missed fields – a check was done as each person finished, but some gaps still went unnoticed. From questions asked by the annotators it is apparent that there was some doubt as to whether the emotion to be annotated was the emotion attributed to the character or the emotion felt by the spectator. Annotators needed frequent reminding to add in information about whether a song happened during particular plot elements.

An interesting observation made by one of the volunteers helping with the annotation process concerns the fact that different people seemed to interpret the plot of a musical in different ways, so that what is three plot elements to one person, is a single plot element to another.

4.2 Discussion of the Annotated Data

The volume of data obtained is substantially large, and only preliminary analyses of it have been carried out. Nevertheless, these reveal some interesting conclusions. Exhaustive listing of all the data would exceed the available space in this kind of paper so an attempt has been made to summarise features that have been considered significant.

Regarding the set of plot elements included in the vocabulary, there was generic concern among the researchers that developed the vocabulary that either some elements may not be employed at all or that some elements crucial for capturing the essence of the plots in question might be missing.

Of the whole set of 116 plot elements in the vocabulary only 6 were never used by evaluators in the annotation of the 42 plots that were annotated. The plot elements in question were: **Rescue from pursuit**, **Erroneous Judgement**, **Involuntary Crimes of Love**, **The Enigma**, **Unfounded Claims**, **Riches**. Of these, **Rescue from pursuit** – originating from Propp's original set of character functions and retained because it more explicitly linked up with the **Pursuit** plot element – may have been displaced by the more intuitive **Rescue** – which was introduced as a generalization and is actually used twice in the annotated plots. The **Involuntary Crimes of Love** plot element originates in Polti's dramatic situations, and has significant overlap with more explicit plot elements such as **Adultery** or **Murderous Adultery**. Plot element **The Enigma** arises from one of Tobias' possible plots, and it is very abstract. Its full description concerns plots that have an underlying mystery that must be solved in parallel by the protagonist and the reader – as in thrillers, mystery novels, and whodunnits. It probably had no role to play as part of a vocabulary for annotating musicals, and it was in any case poorly described in the material provided to annotators. The **Unfounded Claims** plot element is also inherited from Propp's character functions, but may also have been poorly described in the material, so that annotators would be unlikely to choose it.

The remaining 110 plot elements provided were used to label 992 segments of musical theatre plots, giving an average of 9 appearances per plot element. However, observed use differed significantly from this average.

Table 2 shows the most frequently used plot elements in each category, with their associated frequency of use. The most frequently used plot element was **Decision to Take Action** which appears 50 times in the annotated musicals. Five of the plot elements appear more than 20 times: **Aspiration**, **Boy Meets Girl**, **Tested**, **Character's Reaction** and **Bond Strengthened**. In the range of between 10 to 20 appearances we find 22 plot elements. Of these, 8 relate to love (**Class differences**, **Love shift**, **Love Triangle**, **Obstacles to Love**, **Couple Wants to Marry**, **Forbidden Love**, **Reconciliation** and **One-sided Love**); 2 to initiating events (**Call to Action** and **Departure**); 5 to conflicts that require resolution (**Difficult task**, **Imprisoned**, **Struggle**, **Loss of Loved Ones**, **Revenge**); 5 to possible developments from earlier situations (**Reward**, **Solution**, **Exposure**, **Discovery**, **Epiphany**); and 2 cases of help provided to a character (**Guidance** and **Assistance**). Overall, these details constitute an interesting sample of the ingredients of musical theatre as a genre.

There were three plot elements in the vocabulary provided that were deemed important to make it possible to annotate musical plots even though they had not been suggested by any of the sources studied on plot. They refer to either references to the plot itself or references to the annotation vocabulary. Two plot elements were intended to help mark parts of the plot designed to help the audience either by introducing an initial situation (**Initial situation** which was used 45 times in the annotation) or by summarising prior parts of the

Table 2 Extract of the most frequently used plot elements in each category, with their associated frequency of use.

