Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence (IJCAI-22)
Special Track on AI, the Arts and Creativity

Automated Sifting of Stories from Simulated Storyworlds

Wilkins Leong[*], **Julie Porteous**, **John Thangarajah**
RMIT University
leong.wilkins@gmail.com, {julie.porteous, john.thangarajah}@rmit.edu.au

Abstract

Story sifting (or *story recognition*) allows for the exploration of events, stories, and patterns that emerge from simulated storyworlds. The goal of this work is to reduce the authoring burden for creating sifting queries. In this paper, we use the event traces of simulated storyworlds to create Dynamic Character Networks that track the changing relationship scores between characters in a simulation. These networks allow for the fortunes between any two characters to be plotted against time as a story arc. Similarity scores between story arcs from the simulation and a user's query arc can be calculated using the Dynamic Time Warping algorithm. Events corresponding to the story arc that best matches the query arc can then be returned to the user, thus providing an intuitive means for users to sift a variety of stories without coding a search query. These components are implemented in our experimental prototype ARC SIFT. The results of a user study support our expectation that ARC SIFT is an intuitive and accurate tool that allows human users to sift stories out from a larger chronicle of events emerging from a simulated storyworld.

1 Introduction

Interactive narrative systems feature virtual characters whose actions and behaviours play out according to system defined criteria appropriate for the narrative domain of interest. Examples of such systems can be found in areas such as entertainment [Mateas and Stern, 2005], education [Porteous *et al.*, 2017], games [Grinblat and Bucklew, 2017] and healthcare [Lindsay *et al.*, 2015]. Approaches to generating narratives in these systems differ with respect to whether the narrative is determined by a *top-down* narrative structure, the plot, or whether it emerges, *bottom-up*, from the interactions of a population of virtual characters. For the top-down approach, techniques such as AI planning [Riedl and Young, 2010] and Bayesian networks [Mott and Lester, 2006] have been successfully applied but efforts to control narrative structure can impede generative possibilities of the approach. In contrast, bottom-up agent-based simulations offer much greater potential for the emergence of a wide range of interesting narratives, [Grinblat and Bucklew, 2017; Kreminski *et al.*, 2019]. The challenge is how to ensure that such narratives, and behaviours of characters in them, adhere to desired criteria.

Recent work in the area of *story sifting* [Ryan *et al.*, 2015; Kreminski *et al.*, 2019] has introduced new approaches to addressing this challenge. The problem of story sifting can be seen as how to select event sequences and patterns from story simulations, i.e. stories, that adhere to specified criteria that are appropriate for the domain of interest. Note that the criteria will differ across domains, for example an entertainment application's criteria might relate to aesthetic properties of the genre such as tension or suspense [Cheong and Young, 2015], whereas criteria for a healthcare narrative might relate to compliance with required protocols [Lindsay *et al.*, 2015]. Thus, story sifting offers the potential to control the exploration of large spaces to extract specific desirable events and patterns that constitute stories which emerge from the simulation.

A limitation of current approaches to story sifting is that they tend to require manual specification of search queries, e.g. [Behrooz *et al.*, 2015; Kreminski *et al.*, 2019; Ryan, 2018], in a technical language which can be burdensome for non-technical users. Hence, our motivation in this work was to reduce the authoring burden for sifting stories by providing an intuitive and user-friendly method for describing desired story criteria and an automated approach to sift stories corresponding to those criteria. We observed that the use of visual representations of stories, i.e. a *story arc* [Weiland, 2016], would allow users to describe story criteria by drawing the "shape" of stories of interest. A story arc can be thought of as plotting the changes of character fortunes "for better or worse, over the course of the telling" [McKee, 1997]. The use of such a visual representation would reduce the authoring overhead by removing the need for technical knowledge. The problem then is how to represent sequences of event traces from simulated storyworlds which match these story arcs, i.e. the specified criteria.

Our solution to this problem, which we have implemented in a tool we refer to as ARC SIFT, consists of two parts. Firstly, the creation of *Dynamic Character Networks* [Lee and Jung, 2020] that track the changing relationships between

[*]Contact Author

characters over the course of a simulation. These networks allow for the relationships between characters in the simulation to be plotted against time as *simulation* story arcs. Secondly, the use of the time-series comparison algorithm *Dynamic Time Warping* [Müller, 2007], to calculate similarity scores between simulation story arcs and *query* story arcs (e.g. arcs drawn by users). ARC SIFT returns stories (i.e. sequences of events) corresponding to the simulation story arcs that best match query arcs.

