



# Evolutionary combination of connected event schemas into meaningful plots

Pablo Gervás<sup>1,2</sup> · Gonzalo Méndez<sup>1,2</sup> · Eugenio Concepción<sup>1</sup>

Received: 15 October 2022 / Revised: 25 January 2023 / Accepted: 4 May 2023 /  
Published online: 7 June 2023  
© The Author(s) 2023

## Abstract

Many of the stories we are exposed to are built from small schemas of connected events involving a set of characters—*boy meets girl* leads to a *relationship* or *crime* leads to *revenge*. The present paper proposes an evolutionary solution to the task of putting together a story by combining a set of such schemas. This approach presents three challenges: how to mix up the elements in the different schemas, how to instantiate the characters across the schemas and how to tell acceptable combinations from the rest. The present paper applies an evolutionary solution that relies on a genetic representation for these combinations of schemas, and applies as fitness functions a set of metrics on compatibility constraints across schema combinations. Outputs of this procedure are evaluated by human judges in comparison with baseline solutions in which the values for genes are assigned at random. The proposed solution generates a population of story drafts that resemble plot descriptions for simple stories. The results of the comparative evaluation by human judges are positive. The genetic representation of pattern combinations and the metrics on compatibility across pattern pairs provide a valid evolutionary solution for constructing simple plots.

**Keywords** Story generation · Subplot patterns · Evolutionary approach · Metrics on pattern compatibility · Character instantiation

---

✉ Pablo Gervás  
pgervas@ucm.es

Gonzalo Méndez  
gmendez@ucm.es

Eugenio Concepción  
econcepc@ucm.es

<sup>1</sup> Facultad de Informática, Universidad Complutense de Madrid, c/ Profesor García Santesmases, 9, 28040 Madrid, Spain

<sup>2</sup> Instituto de Tecnología del Conocimiento, Universidad Complutense de Madrid, c/ Profesor García Santesmases, 9, 28040 Madrid, Spain

## 1 Introduction

Most of us consume stories regularly in our everyday life, whether in the form of movies, novels, TV series or podcasts. Many of these stories are built from small schemas of connected events involving a set of characters—*boy meets girl* leads to a *relationship* or *crime* leads to *revenge*. As we read or watch the stories we identify such schemas in the stories and remain on the lookout for the events that complete them. A great part of our enjoyment of stories arises from this process—sometimes from seeing the schemas completed, sometimes from seeing them transgressed. Most people can identify this type of schema, and yet there is no consensus on what the set of such schemas might be. Some non-academic efforts have been made to compile instances of these schemas, such as the *TVTropes*<sup>1</sup> web site, a very impressive crowd-sourced compilation [2]. But even if these building blocks are identified very little is known about the process by which they are combined to yield full stories. The present paper proposes an evolutionary solution to the task of putting together a story by combining a set of such schemas. This approach presents three challenges: how to mix up the elements in the different schemas, how to instantiate the characters across the schemas and how to tell acceptable combinations from the rest.

The present paper brings together three separate research lines related to narrative generation: existing work on potential means of representing single plot lines at different levels of granularity—which correspond to the schemas mentioned above—[13], work on development of multi-plot stories by combining a set of individual plot lines [6], and an evolutionary solution for combining several single plot lines into a complex plot [14]. The evolutionary solution relies on a genetic representation for these combinations of schemas, and on fitness functions informed by a set of metrics on compatibility constraints across schema combinations. The main contribution of this paper is the application of an evolutionary solution for the combination of these schemas—small fragments of plot that are not necessarily plot lines in themselves—into complex plots using the procedures previously developed to combine full individual single plot lines into multi-plot stories. Outputs of this procedure are evaluated by human judges in comparison with baseline solutions.

## 2 Related work

Three topics are considered relevant for this paper: prior solutions for the representation of plot, approaches to constructing stories by combining small units of representation, and evolutionary approaches to creation of narratives of some kind.

---

<sup>1</sup> <https://tvtropes.org/>.

## 2.1 Representing plot at the right granularity

The understanding of narrative as a form of communication has been a major subject of study in the field of humanities and became a challenge for computational approaches since the early advent of artificial intelligence. Some relevant approaches are reviewed here.

Russian formalist Vladimir Propp studies a corpus of Russian folk tales and proposed a formal representation for the basic units that made up their plot structures [31]. Propp postulated the concept of a *character function* as a relatively abstract representation of the meaning of an event involving some characters that is relevant to the plot of the story. These events represent the structural elements in a story at a very low level of granularity, because they involve individual actions such as characters meeting, misbehaving, fighting, or travelling. However, Propp identified that these events were connected to other events in the story by virtue of specific characters that necessarily took part in the same set of events. In this way, the victim of an abduction event at the start of the story establishes a connection with the rescue event that happens—to the same character—later in the story. Propp postulates a set of *spheres of action* that define certain specific roles that characters may play in a story: hero, villain, victim, donor, helper...Because they are quite simple and yet provide a semblance of formal structuring, character functions have been used often as basic representation in attempts to generate stories automatically, such as in the authoring tool for interactive fiction Grasbon and Braun [17], the OPIATE interactive story generator [9] or the Teatrix story creation system [25]. But, being limited to representing individual events, they fall short of providing a usable representation of the types of schemas that we want to consider.

Attempts to capture the structure of plot from beginning to end do consider sequences of events that correspond to observed plot archetypes. Existing efforts postulates different numbers of such archetypes as basic structures to understand narrative: one for *the hero's journey* [3], seven by Booker [1], twenty by Tobias [36] or 1462 in Plotto [7]. Such efforts go to the other extreme, because they represent very complex units that completely define the plot of the story. They are therefore too large to represent the type of schemas that we require.

An intermediate degree of granularity has been defined in the *axes of interest* postulated in the PlotAssembler system [13]. These axes of interest (or AoIs) are small sets of events that do not necessarily occur contiguously in the discourse for a story but which are connected by shared characters that give them meaning—like the victim of a kidnapping being rescued later. As axes of interest are chosen to represent subplot schemas in the present paper, they are reviewed below.

Although these axes of interest are presented specifically for representing plot at an intermediate granularity they correspond more closely to the concept of script as defined by Schank and Abelson [33]. In this approach, a *script* is “a structured representation describing a stereotyped sequence of events in a particular context”. The concept of event considered in this definition corresponds to primitive acts performed by an actor on an object. The concept of script carries a certain connotation that the events in a script take place not just in order but, to a certain extent, in close temporal proximity. Scripts have been considered as a possible operational unit for

representation in story understanding and story generation systems by McKenzie et al. [23].

Existing work in natural language processing also considers the concept of a *narrative schema* [4]. In this context, a narrative schema is defined as a “coherent sequences or sets of events (arrested(POLICE, SUSPECT), convicted(JUDGE, SUSPECT)) whose arguments are filled with participant semantic roles defined over words” (Judge = judge, jury, court, Police = police, agent, authorities).

## 2.2 Story construction by combination of plot relevant units

The use of planning technologies for story generation [37] may be considered an instance of processes of construction of stories by combining partially structured fragments of story material. In this case, the basic units used for construction are *planning operators*, which include a story action that represents the main event of the operator—usually in the form of a predicate-argument structure—, and a number of preconditions and postconditions—also represented by similar predicates. When building a plan structure to represent the outcome story, preconditions may be unified with predicates already in the plan—and not necessarily at positions in the story discourse contiguous to the event being added at that stage. Arguments shared across preconditions, main action and postconditions represent connections between different events. Used in this way, planning operators could be seen as possible representations for the type of schema we want to represent. However, for planning techniques to be applicable the relations between the preconditions and the main action of a planning operator need to imply a certain causal relation. This is not necessarily true in many of the schemas we want to represent.

A different approach that also builds stories by combining predefined fragments of material that are partial representations of plot can be found in recent attempts to build stories with more than one plot line. Stories beyond the simpler instances are known to involve often more than one plot line. A *plot line* when used in this context refers to a sequence of plot-relevant elements or *scenes* that make sense in the order in which they appear in the story, and linked by at least a shared set of protagonist and secondary characters. The schemas that we want to represent may indeed be considered very small instances of plot lines, though in general, the concept of plot line has a connotation of slightly more complex sets of events and of the relations between them.

Closely related to the planning approach described above, Porteous et al. [30] present a plan-based procedure for creating multi-plotline stories for an interactive storytelling system. The complete plan is built incrementally as partial selections from the plans that result from attempting to lead the draft at that point towards a predefined goal. At each point, only the next action in the given plan is added to the draft before the initiative is passed to the user. The user intervention usually results in a need to rebuild the plan. In their approach, the different plot lines are represented by different spans of the overall draft involving specific sets of characters.

This concept of subset of a story involving a particular character is sometimes referred to as a *narrative thread*. The work by Fay [10] relies on narrative threads

of this type as building units for constructing complex stories. The approach starts from a set of narrative threads for particular types of characters—extracted from a corpus of existing stories—, and, for a given story request that mentions specific types of character, constructs multi-plotline stories by first selecting threads matching the type of characters in the request, finding a combination of the elements from each thread into a consistent timeline, and identifying valid bindings between characters in different threads that make the story consistent.

The PlotAssembler system [13]—which introduces the concept of axes of interest mentioned above—takes as input a set of axes of interest provided and interweaves the scenes in these in an order designed to maximise the probabilities of character continuity across scenes—as mined from a corpus of prior stories.

At the furthest level of granularity in the representation of plot, the work of Concepción et al. [6] operates on a set of plot templates for complete stories, and it proposes procedures for weaving them together into multi-plotline stories. Some of these procedures are drawn from known techniques used in existing narrative but they also include simple computational approaches that are presented as baselines to compare with.

### 2.3 Evolutionary construction of narratives

Evolutionary solutions have been used in the past to construct stories from smaller units. McIntyre and Lapata [22] use genetic algorithms to explore the search space of possible merges between plot lines previously extracted from a set of stories. Each plot line is represented as a partially ordered graph of events associated with a given entity. A set of entities is received as input and the process is driven by a fitness function designed to maximise story coherence and story interest. Gómez de Silva Garza and Pérez y Pérez [35] build stories by using the GENCAD evolutionary approach for the adaptation stage in case-based solutions to architectural problems [34] to refine an initial population built using the knowledge-based heuristics of the MEXICA knowledge-based story generator [28]. Fredericks and DeVries [11] present a generator of small fragments of narratives—to be used in text-based games—that applies an evolutionary solution driven by novelty search [20]. Kartal et al. [19] generate narratives using a plan-based approach supported by a Monte Carlo tree search driven by a combination of measures of how believable the resulting story is and how many of the goals defined by the user are accomplished by the story. de Lima et al [21] generate quests for games by combining a planner that constructs candidate quests as linear sequences of tasks for the user with an evolutionary search strategy that selects from them those that best match a target curve provided by the user of how tensions should evolve in the quest.

The work of Gervás et al [14] explores an evolutionary solution for the combination of plot templates for complete stories as described by Concepción et al. [6]. This approach proposes a genetic representation for a combination of fragments of plot—such as plot templates or plot lines—that includes genes that govern the order in which elements from different fragments appear in the final discourse and genes that govern how character variables from different fragments may be

instantiated to the same character in the final story. This division corresponds to the two main tasks that make up the process: *discourse planning*—decisions about in what order to present the elements of the story as a sequential discourse—and *character fusion*—decisions about how characters from the different fragments being combined are themselves fused into a single character in the final story. The fitness functions used in this approach relied on metrics that measured how consistent the final story was in terms of basic semantics such as characters being alive in the story at points of the story where they are active. Such constraints had been identified as relevant to human judges in the formative evaluation carried out by by Concepción et al. [6].