BEGINNINGS – STEPS ON THE JOURNEY – CHANGES	
Aspiration	20
Epiphany	18
Departure	15
Call to Action	13
Discovery	12
Guidance	11
BIG DECISIONS - CHOICES	
Decision to Take Action	50
Character's Reaction	23
TESTS, TRIAL AND ORDEALS	
Tested	21
Loss of Loved Ones	17
Difficult task	14
Imprisoned	14
Exposure	11

TRIALS AND TRIBULATIONS OF LOVE – FRIENDSHIP	
Boy Meets Girl	20
One-sided Love	15
Forbidden Love	13
Couple Wants to Marry	12
Obstacles to Love	11
Class differences	10
Love shift	10
Love Triangle	10
CONFLICT	
Bond Strengthened	30
Assistance	15
Reconciliation	13
Solution	10
Reward	10

plot at a given stage (**Summary** which was used 34 times in the annotation). The results support the impression that these plot elements are useful in describing musical theatre plots.

The OTHER plot element was introduced to cover the possibility that some elements crucial for capturing the essence of the plots in question might be missing. Annotators were advised to resort to this label when none of the plot elements provided covered the essence of the fragment they were trying to annotate. The OTHER label was used 19 times. A large variation in usage is observed:

- A small number of annotators used the OTHER field where they felt they could not find a fitting plot element. It is unclear if these choices genuinely do represent plot elements missing from the list provided, or whether the OTHER field was used for ease
- Some people used the OTHER field in addition to choosing a primary plot element, presumably because two plot elements together enabled them to express finer-grained detail about what was happening at that point in the story
- On occasion the OTHER field was used to provide additional non plot-related information (e.g. a diegetic song)

On a related note, at least in one occasion the generic label for a subgroup of plot elements was used to annotate instead of a particular member of the group. Again, it is unclear whether this is due to inadequacy of the lower level plot elements or for simplicity.

Working from insights obtained in the development of story generators based on Propp's character functions [7, 8], it was expected that some of the plot elements introduced in the vocabulary would show some evidence of cross-dependency. This would affect cases like **Aspiration** and **Aspiration Achieved** or **Misunderstanding Arises** and **Misunderstanding Cleared**. While a detailed analysis of this type of relation in individual annotated plots is still pending, a rough analysis of the statistical data collected shows that this expectation is likely to be unfulfilled. Relative frequency for a number of such pairings between plot elements are shown in Table 3.

Along similar lines, Table 4 compiles statistics on the relative frequency between pairs of plot elements that represent different alternatives for resolving a given initial situation. It might seem from these data that, overall, characters in the annotated musicals seem to

Table 3 Relative frequency of appearance of hypothetically related pairs of plot elements.

Misunderstanding Arises	4
Misunderstanding Cleared	1
Aspiration	20
Aspiration Achieved	7
Couple Wants to Marry	12
Wedding	2
Forbidding/Warning	3
Warning/Forbidding Disregarded	2
Lack	7
Lack Fulfilled	2

Table 4 Relative frequency of appearance of related plot elements to correspond to alternative solutions.

Succumbing to Temptation	9
Temptation Resisted	1
Moral Dilemma Triumph	6
Moral Dilemma Failure	4
Boy Meets Girl	20
Boy Loses Girl	7
Couple Wants to Marry	12
Wedding	2

succumb to temptation more often than resist it and triumph more often than fail when faced with a moral dilemma. Also that boys that meet girls do not often lose them, and that couples that want to marry do not always succeed. However, more detailed analysis of the individual annotated plots is required. First, to ensure that these pairings are actually present simultaneously in individual plots. Second, to double check that the expected resolution is not actually implied by some other plot element that may describe the same solution in different terms.

With respect to the annotations on emotions that were collected, it has only been possible to carry out a quick processing of the application of emotional labels to the plot elements as a whole. Labels on emotional content were also collected for each individual character involved in a plot element, but these data have not been processed yet.

The assignment of emotional labels to plot elements as used in the annotation in the context of given musical plots shows little indication that particular emotional labels may be consistently applied to specific plot elements. The annotations collected have been processed with the ultimate goal of identifying cases where similar values for the emotional features considered have been consistently applied to specific plot elements across all their appearances in the set of annotated plots. To achieve this, the standard deviation over the values assigned to the plot element for each emotional feature has been computed. Values of this standard deviation below 0.5 have been taken to be probable indication of consistency across appearances. This occurs only infrequently. Where values are assigned consistently, only cases where either low or high values of some of the emotional features are attributed to a plot element are considered. The resulting data are compiled in Table 5.