We hypothesize that ARC SIFT is an intuitive (easy for users to understand how to use it) and accurate (able to sift a specific story of interest) tool for allowing users to sift stories without the need to author technical search queries. We demonstrate these two properties in a user study, the details of which are presented in Section 4.

2 Background & Related Work

Emergent Narrative is a concept that has been used to describe the "bottom-up" generation of narrative from the interactions between autonomous virtual characters in a simulated storyworld such as a multi-agent simulation. [Riedl and Bulitko, 2013; Louchart *et al.*, 2015]. The strengths of such approaches include their flexibility, generative power and scalability. However, the downside is the potential size of narrative possibilities emerging from such simulations, and the problem of identifying narratives from within, that correspond to desired criteria. Recent work has re-cast this in terms of narrative "curation" and the endeavour to "sift" stories adhering to criteria of interest from the simulation [Ryan *et al.*, 2015; Ryan, 2018]. This is further related to work in "re-telling": the a posteriori creation of stories based on logs of interactive experiences with games and simulations [Behrooz *et al.*, 2015].

This recent interest has led to the development of a number of story sifting systems which differ in their approach [Kreminski *et al.*, 2019; Osborn *et al.*, 2015; Ryan, 2018; Behrooz *et al.*, 2015; Grinblat and Bucklew, 2017]. Many of these systems share a need for sifting queries to be authored using some form of technical language. For instance: Kreminski et al's FELT [Kreminski *et al.*, 2019] introduced a bespoke language for specifying queries to sift stories from a rule-based simulation; and with "Sheldon County", Ryan [Ryan, 2018] used pre-defined chunks of procedural Python code to sift from a simulation.

In recognition of the authoring overhead required for story sifting in such systems, attempts have been made to ease this, such as the interactive example-driven synthesizer that generates sifting patterns introduced in [Kreminski *et al.*, 2020].

Our ARC SIFT approach shares a similar philosophy and aims to reduce the authoring overhead but differs with respect to focus: ours is development of an intuitive method for specifying story criteria of interest via visual representations. It features the use of *Dynamic Character Networks* and *Dynamic Time Warping*. Dynamic character networks [Lee and Jung, 2020], represent and track temporal changes in social affinity that result from interactions between characters in movies. In our approach we generalise this to tracking the positive and negative consequences of interactions with other characters over the course of a virtual simulation. Dynamic Time Warping is an algorithm which can be used to compare the similarity between time-dependant sequences, while allowing for elastic shifting in the time axis. It is useful for comparing sequences that have similar shapes but don't necessarily align along the time axis. We use this to find the best match for an input query story arc from a collection of simulation story arcs [Müller, 2007].

Visual representations, in the form of story arcs, have been used for narrative but in very different contexts to ours. For example: Porteous et al. [Porteous *et al.*, 2011] used a visual representation for the specification of narrative properties within their AI planning based approach to narrative generation; Kybartas et al. [Kybartas *et al.*, 2021] proposed an approach to narrative content authoring using a sketch-based interface via tension space analysis but this did not explicitly feature narrative arcs; and the approach of [Rhodes *et al.*, 2010] developed dramatic devices based on Freytag's Pyramid for sports commentary generation.

3 Approach

Shapes have been used to describe stories for centuries. Early ideas include the Aristotelian arc and Freytag's pyramid to describe the three act story structure, which are now widely used today. Shapes can also be used to describe a central character's emotions, or fortunes. Kurt Vonnegut proposed the idea that stories surrounding a central character have shapes, and can be charted on a "Beginning-End" and "Ill Fortune–Great Fortune" axis. This has been widely referred to in the writing industry as a *story arc*[1] [McKee, 1997; Weiland, 2016].

This concept can also be applied to virtual characters in a simulated storyworld, where a *story arc* can be seen as a time series plot describing the shape of the change of a single character's "fortunes" over the course of the story or simulation. Fortunes are simulation specific, but can be viewed as the negative and positive consequences of interactions between a character and others.

These consequences may relate to a character's physiological needs, safety, esteem, self-actualization, transcendence or love and belonging. For example in fire evacuation simulation, the main character's focus is to find lasting safety, where the story arc will rise and fall according to how close the main character is to reaching safety. Similarly, in a romantic storyworld, the story arc could rise and fall according to the prospects for an emotional relationship between the main character and their love interest.