The schemas that we want to combine in the present paper share with plot lines the characteristics of being an ordered sequence of plot relevant elements related by a set of shared characters. The combination we hope to achieve for these schemas will also require that plot relevant elements from different schemas be interleaved in the final result, and that certain selected characters from some schemas be fused with characters from other schemas. For this reason, we hypothesize for this paper that the evolutionary mechanism for combining plot lines developed by Gervás et al. [14] will be applicable—with relatively little adaptation required—to combine the schemas we we want to consider.

### 3 Evolutionary combination of plot units driven by consistency metrics

The solution described in this paper explores how well the task of combining sub-plot schemas into a simple story can be addressed by a combination of the following four elements: the *representation of plot* as axes of interest [13], the application of the *genetic representation* presented by by Gervás et al. [14] for combining spans of partially ordered plot elements, a set of new *metrics on compatibility of patterns of combination* of plot element for pairs of axes of interest and a *preprocessing stage* that checks a given set of axes of interest for mutual compatibility—in terms of that constraints on relative ordering that arise between the elements involved.

#### 3.1 Knowledge representation for plot

To achieve the goal of the paper a representation is needed for schemas of connected plot relevant events that can be considered building blocks for larger patterns of plot such as plot lines. For lack of a better word we will refer to them as *schemas of connected events*. These building blocks need to be themselves constructed from plot relevant elements that align with the concept of event, and they need to allow representation of the characters that take part in them. In this paper we will use *axes of interest* to represent the concept of a schema of connected events, and the axes of interest will be ordered groupings of *plot atoms*.

### 3.1.1 Plot atoms as basic units of plot

As smallest unit of plot relevant element we will consider the concept of *plot atom*. A plot atom is conceptually similar to a character function in that it represents an action by one or more characters that is relevant to some aspect of the plot of the story. In contrast to character functions, each plot atom explicitly holds additional information to indicate how the roles specific to the plot atom (*kidnapper*, *kidnapped*) are filled in by characters playing roles that are relevant to the plot (**villain**, **victim**).<sup>2</sup> This refinement allows for interesting articulation between roles specific to a plot atom and roles more general to the plot at large. The variables employed in a plot atom to represent the participating entities are separated into three different sorts: characters, objects and locations. In this way, objects and locations may play relevant roles in the plot as well as characters. The use of sorts to separate these types of entities ensures that during evolution there are no instances of characters replaced with objects or objects mistaken for locations.

### 3.1.2 Axes of interest (Aols) as representation of connected event schemas

The type of small schema of related and not necessarily contiguous plot atoms that we want to operate with are represented by *axes of interest* (AoIs). An axis of interest is a sequence of plot atoms related by a conceptual dependency. For example, a schema representing a JOURNEY would include a plot atom for an event of *Departure*—usually somewhere towards the start of the story—and plot atom for an event of *Return*—again often somewhere towards the end of the story— but these two plot atoms are structurally connected.<sup>3</sup> The conceptual dependency may operate over a long range—as in the example of a journey—or at very short range—such as in a **CONFLICT**, where a *Struggle* is closely followed by a *Victory*. An axis of interest has a set of *narrative roles*—those of its constituent plot atoms—that are initially free variables but which can be instantiated to specific constants representing entities when the axis of interest is combined with other AoIs into larger structures. When a variable in this set is instantiated to a particular entity name, all the appearances of it in the associated plot atoms are instantiated as well. In the example above, for the AoI to make sense the traveller in a JOURNEY needs to be the same in the *Departure* and the *Return*, and the *origin* location for the *Departure* needs to match the *end* location of the *Return*.

Three different examples of axes of interest are shown in Table 1.

The set of axes of interest—and the corresponding set of plot atoms—used in the present paper resulted from the knowledge engineering effort described by Gervás

<sup>2</sup> Throughout the paper, the convention has been applied to present role names in *Italic* font and the names used for identifying the characters that fulfill the roles in specific instances in **Bold** font. This has been considered useful to facilitate understanding in cases where the terms used for role names and character names may lead to confusion.

<sup>3</sup> Along the lines of the convention already established for representing roles in *Italic* and characters in **Bold**, names of axes of interest are represented in SMALL CAPS font and names of plot atoms are represent in typewriter font.

**Table 1** Three examples of Axes of Interest, with one (RAGS2RICHES), two (HAPPYLOVE) and three (SHIFTINGLOVE) linked participating characters

AXISofINTEREST =	RAGS2RICHES
Poverty	<i>sufferer = x</i>
Aspiration	<i>aspirer = x</i>
Transformation	<i>transformed = x</i>
AspirationAchieved	<i>achiever = x</i>
AXISofINTEREST =	HAPPYLOVE
BoyMeetsGirl	<i>boy = x, girl = y</i>
FallInLove	<i>lover = x, beloved = y</i>
HappyEverAfter	<i>lover = x, beloved = y</i>
AXISofINTEREST =	SHIFTINGLOVE
GirlMeetsOtherBoy	<i>previous-love = x, girl = y, boy = z</i>
LoveShift	<i>lover = z, beloved = y, rival = x</i>
Reconciliation	<i>lover = x, beloved = y</i>

Co-occurrence in variable names is relevant, so that, in the final story, the character that plays  $x$  in any of the plot atoms for a given AoI must be the same as the one that plays  $x$  in all the other plot atoms for the same AoI

[13]. In this effort, a number of sources in the literature were consulted—including Propp’s character functions [31], Booker’s seven basic plots [1] and Polti’s situations [29]—and a process of abstraction and condensation was applied. As a result a set of 34 basic plot atoms was obtained, together with a set of 19 axes of interest that provide possible schemas of structuring for particular subsets of plot atoms. Interested readers are referred to the original paper for details.

### 3.2 Combining Aois into story drafts

As representations of the kind of small schemas of related events that occur in a plot, we want these axes of interest to be combined together, interleaving the various sequences of atoms of the AoIs involved in an order that makes sense as description of the plot of a story. We consider such a description of the plot of a story as a *story draft*. In a story draft, the ordered sequence of plot atoms from the axes of interest is referred to as the *discourse* for the story draft. In this discourse, each plot atom carries an additional label to indicate the axis of interest that it comes from.

An example of story draft is presented in Table 2, which shows how the HAPPYLOVE, UNRELENTINGGUARDIAN and TASK axes of interest are interleaved to form the basic story draft. It also shows how the narrative roles for the story draft (**hero**, **love-interest**, **parent**) are mapped to the roles specific to the plot atoms of the constituent axes of interest (*boy*, *girl*, *lover* and *beloved* for the HAPPYLOVE axis of interest; *lover*, *beloved* and *guardian* for the UNRELENTINGGUARDIAN axis of interest; and *setter* and *solver* for the TASK axis of interest). This ensures that the various plot atoms



**Table 2** Example of story draft for a basic plot combining axes of interest for HAPPYLOVE, UNRELENTING-GUARDIAN and TASK

HAPPYLOVE	BoyMeetsGirl	boy = <b>hero</b> girl = <b>love-interest</b>
HAPPYLOVE	FallInLove	lover = <b>hero</b> beloved = <b>love-interest</b>
UNRELENTINGGUARDIAN	CoupleWantsToMarry	lover = <b>hero</b> beloved = <b>love-interest</b>
UNRELENTINGGUARDIAN	UnrelentingGuardian	lover = <b>hero</b> beloved = <b>love-interest</b> guardian = <b>parent</b>
TASK	DifficultTask	setter = <b>parent</b> solver = <b>hero</b>
TASK	Solution	solver = <b>hero</b>
UNRELENTINGGUARDIAN	GuardianRelents	lover = <b>hero</b> beloved = <b>love-interest</b> guardian = <b>parent</b> )
UNRELENTINGGUARDIAN	Wedding	lover = <b>hero</b> beloved = <b>love-interest</b>
HAPPYLOVE	HappyEverAfter	lover = <b>hero</b> beloved = <b>love-interest</b>

The first column shows the interweaving of the axes of interest. The co-occurrence of constants in the final column—shown in **bold**—provides the argumental connections between the three AoIs

in the plot are instantiated in a manner coherent with the narrative roles that the characters play in the overall story draft.

The inclusion of this type of connection in terms of shared characters between the constituent AoIs in a story constitutes an instance of *character fusion*. These connections that relate plot atoms across the different axes of interest being combined are going to be used to build the metrics that will be used as fitness functions in the evolutionary procedure.

### 3.3 Metrics for acceptability of stories

Any process of computational construction of stories is likely to yield a large number of potential stories, so there is a need for some means of measuring the quality of drafts that can help identify valuable candidates among this search space. Intuitively, it seems that the perception of story quality is closely related to emotions represented in the story and the emotional impact on the reader [12]. There have been numerous efforts to develop valid metrics for story quality. The work by Gomes et al. [15] shows that the perception of quality for narratives is greatly influenced by many subjective matters such as behavior coherence, change with experience, awareness, behavior understandability, personality, visual impact, predictability, social and emotional expressiveness. However, such features present two serious

difficulties: very little is known about them and they require specific layers of representation to capture.

Recent work on developing metrics for story quality considers factors such as consistent use of entities across the narrative [27] or comparing the probability of each sentence in the story with and without its preceding story context [32]. These metrics are designed to operate on narratives rendered as a full text, therefore they would not be directly applicable to the outputs of our system, which are sketches of plot structure rendered in textual form.

For these reasons, we are restricting the evaluation of quality of candidate stories to considering them acceptable in terms of the two aspects that the procedure is designed to consider: whether the relative order in the sequence of plot atoms in the discourse makes sense, and whether the co-occurrence of entities across the different AoIs in the story is coherent.

To this end, we develop a set of metrics that measure correct sequencing and correct co-instantiation of variables over each potential pairwise combination of two AoIs, designed to cover the following aspects:

- *Role-sharing constraints* on a particular character playing a role X in one of the AoIs and a role Y in the other (say, the traveller in a journey becoming the victim of a kidnapping)
- Particular *sequencing constraints* on the atoms for the AoIs involved, possibly arising from a particular shared role (for instance, a kidnapped traveller should return only after he has been released from his kidnapping)

An example of the way these constraints are represented in the entries for a particular combination of a pair of AoIs is:

```
CallToActionReward+Conflict =
hero + hero =
CallToAction + Struggle / Victory + Reward
```

This expresses the fact that, for a combination of a hero being called to action (CALLTOACTIONREWARD) and involved in a fight (CONFLICT)—first line—, the hero of one should be the hero of the other—second line—, the fight should take place after the hero has been called and the reward should be obtained after the victory—third line.

Not all pairwise combinations of AoIs allow the formulation of this type of constraints. Whereas a combination of an ABDUCTION schema—corresponding to a kidnapping—and a STRUGGLE schema—which corresponds to a fight—does suggest that the fight involve the rescuer and the villain, there is no similar obvious connection when combining, for example CROSSDRESSING schema— which involves someone dressing up as a member of the opposite sex—and a REPENTANCE schema—which involves someone drastically changing their minds about their view on a prior decision. There may be cases in specific stories where a character repents about disguising themselves as a member of the opposite sex, but there is

no generic intuition that this connection might make for a more interesting narrative, as the connection between a kidnapping and a fight does. If pairs of schemas that do not have an immediate intuitive relation co-occur in the same story, they can still end up connected in the story due to the particulars of the construction procedure employed. They may become connected directly—having the cross-dressed character repent—as a result of character fusion operations resulting from the application of the evolutionary operators. They may also become connected in an indirect fashion via their relations with other schemas present, such as the cross-dressed person committing a villainy (VILLAINYPUNISHMENT schema) and then repenting of having done so.