Overall, the data on emotion assignment to plot elements shows few surprises. The **Deception to Fit In** plot element shows low values of both happiness and sadness, and

Table 5 Plot elements consistently assigned significant emotional values across the annotated set of plots.

Emotional feature	Plot elements
Love+	Couple Wants to Marry
	Maturation
	Pursuit
	Recovery of a Lost One
	Rescue
	Wedding
Love-	Abduction
	Madness
	Warning/Forbidding Disregarded
Hate+	Cross-Rank Rivalry
Hate-	Cross-Dressing
	Deception to Fit In
	Deliverance
	Return
	Reward
	Victory
	Wedding
Danger+	Abduction
	Misfortune
Danger-	Adultery
	Boy Loses Girl
	Moral Dilemma Failure
	Rescue
	Victory
	Wedding
Sadness+	Abduction
Sadness-	Couple Wants to Marry
	Deception to Fit In
	Rescue
	Wedding
Humour-	Abduction
	Moral Dilemma Failure
	Shame of Loved One
	Warning/Forbidding Disregarded

Emotional feature	Plot elements
Violence+	Abduction
Violence-	Aspiration Achieved
	Couple Wants to Marry
	Disguise
	High Status Revealed
	Love shift
	Maturation
	Obstacles to Love
	High Status Revealed
	Pursuit
	Recovery of a Lost One
	Reward
	Victory
	Wedding
Fear+	Shame of Loved One
Fear-	Adultery
	Ambition
	Aspiration Achieved
	Deception
	Moral Dilemma Failure
	Reward
	Aspiration Achieved
Happiness+	Aspiration Achieved
	Couple Wants to Marry
	Lack Fulfilled
	Pursuit
	Wedding
Happiness-	Cross-Rank Rivalry
	Deception to Fit In
	Lack
	Shame of Loved One

the **Pursuit** plot element shows high values of happiness. More interesting is the relation between some emotions for certain plot elements. For example, the set of plot elements **Reward**, **Victory**, and **Wedding** presents low values of hate and violence and, in the case of **Victory** and **Wedding**, low values of danger. Other dependencies appear for **Couple Wants to Marry** and **Wedding**, both having low values of sadness and high values of happiness and love, and **Abduction** showing high values of danger and violence. More studies will be performed automatically in order to find more dependencies like these ones.

The low level of consistency in attribution of emotional features to specific plot elements may be the result of having insufficient data. Larger volumes of data would have to be collected to obtain significant results. The current annotation suffers from two disadvantages: each musical has been annotated by a different person, so a single subjective view is available for each musical, and there are insufficient occurrences of most of the plot elements to provide a significant set of values.

5 Conclusions

The annotation reported in this paper was carried out to inform the development of a generator of musical theatre plots. The ultimate aim of the annotation was to provide a reference corpus to test the validity of a Proppian approach to the generation of musical theatre plots, supported by a limited process of adaptation to the new domain. The generator for musical plots had to be ready to be used in the composition of a musical that had to open in the London West End on February 2016. It therefore had tight deadlines to meet. Although every effort was made to apply academic rigour wherever possible, constraints, both on time and resources, made it imperative to reduce both the number of musicals that could be annotated and the amount of annotations that could be collected for them. Nevertheless, the corpus was ready on time and served to inform the development of the PropperWryter software – a generator of plot lines for musical theatre pieces – which was used for the composition of "Beyond the Fence", the first ever experimental computer-generated musical [3]. "Beyond the Fence" opened successfully at the Arts Theatre in London on 22nd February 2016 and has by now enjoyed a successful two-week run.

The present paper reports on the knowledge resources that were developed to inform that process. The team that worked on the initial annotation is currently applying for funding to carry out a more refined annotation, over a larger corpus, with a view to improving the results obtained so far.

References

1 C. Booker. *The Seven Basic Plots: Why We Tell Stories*. The Seven Basic Plots: Why We Tell Stories. Continuum, 2004.

2 J. Campbell, P. Cousineau, and S. L. Brown. *The Hero's Journey: Joseph Campbell on His Life and Work*. Collected works of Joseph Campbell. New World Library, 1990.