Based on the widespread use of story arcs, and our observation that a wide variety of stories can be described using story arcs, our expectation was that this would be an intuitive mechanism for accurately specifying the criteria for stories of interest. We evaluate the intuitiveness and accuracy of ARC SIFT in a user study (see section 4).

3.1 Story Domain Test-bed

For the evaluation of this overall approach, we have modeled a simulated storyworld by creating an agent-based simula-

[1] Also referred to as: character arc, narrative arc or emotional arc.

tion based on the setting of the Prom Week scenario [McCoy *et al.*, 2012] using AgentMaps[2], a JavaScript library for building and visualising agent-based simulations. This narrative setting was chosen because of its popularity in television, films and books, therefore making it understandable to a general public. Further, it allows for the potential to produce a wide variety of stories with different endings, and resembles popular social simulation games such as "the Sims" series[3].

In this simulated storyworld, the characters are modelled as students at a school in the week before the prom where their goal is to find a prom date. Each character is assigned a dorm to sleep at night, and during the day they have classes to attend. Characters move between these locations at set intervals. As they move throughout the day, characters have the opportunity to meet and mingle with other students in order to form relationships. Characters may choose to ask another character to the prom, depending on how their current relationship is going with the character.

At the start of the simulation, characters are randomly assigned to classes and dorms. These randomly assigned factors affect which other characters they meet throughout the simulation making it non-deterministic.

Our simulation allows for adjustable parameters in order to scale up or down the simulation. This includes the number of students simulated, the amount of time spent in the dorm and in classes, and how often characters interact.

Each interaction event that occurs when the simulation runs is captured as output in the form of Abstract Interaction Logs (AIL), as used in [Behrooz *et al.*, 2015]. AIL is a language where events are described as tuples:

```
(subject, object, type, content, context, time)
```

where `subject` is the character instigating the interaction, `object` is the character who is the recipient of the event, `type` is the type of interaction event, `content` is the speech content of the character, `context` is the location in which the event takes place, and `time` is the time segment at which this interaction occurred. A sample of the AIL output from our Prom Week simulation is provided in Figure 1. Note that simulation values are mapped to appropriate names when stories are sifted for presentation to users.

3.2 Dynamic Character Networks

We use *Dynamic Character Networks* to track temporal changes in fortunes resulting from the interactions between virtual characters in our simulation. *Dynamic Character Networks* [Lee and Jung, 2020] represent and track the temporal changes in social affinity resulting from interactions between characters in a story or movie. We observe that these networks can be applied more broadly to simulated storyworlds, such as multi-agent simulations, to track the consequences of interactions between characters in the simulation over time via their *interaction consequence score*, and where the semantics of these interaction consequences are simulation domain specific.

In this broader context we define a *Character Network* and *Dynamic Character Network* as follows:

[2]https://github.com/noncomputable/AgentMaps
[3]https://www.ea.com/en-au/games/the-sims

```
(987, 1017, SpeechPositive, "Having a great time!,
    class3", 1581)
(987, 1017, SpeechProposal, "Will you come to the
    prom with me!", class3, 1581)
(1017, 1025, SpeechDrop, "I'm Sorry 1025 I'm not
    going to the Prom with you anymore.", class3,
    1581)
(1017, 987, SpeechAccept, "Yes 987! I'd love
    to go to the Prom with you!", class3, 1581)
```

Figure 1: Sample AIL output from the adapted Prom Week simulation between three characters Alice (987) and Bob (1017) and Charlie (1025). Here we see Alice have a positive interaction with Bob, before asking him to go to the prom with her. Bob responds by dropping their current date Charlie, before accepting Alice's proposal.

Character Network: $CN(S_t)$ is the character network of a simulation S at time $t \in [1, T]$, which is a matrix of the interaction consequence scores, $ics_{i,j}$, between each pair of characters i and j in the simulation. It can be represented as:

$$CN(S_t) = \begin{bmatrix} ics_{1,1} & \cdots & ics_{1,a} \\ \vdots & \ddots & \vdots \\ ics_{a,1} & \cdots & ics_{a,a} \end{bmatrix}$$

where, a is the total number of characters in the simulation.

Dynamic Character Network: $DCN(S)$ is a dynamic character network of S, which is an ordered sequence of $CN(S_t)$. It can be represented as:

$$DCN(S) = \langle CN(S_{t1}), \ldots, CN(S_{tT}) \rangle$$

where T is the total number of discrete time steps for simulation S.