The fact that constraints will not exist for some possible pairs of AoI combination is actually helpful, because the constraints arising from different pairs may be incompatible with one another, and too many constraints make it difficult to produce acceptable solutions. A specific solution is required to handle this profusion of constraints when these metrics are applied as fitness functions for our evolutionary process (see Sect. 3.5).

A particular pairwise combination of AoIs is assigned a numerical score over a total of 100. Of that score, 50 points are assigned based on role-sharing constraints. Each role-sharing constraint present is scored 100 if met and 0 otherwise, and the average value of all role-sharing constraints taken as the total role-sharing score (normalised to 50). The remaining 50 points are computed by:

- Assigning 100 points to any precedence constraint that is met (for  $A + B$ ,  $A$  appears before  $B$  in the discourse sequence)
- If a required precedence constraint is not met, a partial score over 100 is assigned corresponding to the number of positions that one of the elements would need to shift for the constraint to hold (normalised over the length of the sequence)
- The average of all sequencing constraints is taken as the total sequencing score (normalised to 50).

We have chosen to score over 100 because we reckon the problem that we are addressing requires a broad range of score values and the ability to distinguish between solutions that are different and yet scored to similar values by the metrics. These requirements arise from two specific characteristics of the problem. One, that, in the case of combinations involving a large number of AoIs, the scores are computed from contributions from a set of metrics whose cardinality is exponential on the number of AoIs. If we opt for a simpler scoring solution—such as 0-10 or 1-5—there would a much higher risk of individuals in the population reaching the same score even when there are differences between them. Since the problem requires a fine level of precision we prefer to have a higher top value for the score rather than having to express the scores for specific individuals with decimal places. Two, that the individuals obtained in early generations are likely to score very badly and then progressively improve their scores over the generations. We therefore need a score that can capture this progressive increase, and that will allow at each stage of the process to assign different scores to different individuals in a way that allows their relative ranking.

These metrics provide a progressive scoring, so that drafts where the sequencing constraints are not met are scored relative to how far they need to be modified for the constraints to be met. This allows mutations that modify the sequence in the right direction to be scored progressively higher, allowing evolution to converge towards optimal solutions.

### 3.4 Representing AoI-based stories for evolutionary construction

As mentioned in Sect. 2.2, the difference between a small schema for a subset of events appearing in a plot and a plot line as defined by Gervás et al. [14] is that the small schema does not usually cover a complete story on its own. But in both instances they correspond to a sequence of plot atoms connected by restrictions on shared characters so they are structurally equivalent in computational terms. We can therefore use a very similar formalism for representing the building blocks to construct stories and very much the same evolutionary mechanism for the task of combining them together into stories. The genetic representation presented by Gervás et al. [14] for combining templates for complete plots can be adapted to the combination of AoIs. To ensure that the present paper is understandable as a self-contained unit, this adaptation is described here in detail.

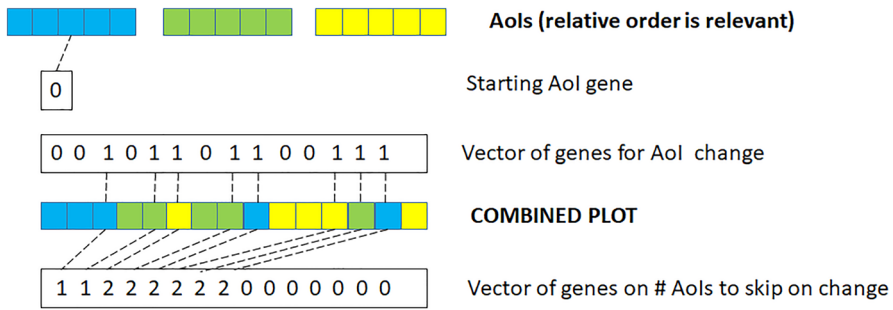
Because the task of combining the AoIs into stories requires decisions at two different levels—*discourse planning* decisions concerning the relative order of presentation of plot atoms and *character fusion* decisions concerning instantiations of shared characters across AoIs—the representation will require separate features to deal with each level.

#### 3.4.1 Genetic representation for discourse planning

Refined instances of storytelling often rely on advanced mechanisms for presenting a story—flashbacks, flashforwards—that involve presenting the scenes in the story in an order that differs from the chronological order in which they are supposed to have happened. Such instances of altered chronology are indeed very powerful tools and very interesting to explore as additional computational challenges, but we consider them beyond the scope of the present paper and we leave them for further work. We will therefore assume that the relative chronological order of the scenes in each AoI is respected in the final discourse.

A genetic representation of the discourse plan for a given story candidate must represent the following information: which AoI the discourse starts on, at what point in an AoI the discourse switches to a different AoI, and to which of the other AoIs in the draft does the discourse jump when it abandons the AoI it was on. To capture this information we use vectors that define the answer to these questions as follows:

- A single digit (0 or 1) defines which AoI the final discourse starts with
- A sequence of digits (0 or 1) defines for the total number of scenes in the final discourse whether the next scene follows on with the prior AoI (0) or it switches to a different AoI (1)



**Fig. 1** Diagram showing relation of discourse planning genes with actual discourse planning as a combination of AoIs. Three different AoIs shown in different colours, and the resulting combination as encoded by the vector of 0 s and 1 s for the discourse planning genes

- A sequence of digits (ranging between 1 and  $N-1$ , where  $N$  is the total number of plots being combined) defines how many of the available AoIs are skipped whenever the discourse switches to a different AoI

An example of a combination of AoIs—shown in different colours—as encoded by a particular assignment of 0 s and 1 s to the vector of discourse planning genes is shown in Fig. 1. In this example you can see how the *starting AoI gene* determines which AoI the combination should start on: *starting AoI gene* set to 0 implies the first plot atom in the combination is taken from the blue AoI. The sequence of values in the *vector of genes for AoI change* indicates the points in the combination when change to plot atoms from a different AoI occurs. Two initial 0 s indicate the first three plot atoms come from the blue AoI, the first 1 indicates that the combination jumps to plot atoms from a different AoI. Which AoI is chosen is determined by the first value in the *vector of genes for number of AoIs to skip on change*. As this value is currently set to 1, the change shifts to the green AoI, which immediately follows the blue one in the relative order. The next 0 in the *vector of genes for AoI change* results in two plot atoms taken from the green AoI. The 1 after that in *vector of genes for AoI change* forces another change of AoI. Again, a 1 in the *vector of genes for number of AoIs to skip on change* shifts the focus to the yellow AoI, which immediately follows the green one in the relative order. At that point a 1 in the *vector of genes for AoI change* results in a single plot atom from the yellow AoI, followed by another AoI change. This time a value of 2 in the *vector of genes for number of AoIs to skip on change* means that the blue AoI—which would follow the yellow one in the relative order since there are no more AoIs to consider—is skipped, and the focus shifts to the green AoI which appears after the blue. The pattern is repeated till the exhaustion of the gene vectors. It is important to note that the genes in the *vector of genes for number of AoIs to skip on change* are not alligned with specific positions in the combination. These genes are only activated when a gene in the *vector of genes for AoI change* is set to 1.

### 3.4.2 Genetic representation for character fusion

A genetic representation of the way in which characters from different AoIs are combined would need to represent which of the terms used for entities in the story is assigned to each variable for entities that participates in each of the AoIs involved.

One important difference with the original representation by Gervás [14] is that the plot templates used there only considered characters as elements to be instantiated, whereas AoIs consider three sorts of elements to be instantiated: characters, objects and locations.

We represent this information in terms of three different vectors—one for each different sort of entity: characters, objects or locations—that define how the entity roles for the different AoIs are instantiated by the entity names that appear in the final story draft. Within the restricted set of variables of a given sort, the procedure for instantiation, initial assignment, mutation and cross over is the same, but applied to the corresponding set of entities.

The set of possible entities for the complete story is defined by the unions of the sets of variable names for the each of the three sorts that appear in each AoI. These variable names need to be distinct across the different AoIs to avoid confusion. This is a challenge because, for instance, many of the definitions for AoIs identify a particular entity as “the hero”, and most of the AoIs dealing with romantic liaisons include variables for “lover” and “beloved”. To avoid this problem, the name of the AoI is assigned as a prefix to all the variable names that feature in that AoI.

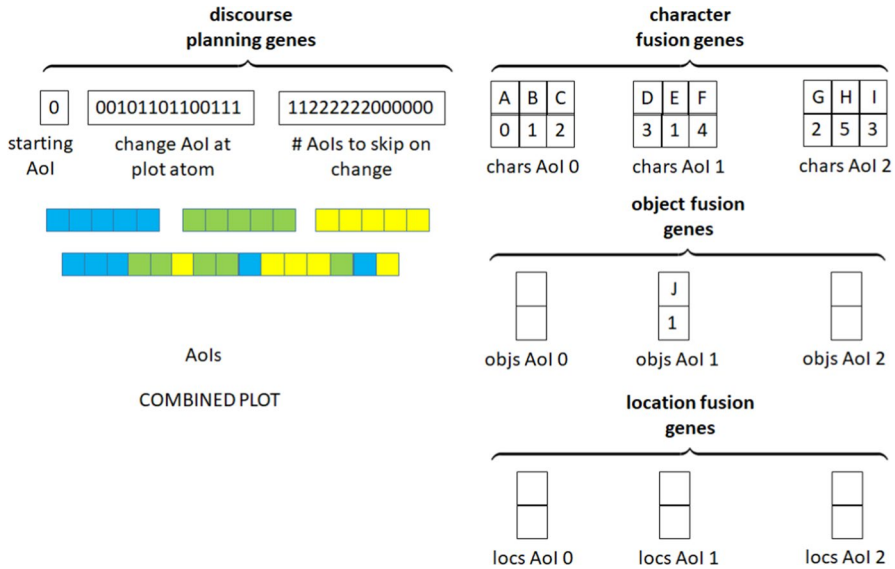
The characterisation of the choices for entity fusion for a given story candidate requires an assignment of entity names to each of the variables in the joint set of variables for the story. This is applied separately for each sort of entity.

For simplicity, the set of potential entity names of each sort for the story is defined to be the set of integers from 0 to  $N$ , with  $N$  being the cardinality of the joint set of variables of that sort for the story. To avoid confusion across sorts, entity names for a particular sort are assigned distinguishing prefix:  $C$  for characters,  $O$  for objects,  $L$  for locations. This is sufficient to represent any choices made in terms of entity fusion (with variables in two different positions in the name-assignment vector being assigned to the same–prefixed–integer). The form of the resulting stories would be significantly improved by a later stage of transforming these integer-based names for the characters into strings representing realistic names.

An example of the complete genetic representation for story draft is shown in Fig. 2. This representation brings together the genes for representing discourse planning and the genes for representing character fusion.