3 Simon Colton, M. T. Llano, R. Hepworth, J. Charnley, C. V. Gale, A. Baron, F. Pachet, P. Roy, P. Gervas, N. Collins, B. Sturm, T. Weyde, D. Wolff, and J. Lloyd. The beyond the fence musical and computer says show documentary. In *Proceedings of the Seventh International Conference on Computational Creativity, ICCC 2016, Paris, 27 June – 1 July 2016.*, 2016.

4 M. A. Finlayson, J. R. Halverson, and S. R. Corman. The N2 corpus: A semantically annotated collection of islamist extremist stories. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014), Reykjavik, Iceland, May 26-31, 2014.*, pages 896–902, 2014. URL: `http://www.lrec-conf.org/proceedings/lrec2014/summaries/48.html`.

5 M.A. Finlayson. *Learning Narrative Structure from Annotated Folktales*. PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, 2011.

6 M.A. Finlayson. The story workbench: An extensible semi-automatic text annotation tool. In *Intelligent Narrative Technologies IV, Papers from the 2011 AIIDE Workshop, Stanford, California, USA, October 10-11, 2011*, 2011. URL: `http://aaai.org/ocs/index.php/AIIDE/AIIDE11WS/paper/view/4091`.

7 P. Gervás. Reviewing Propp's Story Generation Procedure in the Light of Computational Creativity. In *AISB Symposium on Computational Creativity, AISB-2014, April 1-4 2014*, Goldsmiths, London, UK, 2014.

8 P. Gervás. Computational Drafting of Plot Structures for Russian Folk Tales. *Cognitive Computation*, pages 1–17, 2015. `doi:10.1007/s12559-015-9338-8`.

9 P. Gervás, R. Hervás, and C. León. Generating Plots for a Given Query Using a Case-Base of Narrative Schemas. In *Proceedings of the Experience and Creativity Workshop, ICCBR*, 2015.

10 P. Gervás and C. León. The need for multi-aspectual representation of narratives in modelling their creative process. In *5th Workshop on Computational Models of Narrative*, OASIcs-OpenAccess Series in Informatics, 2014.

11 P. Gervás, C. León, and G. Méndez. Schemas for Narrative Generation Mined from Existing Descriptions of Plot. In M.A. Finlayson, B. Miller, A. Lieto, and R. Ronfard, editors, *6th Workshop on Computational Models of Narrative (CMN 2015)*, pages 54–71, Dagstuhl, Germany, 2015. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik. doi:10.4230/OASIcs.CMN.2015.54.

12 R. Hervás and M. Finlayson. The prevalence of descriptive referring expressions in news and narrative. In *48th Annual Meeting of the Association for Computational Linguistics (ACL 2010)*, Uppsala, Sweden, 2010.

13 Boyang Li, S. Lee-Urban, G. Johnston, and M. O. Riedl. Story generation with crowdsourced plot graphs. In *Proceedings of the 27th AAAI Conference on Artificial Intelligence*, AAAI'13, 2013.

14 B. O'Neill. *A Computational Model of Suspense for the Augmentation of Intelligent Story Generation*. PhD thesis, Georgia Institute of Technology, Atlanta, Georgia, 2013.

15 G. Polti and L. Ray. *The Thirty-six Dramatic Situations*. Editor Company, 1916.

16 V. Propp. *Morphology of the Folktale*. University of Texas Press, 1968.

17 M. Taylor and D. Symonds. *Studying Musical Theatre: Theory and Practice*. Palgrave Macmillan, 2014.

18 R. B. Tobias. *20 Master Plots: And How to Build Them*. F+W Media, 2012.

19 P. Weyhrauch. *Guiding interactive drama*. PhD thesis, Carnegie Mellon University, Pittsburgh, PA, 1997.

20 J. Woolford. *How Musicals Work: And how to Write Your Own*. A Nick Hern book. Nick Hern Books, 2012.

A Appendix A

This appendix presents the full list of abstractions proposed as plot elements, with their accompanying descriptions as presented to the annotators.