In the context of our simulated storyworld Prom Week, the interaction consequence scores, $ics_{i,j}$, relate to the state of the relationship between the two characters i and j. Each character in Prom Week seeks to find a prom date, and their interactions with other characters will either increase or decrease their relationship score with the other character. For our evaluations we assumed that the consequences of the interactions between two characters were the same, with either positive or negative consequences on their relationship. That is, an interaction between characters i and j results in a score $ics_{i,j}$ that is equivalent to $ics_{j,i}$, irrespective of which character is the subject or object. An outcome of this is that the produced story arc tracking character i and their relationship over time with j has the same shape as the story arc for j with i. Thus, the total number of story arcs produced by the simulation is $\frac{(a-1)a}{2}$ where a is the number of characters in the simulation.

The Dynamic Character Network tracks the changes for each AIL event from Prom Week, according to a mapping that describes how different interaction event type affects the relationship scores between the characters. This is discussed further in Section 4.

Plotting the values of $ics_{i,j}$ for every $CN(S_t)$ in $DCN(S)$ against time t shows the changing relationship score between characters i and j. This can be interpreted as the changing fortunes of character i in pursuit of their goal of finding a

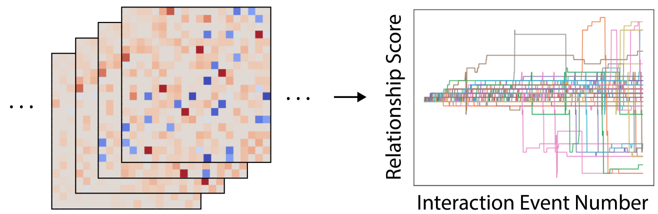


Figure 2: Dynamic Character Networks to Simulation Story Arcs: (Left) shows a fragment of 20x20 Dynamic Character Network from our Prom Week simulation; (Right) shows 190 simulation story arcs produced from this Dynamic Character Network.

prom date with character j, which satisfies the definition of a story arc for character i. We refer to these plots that are generated from Dynamic Character Networks as *simulation* story arcs.

Examples of simulation story arcs generated from a Dynamic Character Network can be found in Figure 2.

3.3 Dynamic Time Warping

Dynamic Time Warping (DTW) is an algorithm used to compare the similarity between time-dependant sequences, while allowing for elastic shifting in the time axis. This is useful when comparing two sequences that have similar shapes, while not necessarily aligning in the time axis. DTW has seen previous applications in a wide range of domains, including speech and sound processing [Myers *et al.*, 1980], handwriting and signature matching [Tappert *et al.*, 1990], and motion and gesture recognition [Varatharajan *et al.*, 2018].

There are two parts to the Dynamic Time Warping algorithm. Firstly DTW builds a *cost matrix* containing the cost values that compares each point in the query time-series with each point in the candidate time-series. The more similar the two values are, the lower the cost. Secondly, DTW seeks to finds the alignment between the two time-series with the lowest overall cost. This optimal alignment runs along the "valley" of low cost cells within the matrix. Upon finding this alignment, DTW returns the minimal overall cost score. The lower the score returned by DTW, the more similar the two time-series are. For further detail and formal definition of the DTW algorithm see [Müller, 2007].

While DTW can be applied to a diverse range of problems, these problems all seek to find, from a pool of candidate time-series, the best match for a query time series. Thus, DTW can also be applied to story sifting, as it enables a query story arc (a time series that describes what story a user is interested in) to be compared to a number of candidate simulation story arcs. The events corresponding to the simulation story arc that best match a given query arc can then be returned as the sifted story.

The Python package DTAIDistance [Meert *et al.*, 2020] was used in this work. Further implementation details of DTW can be found in DTAIDistance's documentation.

3.4 Arc Sift

To develop our story sifting prototype, Arc Sift, we implemented components that built Dynamic Character Networks from our Prom Week simulated storyworld and integrated this with use of Dynamic Time Warping of Query Arcs.

Arc Sift *Input:* consist of the following three inputs. Firstly, it requires AIL output of a simulated storyworld. Secondly, a configuration mapping how each interaction type affects the relationship scores between characters. Thirdly, it requires an input query story arc, describing the shape of the story of interest. This query arc can be plotted by a user or be computer generated.

Arc Sift uses the first two inputs to create a dynamic character network. From this, the changing relationship scores between characters, arising from their interactions are plotted against time as simulation story arcs for each pair of characters.

Arc Sift *Output:* Arc Sift computes a similarity score between a query story arc and all simulation story arcs. Arc Sift then returns the specific AIL events corresponding to the simulated story arc with the lowest cost score. Note that for presentation to users, as for our evaluation, we use template text translation to form text stories. In general the sifted stories could be mapped to other media, such as 2D or 3D visualisations.