### 3.4.3 Combination-specific nature of the genetic representation

The operational details of the problem of constructing a story for a given combination of AoIs is greatly determined by the particular set of AoIs taken as input, because the required genetic representation will differ based on the specific lengths of the AoIs involved and the number of character roles that each AoI contributes to the story. The genetic representation for combining a particular set of AoIs is tailored to that specific set of AoIs, and it will be incompatible with the genetic

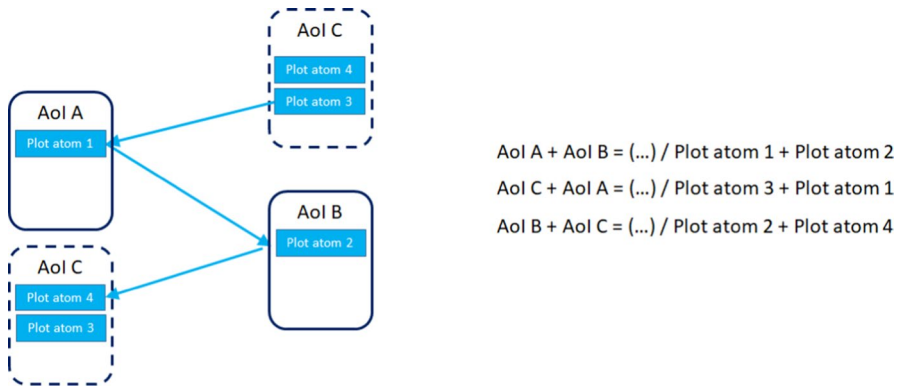


**Fig. 2** Genetic representation for a combination of three AoIs of length 5, each with 3 characters, only one has an object, and none have locations. Fuses characters B (AoI0) / E (AoI1), C (AoI0) / G (AoI2) and D (AoI1) / I (AoI3). Empty boxes have been added to the figure to indicate the position that would have been occupied by information on objects and locations for the AoIs that do not have them

representation for a different combination of AoIs, no matter how similar. This is because the genetic representation includes explicitly the set of AoIs, and the AoIs present will determine the configuration of the gene vectors for representing when to change AoI—the length of which will be determined by the sum of the lengths of the AoIs involved—and the gene vectors for character/object/location fusion, which will have not only their length tailored to the number of instances of each sort, but also the specific names of the elements of each sort that appear in each of the AoIs involved. Genetic representations for different combinations of AoIs are not mutually compatible. This is a shortcoming of the proposed representation that we hope to address in future work. For a particular problem of combining  $N$  AoIs, the length of the final discourse is determined by the total number of scenes in the AoIs being considered, and the maximum number of possible entities featuring in the story is determined by the union of the sets of entities in the AoIs being considered.

### 3.5 Fitness functions

In order to apply the metrics, the genetic representation described needs to be applied to inform a process of construction of story drafts. Then the set of metrics available for all possible combinations of the AoIs in the draft is applied to the resulting story draft. The overall score for a given individual in the population is computed as the average of the scores assigned to the corresponding story draft by the metrics for each of the possible pairwise combinations of the AoIs included in it.



**Fig. 3** Combination of three AoIs with incompatible sequencing constraints. Constraints on character fusion are omitted from the entries shown for the metrics for clarity

**Table 3** Example of data structure of constrained levels for axes of interest for HAPPYLOVE, UNRELENTINGGUARDIAN and TASK, showing relative levels of compatibility between their plot atoms as established by the available metrics

(1.)	HAPPYLOVE		
	{BoyMeetsGirl, FallInLove, HappyEverAfter}		
(2.)	HAPPYLOVE	RELENTINGGUARDIAN	
	{BoyMeetsGirl, FallInLove}	{}	
	{}	{CoupleWantsToMarry, UnrelentingGuardian, RelentingGuardian, Wedding}	
	{HappyEverAfter}	{}	
(3.)	HAPPYLOVE	RELENTINGGUARDIAN	TASK
	{BoyMeetsGirl, FallInLove}	{}	{}
	{}	{CoupleWantsToMarry, UnrelentingGuardian}	{}
	{}	{}	{DifficultTask, Solution}
	{}	{RelentingGuardian, Wedding}	{}
	{HappyEverAfter}	{}	{}

The numbers in the left hand column indicates progress through the construction process as each subsequent AoI is added to the data structure

### 3.6 Selecting sets of compatible AoIs

Because the sequencing constraints for pairs of AoIs force particular positions of atoms in one AoI with respect to the atoms in the other, it is possible that a combination of more than two AoIs prove to be incompatible. This happens for instance if one constraint fixes the position between AoIs A and B, then another constraints forces elements from AoI C to be beyond A and a further constraint forces elements from AoI C to be before B. This type of situation is illustrated graphically in Fig. 3.

This is not a big problem for the procedure we are proposing, because the fitness function will take the average of scores arising from the metrics associated with each of the possible pairings of all the AoIs involved. In cases of conflicts between incompatible AoIs, the overall score will include at least one very low score for one of the conflicting pairings, and so it will always rank lower than other combinations where the AoIs involved are fully compatible. However, in these situations the averaging procedure will result in a score with a relatively low upper threshold, which corresponds to combinations of AoIs that do not quite make sense. In order to



improve the quality of the set of outputs as a whole, we will apply a preprocessing stage to filter the sets of AoIs to be used as input down to sets that are known to be reasonably compatible.

To this end, we have developed a process that constructs from a given set of AoIs a data structure that represents the relative orderings among the plot atoms in the combination as imposed by the corresponding set of constraints. An example for such a data structure of constrained levels for the combination of axes of interest for HAPPYLOVE, UNRELENTINGGUARDIAN and TASK, as used in the example shown in Table 2 above, is shown in Table 3. The example shows how the data structure is progressively built by successive addition of each AoI in the combination.

This data structure is built incrementally by progressively adding the AoIs in the set, and for each addition applying the constraints for the combination of the new AoI with all the AoIs already in the structure. The structure is built of ordered lists of lists of plot atoms. Each of these lists of lists represents a *constrained level*, which contains plot atoms from the different AoIs—one constituent list for the contributions from a particular AoI—on which there is no relative ordering constraint. When the application of a new sequencing constraint establishes a relative ordering between elements from different AoIs that appear at the same constrained level, a new constrained level is built, and the corresponding contributions from the related AoIs are split into two and separated into the resulting two levels. For the example in Table 3, initially (1.) all the plot atoms for the HAPPYLOVE AoI appear as a single list of plot atoms at the first level. When the RELENTINGGUARDIAN AoI is added (2.), the sequencing constraint between FallInLove and CoupleWantsToMarry forces the CoupleWantsToMarry onto a second level, and the sequencing constraint between Wedding and HappyEverAfter forces the HappyEverAfter onto a third level. The levels are represented as rows in the table, with the columns used to keep separate plot atoms from different AoIs. When the TASK AoI is added (3.), the sequencing constraint between UnrelentingGuardian and DifficultTask forces the DifficultTask onto a third level, and the sequencing constraint between Solution and RelentingGuardian forces the RelentingGuardian onto a fourth level. As a result of these shifts, the HappyEverAfter plot atom has been forced onto a fifth level.

The data structure of constrained levels has the advantage that any instances of incompatibility for a particular combination of AoIs result in the plot atoms for one of the AoIs appearing in the data structure out of sequence. This is easy to spot and it allows such problematic combinations to be filtered out as potential inputs.

A set of AoIs needs to be connected for the process to make sense because otherwise the metrics will not be able to score the resulting drafts. Two AoIs are connected in a given draft of a story if there is at least one character that, as a result of character fusion, appears in both AoIs. The minimum requirement is that every AoI in the starting set be at least connected to another AoI in the set, and that all other AoIs in the set can be reached by traversing the connections from the given AoI.

The basic procedure for building a population of drafts for a given combination of AoIs is as follows:

1. A given AoI to act as seed is provided as input, together with the number of AoIs that the combination should have,
2. A set of AoIs that are connected—at least indirectly—to the seed AoI is compiled,
3. The combination is validated in terms of compatibility—if incompatible the compilation is redone to try with a new set of connected AoIs—,
4. An evolutionary process is launched on the given combination.

### 3.7 Constructing an initial population

An initial population of story candidates is built by assigning values to the representation described in Sect. 3.4. For each of the different parts of the representation the process of assignment of values needs to be treated differently.

For the initial digit that defines which AoI to start on, and for the vector of decisions on whether to switch, random choice between 0 and 1 is suitable.

For the vector of decisions on skip size at each switch, random choice between 1 and  $N-1$  (with  $N$  the total number of plots being combined) is suitable.

For the vector of decisions of which entity to assign to each variable, the choice is more complex. This is because variables from the same AoI should not be assigned to the same entity, at the risk of confusing the relations between entities in the corresponding subplot. The process of assignment is carried out separately for the set of variables of each sort for each AoI. For such a set of variables, the process decides at random whether to assign to each variable either an entity name chosen at random from those of the same sort already used in some of the AoIs already processed, or an entirely new entity name chosen at random from the entity names that remain free. This ensures the required constraints are satisfied.

### 3.8 Evolutionary operators

Once a population has been constructed, mutation and cross over operators are applied to it.

Because of the different nature of the various parts of the representation, specific operators of each kind are applied to the different parts.

For the mutation operators:

- For the starting point gene, the value is mutated at random
- For the switch point vector, values at a single point chosen at random are mutated
- For the skip size vector, values at a single point chosen at random are mutated to a value chosen at random within the required range
- For the entity assignment vectors, entity names at each point are either mutated or not depending on a threshold parameter, and, if required, mutated to an entity name chosen at random within the required range

For the cross over operators:

- For the starting point gene, the value of the two individuals being considered is swapped
- For the switch point vector, a point in the vector is chosen at random and the corresponding halves of the vectors for the two individuals are swapped over
- For the skip size vector, a point in the vector is chosen at random and the corresponding halves of the vectors for the two individuals are swapped over
- For the entity assignment vector, the assignments of entities for the two different individuals are swapped over (specific operators are defined for each sort of entity)

### 3.9 Textual rendering for story drafts

The data structures on which the system relies for representing stories—as shown in the examples above—are appropriate for capturing the features that have been considered relevant, but they are not necessarily very user friendly as means of conveying the stories to readers not familiar with the formalism. To facilitate the task of human volunteers charged with providing an evaluation of the acceptability of the stories, a module has been added to the system to produce textual renderings of the resulting story drafts.

The textual rendering module performs four basic tasks: it compiles the set of constants used to refer to the entities that appear in the plot atoms of the final discourse, it assigns to each constant a proper name applicable to a person, it assigns to each plot atom in a story draft a String template that conveys the meaning of the plot atom as a natural language sentence—with place holder tokens for the constants used in the plot atom—and it replaces the place holder tokens for constants in these String templates with the corresponding proper name. The result of this process is a sequence of pseudo-sentences that provide a textual rendering of the discourse for the story draft. The sentences in this textual rendering are repetitive because they refer to all character by a proper name at all mentions, but they are easier for the untrained eye to read than the raw data structures.

## 4 Results and discussion

The results of the proposed system are presented and the relation of the proposed approach with previous work is discussed.

### 4.1 Configuration of a system run

The ability of the system as described to generate acceptable stories is tested in different set ups to explore the impact of the choice of seed AoI and the number AoIs employed in the input. The system has therefore been run with combinations of 3, 4 and 5 AoIs, starting in each case from a different seed AoI.

Because exhaustive testing over the set of 19 AoIs yields a substantial volume of outputs, the initial tests have been run using a limited selection of AoIs as seeds

for the generation process. The AoIs selected to be used as seeds for the generation process are: ABDUCTION, DONOR, RIVALRY and SHIFTINGLOVE. These AoIs have been selected attempting to cover the different kinds of AoI that are present in the set. ABDUCTION represents a classic villainy often used to trigger traditional stories, DONOR represents the donor sequence in Propp's formalism—namely, the hero meets someone who tests him and, on a successful outcome, provides him with a magic object that will help him to achieve his goals, RIVALRY represents a different mechanism for introducing conflict in a story, and SHIFTINGLOVE introduces a specific plot elements dealing with romance—an existing love affair goes through difficult times but eventually succeeds.

For each generation run, a set of additional AoIs is chosen at random—but restricted to AoIs that are connected to the chosen seed—to establish the AoIs that will be included in the combination. The number of AoIs to include in the combination is received as an additional parameter for the generation process.

The proposed system is run in each case with an initial population of 100 individuals generated at random, with the described operators for mutation (probability of mutation set to 0.2) and cross over (probability of cross over set to 0.05), for 100 generations. At each generation populations are culled by selecting the next generation using a best scoring criterion.