Abduction a character is abducted
Adultery two characters commit adultery, deceiving a partner
Ambition character tries to obtain/achieve the possession or status of another
An Enemy Loved one of two lovers is hated by a friend(s) of the other
Aspiration Achieved an aspiration held earlier the story is achieved
Aspiration a member of a group wants or aspires to have, or be, something better
Assistance a character assists another to achieve their goals
Bond Strengthened the bond between two characters is strengthened
Boy Loses Girl two characters that were emotionally attached drift apart
Boy Meets Girl two characters meet and become emotionally attached
Branding a character acquires a mark or a token by which they will later be recognised
Call to Action faced with a misfortune/lack/aspiration, a character is exhorted to take action
Character Flaw a character's misfortune is caused by their own character flaws
Character's Reaction a character reacts to the actions of another character testing them
Class differences relationship impossible due to class differences between partners

Complicity a character colaborates with an enemy, deceiving their friends
Conflict with a God a mortal struggles with a deity
Couple Wants to Marry two characters want to marry
Crimes of Love a lover and beloved incur in questionable acts
Cross-Dressing a character dresses up as one of the opposite sex
Cross-Rank Rivalry two masculine or feminine rivals with different rank
Daring Enterprise character attempts to recover an object or person from an adversary
Deception to Fit In character pretends to be something they're not, in order to fit in
Deception character is deceived, unwittingly helping an enemy
Decision to Take Action a character agrees to or decides to take action
Defeat focal character is defeated by another
Deliverance protector comes to the rescue of the distressed
Departure character leaves the place where they were, the story follows them
Difficult task a difficult task is proposed to a character
Disaster someone is defeated by an enemy or catastrophe
Disconnected from Reality a character becomes disconnected from the reality of their surroundings (either due to overwhelming love, grief, dreaming, flashbacks, etc.)
Discovery character discovers/realises/learns something about themselves
Disguise a character puts on a disguise
Epiphany a character has a significant moment of realization
Erroneous Judgement any kind of mistaken judgement
Escape character that was imprisoned or abducted escapes by their own means
Exposure a character operating under deception is exposed
Forbidden Love unconventional love relation
Forbidding/Warning a character is forbidden from or warned against a specific course of action
Guidance a character is guided towards a new possession or goal
Hatred between Friends characters that were once friends become enemies
High Status Revealed high status of one character is revealed, opening new possibilities
I am what I am character defiant about who they are/their identity
Ill-fated Imprudence recklessness or curiosity that results in loss
Imprisoned a character is confined against their will
Inconstancy one partner is inconstant in their love of the other
Initial situation description of initial situation
Involuntary Crimes of Love character unknowingly commits adultery or incest
Jealousy a character is justifiably jealous of another
Judgement Deferred to Authority conflict between characters to be solved by a decision of someone in authority
Lack Fulfilled an important lack suffered by a character earlier in the story is fulfilled
Lack a member of a group lacks something important in their life
Lesson Learned a character realizes they were wrong because of an event
Loss of Loved Ones a character experiences the loss of a loved one
Love Triangle a third individual is involved in a love relationship
Love shift person that loved one character shifts their love to another character
Madness a madman slays, injures or brings disgrace onto a character
Maturation immature character evolves to maturity, usually as a result of a challenging incident
Metamorphosis character transformed into a beast (literal/metaphorical)

Misfortune character suffers a cruel master or misfortune
Mistaken Jealousy a character is mistakenly jealous of another
Mistaken murder character unwittingly kills a friend
Misunderstanding Arises a misunderstanding arises that constitutes an obstacle to the intentions of some characters
Misunderstanding Cleared a misunderstanding that constituted an obstacle is cleared
Moral Dilemma Failure character faces a moral dilemma and makes the wrong decision
Moral Dilemma Triumph character faces a moral dilemma and makes the right decision
Murderous Adultery a betrayed husband or wife kills one or both adulterers
Obstacles to Love relationship prevented by social norms
One-sided Love character loves another but is not loved back
Parent Convinced parent or guardian that opposed a relationship is now convinced
Persuasion character has to obtain something through eloquence and diplomacy
Poverty character suffers extreme poverty
Provision of a Magical Agent a character acquires the use of a magical agent
Punishment a character is punished
Pursuit a character is pursued
Recognition a character that was in disguise is recognized
Reconciliation two characters that had been at odds get back together
Reconnaissance a character makes an attempt at reconnaissance
Recovery of a Lost One a character recovers a loved one the had lost
Remorse a culprit suffers remorse for a crime or love fault
Repentance Rewarded a character that has repented is rewarded
Repentance a character that has behaved badly repents
Rescue one character rescues another character imprisoned against their will
Rescue from pursuit a character that was being pursued manages to get away
Return a characters returns to to a place where they had been earlier in the story
Revenge character enacts revenge
Revolt character conspires to revolt against a tyrant
Reward character receives a reward
Riches character achieves material fortune
Rivalry a character and an antagonist of balanced power clash
Sacrifice for Family character makes sacrifices for happiness of a relative
Sacrifice for Passion character makes sacrifices for a vice or passion
Sacrifice for an Ideal character sacrifices life, love or well-being to a cause
Sacrifice of Loved Ones character sacrifices a loved one for a necessity or vow
Shame of Loved One a character discovers the shame of a loved one
Solution a task is resolved
Someone Leaves one of the members of a group goes away, the story follows those who remain
Struggle two characters join in direct combat or open confrontation
Succumbing to Temptation a character succumbs to temptation
Summary scene that summarises a particular moment in the story, though nothing much actually happens
Temptation Resisted a character resists temptation
Tested a character is tested, interrogated, attacked, asked to perform a service...
The Enigma a combat of the intelligence to find a person or object
Transfiguration a character is given a new appearance