4 Evaluation

4.1 Setup

In order to evaluate our approach to story sifting, we set up an instance of our Prom Week simulation (as described in Section 3.1) with 20 student characters. Each character was assigned 1 out of 7 possible dorm rooms, and 3 out of 5 possible classrooms. Students were all on the same daily schedule of spending 300 ticks per class, and 400 ticks at their dorms.

The simulation produced 782 interaction events. These interaction events were captured as output in the form of Abstract Interaction Log (AIL) events (as described in Section 3.1, which was fed into Arc Sift).

The configuration used which specified how each event type affected the relationship score was as follows: +10 for `Greet` and `SpeechPositive`, -10 for `SpeechNegative`, 0 for `InnerConflict` and `Proposal`, +100 for `Accept` and -100 for `Reject` and `Drop`.

From this, Arc Sift generated a 782 by 20 by 20 Dynamic Character Network (as there were 782 interaction events and 20 characters). This meant that 190 simulation story arcs were generated from the Dynamic Character Network. These simulation story arcs are shown in Figure 2.

4.2 Sifting Stories with Arc Sift

For the purposes of the evaluation, a selection of stories were sifted from our Prom Week simulation. The selection was based on a set of common shapes described in [Reagan *et al.*, 2016] and from which query arcs were generated. These common story shapes, descriptions and snippets of sifted stories are shown in Figure 3.

The text of the stories was produced from the simulation abstract interaction logs: mapping from simulation values to names, locations, and so on. The rise and fall reflect changing interaction consequence scores between characters (calculated using the different parameters shown in 4.1). For example, some of the snippets in Figure 3, such as the positive greetings ("Nice to meet you"), and acceptance ("I will go to the Prom"), increase this score between the characters, whilst others decrease it ("Sorry. I'm not going to the Prom with you").

4.3 User Study

We conducted a user study to assess how intuitive users found the use of story arcs as a mechanism to describe criteria for selecting stories of interest. The study consisted of 32 non-technical adult participants who were proficient English speakers, and was conducted using an online form. There was no time limit for participants to complete the online form. The participants were given 2 tasks as follows:

Task 1 (Intuitiveness): To test the intuitiveness of ARC SIFT, participants were presented with 6 summarised stories sifted from the Prom Week simulation using ARC SIFT. Users were asked to draw a story arc they felt best represented the summarised story presented to them (online, in a free draw box, using a mouse). The user drawn story arcs were collected, analysed and compared with the original query arcs used to sift the presented stories. Intuitiveness of ARC SIFT is shown if users can accurately draw an arc corresponding to a story which matches the arc that sifted it.

Task 2 (Accuracy): To assess the accuracy of ARC SIFT, participants were presented with 6 story arcs. For each story arc presented, participants were asked to select, from a set of story summaries, the summarised story they thought best matched the presented arc. Each set of story summaries contained the story summary that was sifted by ARC SIFT when presented with that story arc. Participants could also choose "other" if they felt none of the choices matched the presented story arc. Accuracy of ARC SIFT is shown, if given the same story arc, ARC SIFT sifts the story that corresponds to what the majority of participants selected.

The stories and story arcs chosen to present to the participants for both tasks are based on the six common story shapes as discussed in the previous section and shown in Figure 3.

At the end of each task, participants were asked to provide a confidence score for their answers for the completed task, using a 5-point Likert scale (where 1=low confidence and 5=high confidence). Participants also had the opportunity to provide any explanation for their ratings or general feedback via a free text field.

4.4 Results

Task 1 (Intuitiveness): In order to objectively assess the match between the participant drawn query arcs and the "target" query arcs that we used to sift the presented stories, we used the following 3 criteria based on properties of the arcs: the number of local troughs, MIN, local peaks, MAX, and the direction of the first inflection, FIRST.

This criteria is used to generate a vector of values of the form: (MIN, MAX, FIRST). We ranked a pair of arcs as matching when they had the exact same arc vector.

The results of matching analysis of the participants drawn story arcs and the target query arc for each of the 6 presented stories are shown in Figure 4. Note that for all of these presented stories, the story arc most commonly drawn by the participants matched the target query arc used to sift the presented story, i.e. the "correct" shape. The figure shows that 84.4% of participant drawn arcs matched the target arc for RagsToRiches, 81.3% of the participants matched the arc for RichesToRags and Icarus, 62.5% of the participants matched the arc for ManInHole, 50.0% of the participants match the target arc for Oedipus and Cinderella.