In the experiments reported here, each seed AoI of the selected set has been tested in combinations of 3, 4 and 5 AoIs.

Over this general set up, two types of evaluation are presented: a qualitative evaluation that analysis some examples of output, and a quantitative evaluation that compares outputs of the evolutionary solution with a randomly instantiated version of the genetic representation.

## 4.2 Qualitative evaluation of selected examples

Some examples of results for the three different lengths of combination are shown below. An attempt has been made to use different AoIs as seed in each case. The examples constitute random samples from the potential search space in the sense that they are the result of the first successful run for each input configuration. The only exception is where a later run produced a combination too similar to those in the examples generated for different seeds used earlier, in which case a different result was generated to ensure broader coverage of the spectrum of possible stories in the selected examples.

Table 4 shows the result for a combination of three AoIs using ABDUCTION as seed AoI. The two AoIs chosen at random for combination are UNRELENTINGGUARDIAN and HAPPYLOVE. This example shows that, in spite of having tailored the metrics to capture basic compatibility constraints as found in traditional stories, the combinations produced by the system do not always match traditional expectations. In this example, the guardian opposing the union of his ward with a suitor kidnaps the suitor, who then fights him successfully before being rescued by someone else; and this leads to the guardian relenting and allowing the proposed union.

**Table 4** Example of story draft for a basic plot combining three axes of interest, using ABDUCTION as input seed and adding UNRELENTINGGUARDIAN and CONFLICT as random-connected-extensions

Discourse structure			
UNRELENTINGGUARDIAN	CoupleWantsToMarry	<i>lover = Krull</i>	
		<i>beloved = Alan</i>	
UNRELENTINGGUARDIAN	UnrelentingGuardian	<i>lover = Krull</i>	
		<i>beloved = Alan</i>	
		<i>guardian = West</i>	
ABDUCTION	Abduction	<i>abducted = Krull</i>	
		<i>abductor = West</i>	
CONFLICT	Struggle	<i>attacker = Krull</i>	
		<i>defender = West</i>	
CONFLICT	Victory	<i>winner = Krull</i>	
		<i>loser = West</i>	
ABDUCTION	Rescue	<i>abducted = Krull</i>	
		<i>rescuer = Mina</i>	
		<i>abductor = West</i>	
UNRELENTINGGUARDIAN	GuardianRelents	<i>lover = Krull</i>	
		<i>beloved = Alan</i>	
		<i>guardian = West</i>	
UNRELENTINGGUARDIAN	Wedding	<i>lover = Krull</i>	
		<i>beloved = Alan</i>	
HAPPYLOVE	HappyEverAfter	<i>lover = Krull</i>	
		<i>beloved = Alan</i>	
Fitness scoring			
ABDUCTION+RELENTINGGUARDIAN			
RS	Abduction victim = RelentingGuardian hero	100 %	50
	Abduction villain = RelentingGuardian obstacle	100 %	
SQ	UnrelentingGuardian < Abduction	100 %	50
	Rescue < RelentingGuardian	100 %	
ABDUCTION+CONFLICT			
RS	Abduction villain = Conflict villain	100 %	50
SQ	Abduction < Struggle	100 %	50
	Victory < Rescue	100 %	
CONFLICT+RELENTINGGUARDIAN			
RS	Conflict hero = RelentingGuardian hero	100 %	50
SQ	UnrelentingGuardian < Struggle	100 %	50
	Victory < RelentingGuardian	100 %	

**Textual Rendering** Krull wants to marry Alan. Krull finds proposed union with Alan opposed by guardian West. West kidnaps Krull. Krull fights with West. Krull achieves victory over West. Mina rescues Krull from West. Krull vanquishes the opposition of West to the proposed union with Alan. Krull marries Alan

The top part of the table shows the structure of the discourse following the conventions used in Table 2. The middle part of the table shows the scores assigned by the metrics to each pair of Aols in the combination, where RS refers to role sharing, SQ to sequencing. The bottom part of the table presents the textual rendering for the story draft

**Table 5** Example of story draft for a basic plot combining four axes of interest, using SHIFTINGLOVE as input seed and adding RIVALRY, VALIDATOR and RAGS2RICHES as random-connected-extensions

Discourse structure			
SHIFTINGLOVE	GirlMeetsOtherBoy	<i>previous-love= Alan</i> <i>girl = Benson</i> <i>boy = Hans)</i>	
RIVALRY	Rivalry	<i>rival1 = Alan</i> <i>rival2 = Benson</i>	
SHIFTINGLOVE	LoveShift	<i>lover = Benson</i> <i>beloved = Hans</i> <i>rival = Alan</i>	
RAGS2RICHES	Poverty	<i>sufferer = Alan</i>	
RAGS2RICHES	Aspiration	<i>aspirer = Alan</i>	
RAGS2RICHES	Transformation	<i>transformed = Alan</i>	
VALIDATOR	Tested	<i>tested = Alan</i> <i>tester = Benson</i>	
RIVALRY	Cooperation	<i>rival1 = Alan</i> <i>rival2 = Benson</i>	
VALIDATOR	Character'sReaction	<i>tested = Alan</i> <i>tester = Benson</i>	
RIVALRY	RivalReconciliation	<i>rival1 = Alan</i> <i>rival2 = love-interest</i>	
VALIDATOR	Validation	<i>validated = Alan</i> <i>validator = Benson</i>	
VALIDATOR	ValidationRecognised	<i>validated = Alan</i>	
RAGS2RICHES	AspirationAchieved	<i>achiever = Alan</i>	
SHIFTINGLOVE	Reconciliation	<i>lover = Benson</i> <i>beloved = Alan</i>	
Fitness scoring			
RIVALRY+SHIFTINGLOVE			
RS	Rivalry shadow = ShiftingLove love-interest	100 %	50
	Rivalry hero = ShiftingLove hero	0 %	
SQ	Rivalry < LoveShift	100 %	50
	RivalReconciliation < Reconciliation	100 %	
RAGS2RICHES+SHIFTINGLOVE			
RS	Rags2Riches hero = ShiftingLove former	100 %	50
SQ	Aspiration < GirlMeetsOtherBoy	71 %	50
	LoveShift < Transformation	100 %	
	AspirationAchieved < Reconciliation	100 %	
RAGS2RICHES+VALIDATOR			
RS	Rags2Riches hero = Validator hero	100 %	50
SQ	ValidationRecognised < AspirationAchieved	100 %	
SHIFTINGLOVE+VALIDATOR			
RS	ShiftingLove hero = Validator hero	0 %	

**Table 5** (continued)

Fitness scoring			
SQ	LoveShift < Tested	100 %	50
	ValidationRecognised < Reconciliation	100 %	
RAGS2RICHES+RIVALRY			
RS	Rags2Riches hero = Rivalry hero	100 %	50
SQ	Transformation < Cooperation	100 %	
RIVALRY+VALIDATOR			
RS	Rivalry hero = Validator hero	100 %	50
SQ	Cooperation < Character'sReaction	100 %	

**Textual Rendering** Benson who loved Alan meets different person Hans. Alan develops rivalry with Benson. Benson no longer cares for Alan and now loves Hans. Alan suffers poverty. Alan has aspiration. Alan is transformed. Alan is tested by Benson. Alan cooperates with Benson. Alan reacts to the test by Benson. Alan ends rivalry with Benson. Alan is validated by Benson. Alan sees validation recognised. Alan fulfills their aspiration. Benson makes up with Alan

The top part of the table shows the structure of the discourse following the conventions used in Table 2. The middle part of the table shows the scores assigned by the metrics to each pair of AoIs in the combination. The bottom part of the table presents the textual rendering for the story draft

Table 5 shows the result for a combination of four AoIs using SHIFTINGLOVE as seed AoI. The three AoIs chosen at random for combination are RIVALRY, VALIDATOR and RAGS2RICHES. It is interesting to see in this case, that, although the score (86) is not 100 %, the result is quite acceptable. In fact, some of the constraints that are not satisfied—like not having the protagonist of the SHIFTINGLOVE AoI also be the protagonist of the VALIDATOR AoI—yield interesting results—in this case, having the validation of the former lover play a role in the following reconciliation.

Table 6 shows the result for a combination of four AoIs using CROSSDRESSING as seed AoI. The four AoIs chosen as random—connected—extensions are RAGS2RICHES, TASK, ABDUCTION and SHIFTINGLOVE. In this case the score is even lower (76) and the result is still acceptable. Transgressions of the expected combination patterns include: the protagonist of the story, Lilly, is the person who sets the task rather than the person trying to solve it—the solving of the task becomes the context in which Lilly's adventures take place—, the protagonist's partner shifts their romantic interest to the person that Lilly has charged with solving the task, the person in charge of solving the task commits a villainy—a kidnapping—, Lilly disguises herself as a man to rescue the victim, she also achieves her aims, the task gets solved and Lilly recovers her lover. It is interesting to note that the increase in the number of AoIs involved in the combination increases very significantly the number of constraints that need to be considered. This in its turn leads to a lower overall score, as it becomes more difficult for all the constraints to be satisfied at the same time. However, the satisfaction of those constraints that do hold contributes to the overall appearance of coherence of the final story.

**Table 6** Example of story draft for a basic plot combining five axes of interest, using CROSSDRESSING as input seed and adding RAGS2RICHES, TASK, ABDUCTION and SHIFTINGLOVE as random-connected-extensions

Discourse structure			
TASK	DifficultTask	<i>setter= Lilly</i>	
		<i>solver = Sarah</i>	
RAGS2RICHES	Poverty	<i>sufferer = Lilly</i>	
SHIFTINGLOVE	GirlMeetsOtherBoy	<i>previous-love= Lilly</i>	
		<i>girl = Ernest</i>	
		<i>boy = Sarah</i>	
RAGS2RICHES	Aspiration	<i>aspirer = Lilly</i>	
SHIFTINGLOVE	LoveShift	<i>lover = Ernest</i>	
		<i>beloved = Sarah</i>	
		<i>rival = Lilly</i>	
RAGS2RICHES	Transformation	<i>transformed = Lilly</i>	
ABDUCTION	Abduction	<i>abducted = Hugo</i>	
		<i>abductor = Sarah)</i>	
CROSSDRESSING	CrossDressing	<i>cross-dresser = Lilly</i>	
ABDUCTION	Rescue	<i>abducted = Hugo</i>	
		<i>rescuer = Lilly</i>	
		<i>abductor = Sarah</i>	
CROSSDRESSING	Recognition	<i>recognised = Lilly)</i>	
RAGS2RICHES	AspirationAchieved	<i>(achiever = Lilly</i>	
TASK	Solution	<i>(solver = Sarah)</i>	
SHIFTINGLOVE	Reconciliation	<i>lover = Ernest</i>	
Fitness scoring			
ABDUCTION+RAGS2RICHES			
RS	=is also Rags2Riches hero	100 %	50
SQ	Transformation < Rescue	100 %	
CROSSDRESSING+SHIFTINGLOVE			
RS	CrossDressing someone = ShiftingLove hero	100 %	50
SQ	LoveShift < CrossDressing	100 %	
	Recognition < Reconciliation	100 %	
RAGS2RICHES+SHIFTINGLOVE			
RS	Rags2Riches hero = ShiftingLove former	100 %	50
SQ	Aspiration < GirlMeetsOtherBoy	92 %	50
	LoveShift < Transformation	100 %	
	AspirationAchieved < Reconciliation	100 %	
ABDUCTION+SHIFTINGLOVE			
RS	Abduction victim = ShiftingLove hero	0 %	50
	Abduction villain = ShiftingLove shadow	0 %	
SQ	Abduction < LoveShift	84 %	50
	Rescue < Reconciliation	100 %	