Transformation because of a crisis, a character's behaviour or sense of self is transformed
Trickery a character attempts to deceive another to get something they want
Underdog a character and an antagonist of balanced power clash, and the character is at disadvantage
Unfounded Claims a character presents unfounded claims
Unrecognized Arrival unrecognized, a character arrives in a new place
Unrelenting Guardian parent or guardian of one partner opposes a desired relationship
Useful Information character receives information relevant to their intentions
Victory focal character achieves victory over another
Villainy a character causes harm or injury to another character
Warning/Forbidding Disregarded a character disregards a warning or a prohibition received earlier in the story
Wedding a character is married

B Appendix B

This appendix presents the way in which the plot elements were grouped into subsets that fell under one of 6 generic labels:

BEGINNINGS – STEPS ON THE JOURNEY – CHANGES Initial situation, Summary, Aspiration, Call to Action, Cross-Dressing, Departure, Deliverance, Disconnected from Reality, Discovery, Disguise, Epiphany, Escape, Guidance, High Status Revealed, Maturation, Metamorphosis, Pursuit, Reconnaissance, Rescue, Return, Someone Leaves, Transfiguration, Transformation, Unrecognized Arrival

BIG DECISIONS – CHOICES Character's Reaction, Decision to Take Action, Deception to Fit In, Erroneous Judgement, Ill-fated Imprudence, Moral Dilemma Triumph, Moral Dilemma Failure, Mistaken Jealousy, Sacrifice for an Ideal, Sacrifice for Family, Sacrifice for Passion, Sacrifice of Loved Ones, Succumbing to Temptation, Temptation Resisted, Warning/Forbidding Disregarded, Character Flaw

TRIALS AND TRIBULATIONS OF LOVE – FRIENDSHIP Boy Meets Girl, Boy Loses Girl, Wedding, Class differences, Forbidden Love, Inconstancy, Involuntary Crimes of Love, Adultery, An Enemy Loved, Crimes of Love, Love shift, Love Triangle, Murderous Adultery, One-sided Love, Obstacles to Love, Parent Convinced, Recovery of a Lost One, Couple Wants to Marry

TESTS, TRIAL AND ORDEALS Abduction, Branding, Deception, Difficult task, Disaster, Shame of Loved One, Exposure, Forbidding/Warning, Mistaken murder, Lack, Loss of Loved Ones, Madness, Misfortune, Persuasion, Poverty, Punishment, Recognition, Remorse, Tested, The Enigma, Defeat, Villainy, Imprisoned, Lesson Learned

CONFLICT Ambition, I am what I am, Complicity, Conflict with a God, Cross-Rank Rivalry, Daring Enterprise, Hatred between Friends, Jealousy, Misunderstanding Arises, Revenge, Revolt, Rivalry, Struggle, Judgement Deferred to Authority, Trickery, Underdog, Unfounded Claims, Unrelenting Guardian

HELP – REWARD – RESOLUTION Assistance, Bond Strengthened, Useful Information, Lack Fulfilled, Aspiration Achieved, Provision of a Magical Agent, Repentance Rewarded, Reward, Riches, Victory, Misunderstanding Cleared, Reconciliation, Repentance, Solution