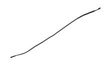
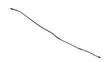
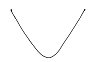
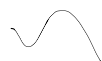
Query Arc	Description	Sifted Story Snippets
	RagsToRiches: fortunes constantly **rise**	(Alice, Bob, "Nice to meet you"), ..., (Alice, Bob, "Having a great time!"), ...
	RichesToRags: fortunes constantly **fall**	(Alice, Bob, "Nice to meet you"), ..., (Alice, Bob, "This is not fun"), ...
	ManInHole: fortunes **fall** then **rise**	..., (Alice, Bob, "I don't want to go to Prom ..."), ..., (Bob, Alice, "Should I go?"), ..., (Bob, Alice, "I will go .. "), ...
	Icarus: fortunes **rise** then **fall**	..., (..., "... a great time!), (Alice, Bob, "I will go .. "), ..., (..., "Sorry. I'm not ..."), ...
	Cinderella: fortunes **rise**, **fall**, then **rise**	(Alice, Bob, "Nice to meet .."), ..., (..., "This is not fun"), ..., (..., "Will you come ...?"), ..., (..., "I'd love to go!"), ...
	Oedipus: fortunes **fall**, **rise**, then **fall**	..., (..., "This is not fun"), ..., (..., "Having a great time!), ..., (..., ""Sorry. I'm not ..."), ..

Figure 3: Sample ARC SIFT sifted stories. Input query arcs shown are common patterns [Reagan *et al.*, 2016], accompanied by descriptions. Also shown are snippets of ARC SIFT sifted stories. These samples were used in the Evaluation (see text).

These results support our expectation that this approach is intuitive. A majority of participants drew arcs that matched the target query arc. This shows that they understood the concept of a story arc and how to describe a story using it. They agreed on what the shape of a story arc should look like when describing a story, even when participants were free to draw anything in a free draw box.

Task 2 (Accuracy): A summary of these results can be found in figure 4. For all of the 6 cases, the story ARC SIFT sifted was the same as the most commonly chosen story summary by the participants.

The results show that 84.4% of the participants selected the same story as ARC SIFT when presented with the RagsToRiches story arc, 87.5% of the participants selected the same story for RagsToRiches and Cinderella story arcs, 81.3% of the participants selected the same story for Oedipus, 56.3% of the participants selected the same story for ManInHole, while 28.1% of the participants selected the same story for Icarus.

These results suggest that ARC SIFT was accurately sifting the "correct" story when presented with a query story arc, as the story matched what the majority of participants expected to be sifted when presented with the that query arc.

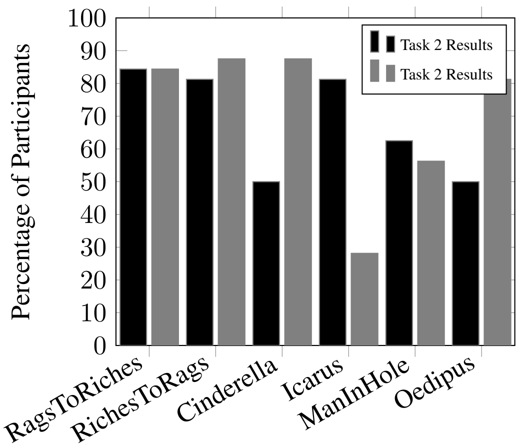


Figure 4: Task 1 Results: % participants who drew story arcs that correctly matched the target query arc. Task 2 Results: % participants who correctly selected the target sifted text story corresponding to a visual story arc.

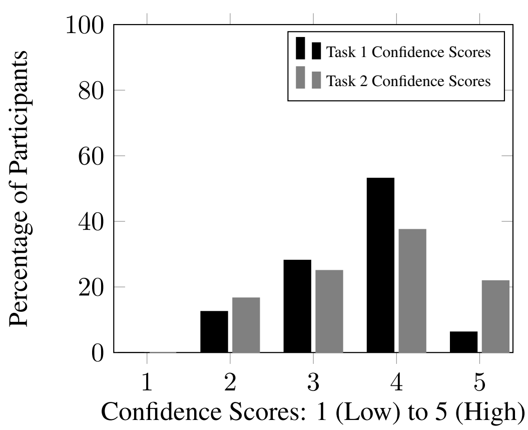


Figure 5: Participant Confidence Scores. For both tasks the majority of participants, 65%, reported feeling confident in their responses.