**Table 6** (continued)

Fitness scoring			
CROSSDRESSING+RAGS2RICHES			
RS	CrossDressing someone = Rags2Riches hero	100 %	50
SQ	Aspiration < CrossDressing	100 %	
	CrossDressing < Transformation	84 %	
	Transformation < Recognition	100 %	
	Recognition < AspirationAchieved	100 %	
ABDUCTION+CROSSDRESSING			
RS	Abduction hero = CrossDressing someone	100 %	
SQ	Abduction < CrossDressing	84 %	50
	Rescue < Recognition	100 %	
ABDUCTION+TASK			
RS	Abduction villain = Task hero	100 %	
SQ	DifficultTask < Abduction	84 %	50
SHIFTINGLOVE+TASK			
RS	ShiftingLove hero = Task hero	0 %	
SQ	LoveShift < DifficultTask	69 %	50
	Solution < Reconciliation	100 %	
RAGS2RICHES+TASK			
RS	Rags2Riches hero = Task hero	0 %	50
SQ	Transformation < Solution	100 %	

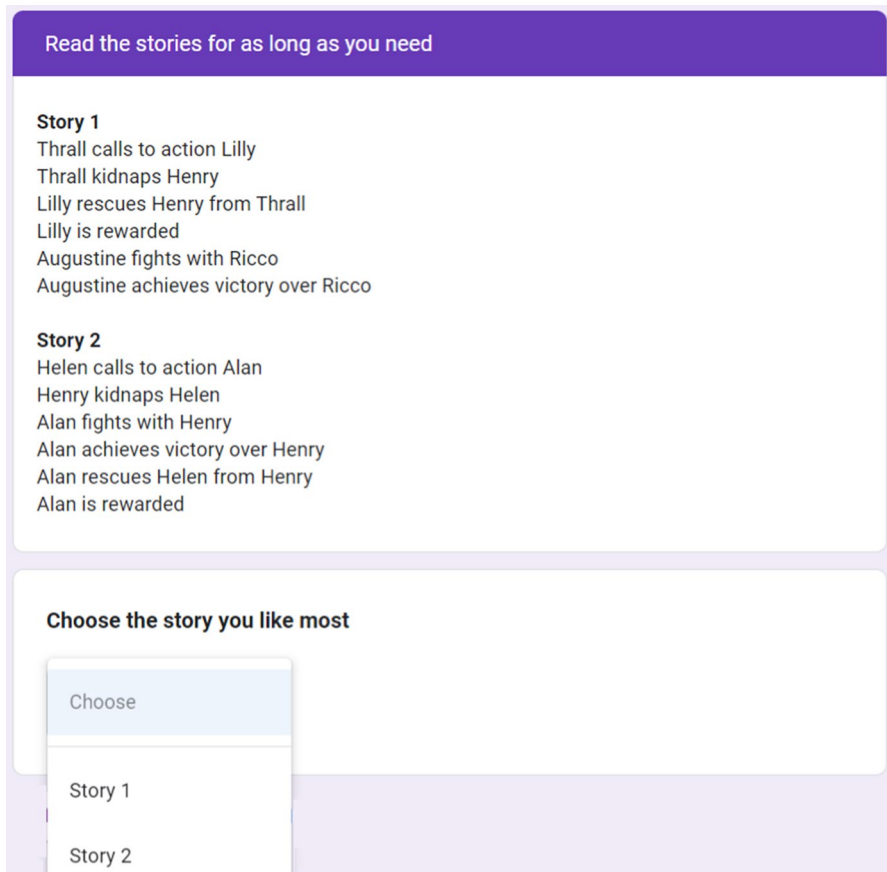
**Textual Rendering** Sarah is set a difficult task by Lilly. Lilly suffers poverty. Ernest who loved Lilly meets different person Sarah. Lilly has aspiration. Ernest no longer cares for Lilly and now loves Sarah. Lilly is transformed. Sarah kidnaps Hugo. Lilly dresses up as a member of the opposite sex. Lilly rescues Hugo from Sarah. Lilly is recognised. Lilly fulfills their aspiration. Sarah solves the task. Ernest makes up with Lilly.

The top part of the table shows the structure of the discourse following the conventions used in Table 2. The middle part of the table shows the scores assigned by the metrics to each pair of AoIs in the combination. The bottom part of the table presents the textual rendering for the story draft

### 4.3 Quantitative comparative evaluation informed by human judgements

To obtain a quantitative measure of the relative quality of the story drafts generated by the proposed evolutionary solution, we carried out a comparative evaluation between the results obtained by the application of the proposed metrics in the evolutionary process and results of a baseline procedure that did not take the proposed metrics into account. The baseline procedure employed relies on the process of random instantiation of the genetic representation used to build the initial populations for the evolutionary procedures. Because that process does not consider the proposed metrics at any point, any observed improvements in quality between the baseline and the outputs of the system should be considered an indication of the added value that the metrics used as fitness functions provide.

To discern between the two competing approaches in terms of perceived quality of the stories we rely on a set of human volunteers that were asked to consider pairs



**Fig. 4** Sample evaluation screen where the user was presented two plots to select the one with the highest perceived quality

of stories and select one of them as more acceptable than the other. Each pair contained—for the same combination of axes of interest—a story draft produced by the evolutionary procedure and a story draft produced by the baseline random instantiation procedure. The pairs were presented in random order to avoid biases arising from the presentation order (see Fig. 4).

This evaluation process is intended to identify whether there is indeed some added value in applying the proposed evolutionary procedure instead of a process of random instantiation. The drafts being considered for evaluation are, in their present form, sketches for the plot for stories. This implies that evaluators are not exercising an already acquired skill, they are actually being asked to develop on the fly a criterion for deciding on these cases when they are presented to them. For this reason, we have decided not to ask evaluators to assign an absolute score to individual drafts, but rather to present them with two contrasting samples and ask them which one they consider preferable. There is however a risk that

**Table 7** Configuration of the evaluation sets used for the evaluation by human judges

AoI	Set 1	Set 2	Set 3
3-Abduction	AB-CA-CO-bl	AB-CA-IV-bl	AB-VP-HL-bl
	AB-CA-CO-ev-99-0	AB-CA-IV-ev-83-0	AB-VP-HL-ev-83-0
3-Donor	DO-IV-VP-bl	DO-SL-IV-bl	DO-TA-RE-bl
	DO-IV-VP-ev-75-0	DO-SL-IV-ev-75-0	DO-TA-RE-ev-100-0
3-Rivalry	RI-RR-IV-ev-100-0	Ri-RR-UG-ev-100-0	RI-TA-PU-ev-100-0
	RI-RR-IV-bl	Ri-RR-UG-bl	RI-TA-PU-bl
3-ShiftingLove	SL-RR-RI-bl	SL-UG-JO-bl	SL-UG-RI-bl
	SL-RR-RI-ev-91-0	SL-UG-JO-ev-66-0	SL-UG-RI-ev-75-1
4-Abduction	AB-JO-PU-RG-ev-97-1	AB-RG-RR-TA-ev-83-0	AB-TA-SL-IV-ev-69-1
	AB-JO-PU-RG-bl	AB-RG-RR-TA-bl	AB-TA-SL-IV-bl
4-Donor	DO-CA-CO-RG-ev-99-1	DO-HL-RI-CD-ev-75-0	DO-IV-CA-CO-ev-99-0
	DO-CA-CO-RG-bl	DO-HL-RI-CD-bl	DO-IV-CA-CO-bl
4-Rivalry	RI-AB-VP-DO-ev-87-0	RI-PU-RG-CO-ev-98-0	RI-RG-JO-CA-ev-100-1
	RI-AB-VP-DO-bl	RI-PU-RG-CO-bl	RI-RG-JO-CA-bl
4-ShiftingLove	SL-AB-DO-CD-bl	SL-HL-CD-RR-bl	SL-VA-CA-RR-bl
	SL-AB-DO-CD-ev-66-0	SL-HL-CD-RR-ev-76-1	SL-VA-CA-RR-ev-82-0
5-Abduction	AB-CA-CO-SL-VA-bl	AB-CD-JO-RI-VI-bl	AB-RG-RR-PU-CD-bl
	AB-CA-CO-SL-VA-ev-75	AB-CD-JO-RI-VI-ev-78	AB-RG-RR-PU-CD-ev-81
5-Donor	DO-CO-VA-SL-HL-ev-78	DO-RG-RR-IV-JO-ev-93	DO-TA-VA-AB-CA-ev-82
	DO-CO-VA-SL-HL-bl	DO-RG-RR-IV-JO-bl	DO-TA-VA-AB-CA-bl
5-Rivalry	RI-CO-RR-SL-CA-bl	RI-HL-SL-UG-VP-bl	RI-VA-VP-PU-IV-bl
	RI-CO-RR-SL-CA-ev-77-0	RI-HL-SL-UG-VP-ev-73-2	RI-VA-VP-PU-IV-ev-82-0
5-ShiftingLove	SL-CA-DO-AB-JO-ev-69-0	SL-CA-DO-RI-VP-ev-64-2	SL-CO-JO-VA-AB-ev-70-0
	SL-CA-DO-AB-JO-bl	SL-CA-DO-RI-VP-bl	SL-CO-JO-VA-AB-bl

evaluators with prior exposure to similar tasks—as maybe the case in researchers working in the field of generation of narrative—may have developed a pre-existing criterion that may introduce a bias. This possibility has been considered in the selection of set of evaluators and it is discussed over the results obtained.

The proposed evaluation is as much a test of the evaluation procedure as a test for the proposed solution. For this reason as well we have decided to restrict the size of the set of evaluators in this initial trial. More refined evaluations may be carried out, conveniently revised based on the insights obtained from this one, as further work.

A set of 10 human volunteers participated in the evaluation, including 7 men and 3 women, with ages ranging from 20 to 60 years old. The level of expertise ranged from Novice to Expert, with 2 considered experts in the field, 3 considered competent, 3 with limited experience and 2 considered novices without any previous experience in narrative generation. An effort has been made to ensure the cohort includes a reasonable variation over the range of variables that may affect the results.

**Table 8** Key for the two-letter labels for the AoIs used in the examples mentioned in other tables

Label	AoI name	Label	AoI name	Label	AoI name
IV	INTERDICTIONVIOLATED	CO	CONFLICT	HL	HAPPYLOVE
AB	ABDUCTION	TA	TASK	SL	SHIFTINGLOVE
VP	VILLAINYPUNISHMENT	PU	PURSUIT	UG	UNRELENTINGGUARDIAN
CA	CALL2ACTIONREWARD	DO	DONOR	RE	REPENTANCE
VA	VALIDATION	RI	RIVALRY	RR	RAGS2RICHES
JO	JOURNEY	CD	CROSSDRESSING		

A set of 36 pairs of plots were generated for the evaluation, each pair consisting of the combination of 3, 4 or 5 AoIs as shown in Table 7. The resulting set was divided into 3 subsets with 12 pairs each, and each of them was evaluated by four evaluators, giving rise to 144 evaluations. For each combination, the name that appears in the tables includes short labels for each of the AoIs involved, according to the key given in Table 8.

The results of this quantitative evaluation are presented grouped by the number of AoIs combined in each case, to allow consideration of the differences in score results arising from the increase in the number of constraints as the number of AoIs rises (see Tables 9 and 11) as well as by evaluator and evaluation set to show possible differences in the distribution of the plot combinations in each evaluation set (see Table 10).

Table 9 shows the decisions made by the evaluators for each pair of plots in each of the 3 evaluation subsets.