We note that the results for "Icarus" were lower than expected. This could indicate that our simulation produced no good examples of an Icarus story type when given the query arc it was presented with, or that there were multiple suitable choices for "Icarus".

Participant Confidence Rankings: The participant confidence rankings for both tasks are visualised in Figure 5. The mean rating for how confident participants felt in their answers for task 1 and task 2 were 3.53 and 3.66 respectively. For both tasks, the median rating was 4. For task 1, 59.4% of participants indicated that they were either confident or highly confident in their answers, with 28.1% of participants who were neutral (rating 3), and only 12.5% of participants who didn't feel confident in their responses. For task 2, 59.4% of the participants indicated that they were either confident or highly confident in their answers, 25.0% indicated that they were neutral, and 15.6% of participants indicated that they were not confident.

The positive confidence ratings suggest that the process of describing a story using shapes and matching a shape to stories is intuitive for users. This is also supported by the sentiment gathered by the free text fields, presented below.

Free text responses: For both tasks participants were asked to provide comments about the tasks. The following provides a flavour of responses for task 1: "... *watched enough TV dramas to expect those arcs, and anticipate emotions of characters.*"; "*The prompts were easy to read and comprehend whether the story was flowing in a positive or negative direction.*".

The following comments gives a flavour of responses for task 2: "*It was easy to trace the arc as I read the story*"; "*The curve rises when positive events occur and drops when negative events occur.*".

Overall these results match our expectations that sifting stories using shapes such as a story arc is feasible, and that it is an intuitive and accurate process.

5 Conclusion

In this paper we presented an intuitive and user-friendly method for describing desired story criteria and an automated approach to sift stories generated by a simulated storyworld corresponding to those criteria. This method used *dynamic character networks* to track the changing relationships between characters over the course of the simulation and *dynamic time warping* to find the most similar story to that described by a query story arc. These components were implemented in our experimental prototype ARCSIFT. We presented the results of a user study evaluating the inputs and outputs of ARCSIFT. These results were consistent with our expectations that story sifting using story arcs was accurate and intuitive for users.

The work presented in this paper provides a strong base for future research, for example we plan to apply ARC SIFT to sift stories from a wider set of domains with larger search spaces, sift stories involving more than two characters, sift by matching to granular segments of a simulation story arc, and tailor the dynamic character network based on each individual user's interpretation of events.

Acknowledgements

This research was supported by funding from the Commonwealth of Australia.

References

[Behrooz *et al.*, 2015] Morteza Behrooz, Reid Swanson, and Arnav Jhala. Remember That Time? Telling Interesting Stories from Past Interactions. In *Interactive Storytelling*, Lecture Notes in Computer Science, pages 93–104, 2015.

[Cheong and Young, 2015] Yun-Gyung Cheong and Robert Michael Young. Suspenser: A story generation system for suspense. *IEEE Trans. Comput. Intell. AI Games*, 7(1):39–52, 2015.

[Grinblat and Bucklew, 2017] Jason Grinblat and C. Brian Bucklew. Subverting historical cause and effect: generation of mythic biographies in Caves of Qud. In *Foundations of Digital Games (FDG)*, 2017.

[Kreminski *et al.*, 2019] Max Kreminski, Melanie Dickinson, and Noah Wardrip-Fruin. Felt: A simple story sifter. *Lecture Notes in Computer Science*, 11869 LNCS:267–281, 2019.

[Kreminski *et al.*, 2020] Max Kreminski, Noah Wardrip-Fruin, and Michael Mateas. Toward Example-Driven Program Synthesis of Story Sifting Patterns. In *Proceedings of the AIIDE 2020 Workshops on Artificial Intelligence and Interactive Digital Entertainment (AIIDE)*, 2020.

[Kybartas *et al.*, 2021] Ben Anderson Kybartas, Clark Verbrugge, and Jonathan Lessard. Tension Space Analysis for Emergent Narrative. *IEEE Trans. Games*, 13(2):146–159, 2021.

[Lee and Jung, 2020] O-Joun Lee and Jason J. Jung. Story embedding: Learning distributed representations of stories based on character networks. *Artificial Intelligence*, 281:103235, April 2020.

[Lindsay *et al.*, 2015] Alan Lindsay, Fred Charles, Jonathon Read, Julie Porteous, Marc Cavazza, and Gersende Georg. Generation of Non-compliant Behaviour in Virtual Medical Narratives. In *Proceedings of Intelligent Virtual Agents - 15th International Conference, IVA*, Lecture Notes in Computer Science, pages 216–228, 2015.