The numeric results in Table 10 point out that there is no relationship between the level of expertise and the number of times the evaluators preferred the evolutionary version of the plot over the randomly generated one, as it might have been expected. For example, Evaluator 3 in evaluation set 2 has preferred the evolutionary version of the plots only 50% of the times (6 out of 12), while the other expert evaluators (green cells in Table 10) have chosen the evolutionary version almost in all cases. In contrast, the novice evaluators (red cells in Table 10) have consistently chosen the evolutionary versions most of the times (83.33%—10 out of 12—in the first case, 75%—9 out of 12—in the second).

Table 10 also shows that there was no significant difference in the composition of the three evaluation sets, as the evaluators of the first set chose the evolutionary versions of the stories 68.75% of the times, whereas the percentage in the other two subsets was slightly higher: 72.92%. While there are some stories in each subset that had a unanimous response by all evaluators (e.g. Table 9, rows 4, 6 and 12 in evaluation set 1, where all the evaluators chose the evolutionary version as more acceptable than the baseline), or almost unanimous (e.g. Table 9, first three rows in evaluation set 1), evaluation set 1 also shows that other combinations were not so clearly preferable in the evolutionary version (e.g. Table 9, rows 5, 8, 10 and 11) or were definitely worse (e.g. Table 9, row 9, with 3 evaluators choosing the baseline version over the evolutionary one). This explains the slight difference in the results of the

**Table 9** Results of the human judgments on the comparison between story drafts produced by the evolutionary procedure and story drafts produced by random instantiation (Color figure online)

Evaluation Set 1	Expected	Ev1	Ev2	Ev3	Ev4
AB-CA-CO-bl	Story 2	Story 2	Story 2	Story 2	Story 1
AB-CA-CO-ev-99-0		Story 2	Story 1	Story 2	Story 2
DO-IV-VP-bl	Story 2	Story 2	Story 1	Story 2	Story 2
DO-IV-VP-ev-75-0		Story 1	Story 1	Story 1	Story 2
RI-RR-IV-ev-100-0	Story 1	Story 1	Story 1	Story 1	Story 2
RI-RR-IV-bl		Story 2	Story 2	Story 2	Story 2
SL-RR-RI-bl	Story 2	Story 2	Story 2	Story 2	Story 2
SL-RR-RI-ev-91-0		Story 1	Story 2	Story 1	Story 2
AB-JO-PU-RG-ev-97-1	Story 1	Story 1	Story 2	Story 1	Story 2
AB-JO-PU-RG-bl		Story 1	Story 1	Story 1	Story 1
DO-CA-CO-RG-ev-99-1	Story 1	Story 1	Story 1	Story 1	Story 1
DO-CA-CO-RG-bl		Story 1	Story 1	Story 1	Story 2
RI-AB-VP-DO-ev-87-0	Story 1	Story 1	Story 1	Story 1	Story 2
RI-AB-VP-DO-bl		Story 2	Story 1	Story 1	Story 2
SL-AB-DO-CD-bl	Story 2	Story 1	Story 1	Story 2	Story 2
SL-AB-DO-CD-ev-66-0		Story 1	Story 1	Story 2	Story 1
AB-CA-CO-SL-VA-bl	Story 2	Story 1	Story 1	Story 2	Story 1
AB-CA-CO-SL-VA-ev-75		Story 1	Story 1	Story 2	Story 1
DO-CO-VA-SL-HL-ev-78	Story 1	Story 1	Story 2	Story 1	Story 2
DO-CO-VA-SL-HL-bl		Story 2	Story 2	Story 2	Story 1
RI-CO-RR-SL-CA-bl	Story 2	Story 2	Story 2	Story 1	Story 1
RI-CO-RR-SL-CA-ev-77-0		Story 1	Story 1	Story 1	Story 1
SL-CA-DO-AB-JO-ev-69-0	Story 1	Story 1	Story 1	Story 1	Story 1
SL-CA-DO-AB-JO-bl		Story 1	Story 1	Story 1	Story 1
Evaluation Set 2	Expected	Ev1	Ev2	Ev3	Ev4
AB-CA-IV-bl	Story 2	Story 2	Story 1	Story 1	Story 2
AB-CA-IV-ev-83-0		Story 2	Story 2	Story 1	Story 2
DO-SL-IV-bl	Story 2	Story 2	Story 2	Story 1	Story 2
DO-SL-IV-ev-75-0		Story 1	Story 2	Story 1	Story 1
RI-RR-UG-ev-100-0	Story 1	Story 2	Story 1	Story 1	Story 1
RI-RR-UG-bl		Story 2	Story 1	Story 1	Story 2
SL-UG-JO-bl	Story 2	Story 1	Story 2	Story 1	Story 2
SL-UG-JO-ev-66-0		Story 1	Story 1	Story 2	Story 1
AB-RG-RR-TA-ev-83-0	Story 1	Story 1	Story 1	Story 2	Story 1
AB-RG-RR-TA-bl		Story 1	Story 1	Story 2	Story 1
DO-HL-RI-CD-ev-75-0	Story 1	Story 2	Story 1	Story 2	Story 1
DO-HL-RI-CD-bl		Story 1	Story 1	Story 1	Story 1
RI-PU-RG-CO-ev-98-0	Story 1	Story 1	Story 1	Story 1	Story 1
RI-PU-RG-CO-bl		Story 2	Story 1	Story 2	Story 2
SL-HL-CD-RR-bl	Story 2	Story 2	Story 1	Story 2	Story 2
SL-HL-CD-RR-ev-76-1		Story 2	Story 2	Story 2	Story 2
AB-CD-JO-RI-VP-bl	Story 2	Story 2	Story 2	Story 2	Story 2
AB-CD-JO-RI-VP-ev-78		Story 1	Story 1	Story 1	Story 1
DO-RG-RR-IV-JO-ev-93	Story 1	Story 1	Story 1	Story 1	Story 1
DO-RG-RR-IV-JO-bl		Story 2	Story 1	Story 2	Story 2
RI-HL-SL-UG-VP-bl	Story 2	Story 1	Story 2	Story 2	Story 1
RI-HL-SL-UG-VP-ev-73-2		Story 1	Story 1	Story 1	Story 2
SL-CA-DO-RI-VP-ev-64-2	Story 1	Story 1	Story 1	Story 2	Story 1
SL-CA-DO-RI-VP-bl		Story 1	Story 1	Story 1	Story 1
Evaluation Set 3	Expected	Ev1	Ev2	Ev3	Ev4
AB-VP-HL-bl	Story 2	Story 1	Story 2	Story 2	Story 2
AB-VP-HL-ev-83-0		Story 2	Story 2	Story 2	Story 1
DO-TA-RE-bl	Story 2	Story 2	Story 2	Story 1	Story 1
DO-TA-RE-ev-100-0		Story 1	Story 1	Story 1	Story 1
RI-TA-PU-ev-100-0	Story 1	Story 1	Story 1	Story 1	Story 1
RI-TA-PU-bl		Story 2	Story 1	Story 2	Story 1
SL-UG-RI-bl	Story 2	Story 1	Story 1	Story 2	Story 1
SL-UG-RI-ev-75-1		Story 1	Story 1	Story 1	Story 1
AB-TA-SL-IV-ev-69-1	Story 1	Story 2	Story 1	Story 1	Story 1
AB-TA-SL-IV-bl		Story 1	Story 2	Story 1	Story 1
DO-IV-CA-CO-ev-99-0	Story 1	Story 2	Story 1	Story 1	Story 1
DO-IV-CA-CO-bl		Story 1	Story 2	Story 1	Story 1
RI-RG-JO-CA-ev-100-1	Story 1	Story 2	Story 1	Story 1	Story 2
RI-RG-JO-CA-bl		Story 2	Story 1	Story 2	Story 2
SL-VA-CA-RR-bl	Story 2	Story 1	Story 2	Story 2	Story 2
SL-VA-CA-RR-ev-82-0		Story 2	Story 2	Story 2	Story 2
AB-RG-RR-PU-CD-bl	Story 2	Story 2	Story 2	Story 2	Story 2
AB-RG-RR-PU-CD-ev-81		Story 1	Story 1	Story 1	Story 1
DO-TA-VA-AB-CA-ev-82	Story 1	Story 1	Story 1	Story 1	Story 1
DO-TA-VA-AB-CA-bl		Story 2	Story 2	Story 2	Story 2
RI-VA-VP-PU-IV-bl	Story 2	Story 1	Story 2	Story 2	Story 2
RI-VA-VP-PU-IV-ev-82-0		Story 1	Story 1	Story 1	Story 1
SL-CO-JO-VA-AB-ev-70-0	Story 1	Story 1	Story 1	Story 2	Story 1
SL-CO-JO-VA-AB-bl		Story 1	Story 1	Story 1	Story 1

**Table 9** (continued)

The colors in the evaluator columns (Ev1 - Ev4) show the level of expertise of each evaluator: green for expert, yellow for competent, orange for limited experience and red for novice. Cells in bold face highlight matches between the evaluators preferences and the expected selection

**Table 10** Quantitative results of the human judgments on the comparison between story drafts produced by the evolutionary procedure and story drafts produced by random instantiation, per evaluator and evaluation set

Results Evaluation Set 1				Results Evaluation Set 2				Results Evaluation Set 3			
Ev1	Ev2	Ev3	Ev4	Ev1	Ev2	Ev3	Ev4	Ev1	Ev2	Ev3	Ev4
10	7	11	5	8	10	6	11	5	11	10	9
83,33	58,33	91,67	41,67	66,67	83,33	50,00	91,67	41,67	91,67	83,33	75,00
68,75				72,92				72,92			

The colors in the evaluator columns (EV1 - EV4) show the level of expertise of each evaluator: green for expert, yellow for competent, orange for limited experience and red for novice

The values in the table show the number of answers of each evaluator that matched the expected answer (row 1), the percentage of answers of each evaluator that matched the expected result (row 2) and the percentage of answers that matched the expected result for each evaluation set (row 3)

**Table 11** Quantitative results of the human judgments on the comparison between story drafts produced by the evolutionary procedure and story drafts produced by random instantiation, per number of AoI combinations

AoI	#	%	n-AoI %	Overall %
3-Abduction	8	66,67	68,75	71,53
3-Donor	8	66,67		
3-Rivalry	10	83,33		
3-ShiftingLove	7	58,33	70,83	
4-Abduction	8	66,67		
4-Donor	9	75,00		
4-Rivalry	9	75,00	75,00	
4-ShiftingLove	8	66,67		
5-Abduction	9	75,00		
5-Donor	10	83,33		
5-Rivalry	7	58,33		
5-ShiftingLove	10	83,33		

The values in the table show the number of times the evaluators preferred the evolutionary version of the stories for a specific combination of AoIs (column 2) and the corresponding percentage (column 3), the accumulated percentage for the combination of 3, 4 and 5 AoIs (column 4) and the overall preference for the evolutionary versions of the stories (column 5)

three subsets, as in the first subset there are more combinations of the last two cases than in the other two subsets.

As for the results shown in Table 11, two interesting outcomes can be highlighted. The first one is that, counting the number of times the evolutionary version was preferred over the baseline (columns 2 and 3), for each combination of 3, 4 and 5 AoIs, we can see that the percentage of positive responses increases with the number of AoIs that must be combined (column 4, 68.75% for 3 AoIs, 70.83% for 4 AoIs and 75% for 5 AoIs). This means that, as the stories gain complexity, it

is more difficult to generate meaningful stories randomly, so the evolutionary versions are favoured over the baselines. The second outcome is that, out of the 144 evaluated pairs of plots, 71.53% of the times (i.e. 103 out of 144, column 5), the evolutionary versions were considered to have better quality than the baselines. This means that the proposed method to combine subplots generates high quality plots that improve the results provided by the baseline method. Although there is still a wide margin for improvement, the results prove that the proposed method can be successfully used to generate rich, complex stories as the result of combining simpler plots that can be subsequently generated using other, well established methods.