[Louchart *et al.*, 2015] Sandy Louchart, John Truesdale, Neil Suttie, and Ruth Aylett. *Emergent Narrative, Past, Present and Future of an Interactive Storytelling Approach*, pages 185–199. 2015.

[Mateas and Stern, 2005] Michael Mateas and Andrew Stern. Structuring Content in the Facade Interactive Drama Architecture. In *Proceedings of the 1st Conf. on AI and Interactive Digital Entertainment (AIIDE)*, 2005.

[McCoy *et al.*, 2012] Josh McCoy, Mike Treanor, Ben Samuel, Aaron A. Reed, Noah Wardrip-Fruin, and Michael Mateas. Prom week. In *Proceedings of the International Conference on the Foundations of Digital Games*, pages 235–237, Raleigh, North Carolina, May 2012.

[McKee, 1997] Robert McKee. *Story: substance, structure, style, and the principles of screenwriting*. NY: ReganBooks, 1997.

[Meert *et al.*, 2020] Wannes Meert, Kilian Hendrickx, and Toon Van Craenendonck. wannesm/dtaidistance v2.0.0. Aug 2020.

[Mott and Lester, 2006] Bradford W. Mott and James C. Lester. U-director: A decision-theoretic narrative planning architecture for storytelling environments. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, 2006.

[Myers *et al.*, 1980] C. Myers, L. Rabiner, and A. Rosenberg. Performance tradeoffs in dynamic time warping algorithms for isolated word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(6):623–635, December 1980. Conference Name: IEEE Transactions on Acoustics, Speech, and Signal Processing.

[Müller, 2007] Meinard Müller. Dynamic Time Warping. In Meinard Müller, editor, *Information Retrieval for Music and Motion*, pages 69–84. Springer, Berlin, Heidelberg, 2007.

[Osborn *et al.*, 2015] Joseph Carter Osborn, Ben Samuel, Michael Mateas, and Noah Wardrip-Fruin. Playspecs: Regular Expressions for Game Play Traces. In *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*, September 2015.

[Porteous *et al.*, 2011] Julie Porteous, Jonathan Teutenberg, David Pizzi, and Marc Cavazza. Visual programming of plan dynamics using constraints and landmarks. In *Proceedings of the Twenty-First International Conference on International Conference on Automated Planning and Scheduling*, ICAPS'11, page 186–193. AAAI Press, 2011.

[Porteous *et al.*, 2017] Julie Porteous, Fred Charles, Cameron Smith, Marc Cavazza, Jolien Mouw, and Paul van den Broek. Using Virtual Narratives to Explore Children's Story Understanding. In *Proceedings of the 16th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2017.

[Reagan *et al.*, 2016] Andrew J. Reagan, Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1):31, November 2016.

[Rhodes *et al.*, 2010] Martin Rhodes, Simon Coupland, and Tracy Cruickshank. Enhancing real-time sports commentary generation with dramatic narrative devices. In *Interactive Storytelling*, 2010.

[Riedl and Bulitko, 2013] Mark O. Riedl and Vadim Bulitko. Interactive Narrative: An Intelligent Systems Approach. *AI Magazine; La Canada*, 34(1):67–77, 2013.

[Riedl and Young, 2010] Mark O. Riedl and R. Michael Young. Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research*, 39(1):217–268, 2010.

[Ryan *et al.*, 2015] James Ryan, Michael. Mateas, and Noah Wardrip-Fruin. Open design challenges for interactive emergent narrative. *Lecture Notes in Computer Science*, 9445:14–26, 2015.

[Ryan, 2018] James Ryan. *Curating Simulated Storyworlds*. PhD thesis, UC Santa Cruz, 2018.

[Tappert *et al.*, 1990] Charles Tappert, Ching Yee Suen, and Toru Wakahara. The state of the art in online handwriting recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(8):787–808, 1990.

[Varatharajan *et al.*, 2018] Ramachandran Varatharajan, Gunasekaran Manogaran, Malarvizhi Kumar Priyan, and Revathi Sundarasekar. Wearable sensor devices for early detection of Alzheimer disease using dynamic time warping algorithm. *Cluster Computing*, 21(1):681–690, March 2018.

[Weiland, 2016] K. M. Weiland. *Creating Character Arcs: The Masterful Author's Guide to Uniting Story Structure, Plot, and Character Development*. PenForASword Publishing, 2016.