From the point of view of how these quantitative results can be interpreted in terms of the specific details of the proposed solution, there are two aspects worth discussing. The differences between the evaluation sets may be explained in terms of the interaction between two different factors: the likelihood that the random baseline sometimes produce acceptable results, and the possibility that specific combinations of AoIs are ill-suited for being combined together unless specific metrics are added to consider romantic affinities between characters.

Because the method used as baseline is based on random assignment of genetic information, there is a non-zero chance that it lead to acceptable story drafts. The likelihood of this happening is higher for combinations of a small number of AoIs, where the search space in question is smaller. As the number of AoIs involved increases, the size of the corresponding search space increases exponentially so the likelihood of acceptable results being produced by the random procedure is significantly reduced. This explains the results shown in Table 11, where evaluator preference for the evolutionary versions rises with size of the combinations. The same phenomenon also increases the likelihood that the baseline procedure sometimes produce results that compete in quality with those of the evolutionary approach. This may explain some of the irregularities observed in the results in Table 10. This second consequence may be compounded where it interacts with an observed shortcoming of the solution as it stands, involving conflicts between certain types of AoIs that are not captured by the current set of metrics.

A close examination of the specific results produced shows that there are cases where the chosen combination of AoIs—which are selected at random except for the described filter on combinations that imply temporal inconsistencies—suffer from conflicts at a different level. This happens for instance when two AoIs that involve romantic relationships are combined together in the same story. Each of the AoIs in such a case will postulate specific relations of affinity—or lack thereof—between the characters. The current set of metrics does not include constraints on consistency in affinities between the characters through the story. This is because affinities between the characters are not explicitly modelled in the chosen representation. For this reason, story drafts produced by the evolutionary solution in such cases are likely to include inconsistencies in affinities between characters, such as for instance, having character A make up with character B as resolution of a `SHIFTINGLOVE` AoI but then marry character C as resolution to a `RELENTINGGUARDIAN` AoI that has been combined with it. When the outcomes of the evolutionary procedure suffer from this

problem, there is a much higher chance that evaluators prefer the outcome of the random baseline. This problem affects only combinations with more than one AoI that involves romantic relationships, so it does not affect combinations without a love interest subplot, or any that have a single AoI involved in the love interest subplot.

To address this particular problem, a future extension of the proposed solution should include explicit consideration of affinities between characters as an additional feature to consider in the metrics that drive the fitness function.

#### 4.4 Relation with previous work

The knowledge representation required for the problem addressed in this paper differs slightly from the classic concept of a script [33] in two senses. First, that the basic elements considered in the present paper as events correspond to a much larger granularity than that considered in traditional scripts. In fact, each of the events considered in this paper might be aligned with a particular script in the Schankian sense—such as a wedding, or a fight. Second, because the traditional concept of script carries a certain connotation that the events in a script take place not just in order but, to a certain extent, in close temporal proximity, whereas the knowledge representation that we will be considering is explicitly intended to allow capture of long term dependencies across events very distant in time—such as a character at the end of a story returning from a journey started at the beginning of the story.

The representation for plot defined in this paper shares with the concept of narrative schema [4] the fact that they both refer to an ordered sequence of events that arises in the context of a narrative, and that they both include in their specification predicates to define the events and a set of labelled roles to identify the arguments of the predicates. Because the narratives schemas defined by Chambers and Jurafsky are extracted automatically from narrative text, they have a higher likelihood of being considered similar to the proposed schemas of connected events. However, there is a significant difference in that the schemas described by Chambers and Jurafsky are postulated as possible abstractions extracted from text based on a certain probability of recurrence over different narratives, whereas the schemas proposed in the present paper constitute the result of a knowledge engineering effort that defines a set of knowledge resources validated by the judgment of human engineers. This is an important requirement if these resources are to be used as building blocks for generated stories.

The representation of plot in terms of axes of interest had been used before to generate story plot drafts [13]. The procedure employed for combining axes of interest in that instance exhaustively generated all combinations deemed to be valid in terms of whether they matched the probabilities of character continuity across scenes as obtained from a prior corpus. This basically means that two scenes—or plot atoms—are placed contiguously in a candidate story if some character can be found that appears in each of these scenes playing a pair of roles that has been observed before in the corpus. This criterion ensures local consistency, but it has some potential shortcomings. First, it does not take into consideration long ranging connections



across non-contiguous scenes. Second, where more than one character from scene A carries on to scene B, a probability-based criterion may validate both links based on different prior stories, but it will not be able to consider the importance of both links occurring together. The new evolutionary solution, by relying on a fitness function based on metrics built heuristically to capture common sense connections across AoIs, improves upon the original on both of these aspects.

The character fusion operation considered here is comparable to binding between characters as used by [10]. In Fay's work, the units being combined are character threads—which tend to gather together all the events in a story in which a given character participates. The procedure proposed by Fay therefore uses fusion—usually involving secondary characters—to combine together threads for different characters into a more elaborate story. The procedure proposed here differs from that approach in two different senses. First, in that the units being combined here are intended to be schemas that focus on plot-relevant connections across elements. This makes it less likely that elements that are not entirely relevant to a particular plot end up included in a story draft simply by virtue of appearing in an existing character thread for a previous story. Second, because the use of the metrics proposed here increases the probability that the bindings established between characters play a relevant role in the narrative structure of the resulting story draft.

With respect to prior approaches that consider evolutionary solutions for story generation, the proposed solution shares characteristics with some of them, but it can also be improved by enrichment with additional features considered in some of them. The use of metrics designed to ensure story consistency is a characteristic shared with the work of [22], and it may be comparable to the use of the knowledge-based heuristics of the MEXICA knowledge-based story generator as fitness function as used by [35]. Further features that may be considered to improve the quality of system outputs are: some measure of story interest as used by [22], measures of story novelty as used by [11], measures of whether the stories satisfy user established goals—as used by [19]—or a specific curve to describe tensions in the story—as used by [21]. Regardless of these potential extensions, the proposed solution captures typical narrative structures by virtue of the choice of representation units and fitness function metrics.

## 5 Conclusions and future work

A number of conclusions that are relevant for the task of narrative generation are outlined, and some avenues to be explored in further research are described.

### 5.1 Conclusions

The introduction of an intermediate knowledge structure for representing connections between events at smaller granularity than a full plot—discussed here as connected event schemas or axes of interest—has been shown to provide a useful abstraction that allows a successful balance between the challenge of capturing the

connection between events necessary for a draft to make sense while allowing the wide range of patterns of articulation that avoids the repetition of known patterns of plot.

The operation of establishing bindings between variables in different schemas—referred here as character fusion—has proven to be a valuable mechanism in tying together different schemas when they are included in the same draft.

The proposed set of metrics for quality of drafts in terms of event coherence across schemas—used to inform the fitness function in the evolutionary procedure—provides a set of criteria for assessing combinations of schemas for coherence. These criteria rely only on the structure of the draft in terms of its constituent schemas and the bindings between their variables established by character fusion, and in particular they do not rely on the genetic representation of the drafts. This makes them applicable to drafts built by procedures other than the evolutionary, as long as the drafts are constructed as combinations of schemas. This is an important contribution since there is a need in the field to develop not just automated procedures for generating artifacts of a given kind, but also automated procedures for evaluating the quality of the resulting artifacts [5].

The evolutionary approach to constructing plot outlines for stories by combining axes of interest based on metrics for common sense connections between them provides efficient means for building a population of drafts that satisfy constraints on semantic validity over the final linear discourse for the story. Due to the progressive nature of the metrics used as fitness function the population converges reasonably quickly for a low number of constituent axes of interest. It remains to be seen whether the solution will scale well towards higher numbers of constituents.

The quantitative evaluation that was carried out with human evaluators, even though it lacks statistical significance, points out that the proposed metrics to evaluate the quality of the generated drafts work reasonably well, as the evolutionary versions of the drafts were consistently preferred over their randomly generated counterparts. Further work needs to be addressed in order to take into consideration the internal relationships between specific AoIs along the story, but the current solution shows that, in the cases where these relationships are not too strong, the proposed method produces coherent plots, even more as the number of AoIs to be combined increases

## 5.2 Future work

In terms of future work, a number of avenues for possible further research open up as a result of the insights presented in this paper.

The set of AoIs used in the reported experiments is restricted to those originally developed by Gervás [13]. That set of AoIs was designed as a proof of concept for the additional level of intermediate granularity for plot representation, and it was never intended to provide exhaustive coverage of the broad set of possibilities open in terms of possible schemas to be used in plot construction. Now that a successful operational procedure for combining these type of elements has proved successful,

an extension of the set of AoIs would significantly enlarge the range of drafts that can be constructed. Such attempts can draw on additional sources for candidate schemas, such as the TVTropes web site mentioned in the introduction or the set of narrative schemas extracted automatically by [4]. In addition, such an effort should take into account insights arising from existing work on developing knowledge resources for narrative generation, such as attempts to build planning operators for planning-based generators by means of advanced methods of knowledge engineering [24] and crowdsourcing [18], or work on automated analysis of the narrative structure of movies, by means of techniques such as graph-based scene extraction [16], turning point identification [26] or scene segmentation and alignment [8].

Detailed examination of the drafts produced by the proposed procedure show some instances of plots that strike the reader as incoherent in spite of having acceptable scores in terms of the metrics applied. This happens in cases where more than two schema relating to romantic entanglement—HAPPYLOVE, SHIFTINGLOVE, UNRELENTINGGUARDIAN—are included in the same AoI. The constraints that the current metrics establish between pairs of AoIs appear to be incapable of guaranteeing coherence in this case because they lack the ability to construct a model of how the affinities between characters evolve over time as a result of the events they are involved in. This sometimes results in plots where a character that appeared to have settled happily for a given romantic partner then decides to take on another. An extension of the set of metrics to handle this feature would improve the quality of the results. The potential extensions to improve the set of metrics to cover additional features listed in Sect. 4.4 will also be considered.

The genetic representation presented in this paper does not allow consideration in the same population of drafts that include different combinations of AoIs (see discussion of this problem in Sect. 3.4.3). This constitutes an important restriction, as it prevents the evolutionary solution to be used as means of searching over the space of possible combinations of AoIs. We intend to address this problem in future work, hoping to find a modified genetic representation that makes this possible.

The mechanism employed for rendering the drafts as text in this paper is extremely basic. This was the result of an intentional decision designed to avoid the risk of evaluators being impressed by features introduced by the text rendering mechanism rather than the plot construction procedure. The design of the evaluation as a comparative task instead of a task of scoring individual drafts in absolute terms was also intended to reduce this risk. However, we are aware that the general impression that a reader gets of system output when presented in this form is poor. Enhancement to the text rendering module may improve the overall impact on the reader of system outcomes, and as long as comparative evaluations are maintained they should not cloud the issue of quality of the underlying plot structures.

The introduction of additional intermediate levels of representation of plot-as defended in this paper—opens additional possibilities for decomposing the task of narrative generation into a set of collaborating modules, each focused on a different level of abstraction. This could be applied to explore specific modules both for developing material at lower levels of granularity—construction of more detailed elaborations of each scenes—or higher levels of granularity—combining single plot lines into multi-plot stories. Each of these options is discussed in some more detail.

The procedure described here operates at the level of plot atoms, which corresponds to scenes in a narrative, but in the resulting drafts each scene covers a complete scene described by a single label: `BoyMeetsGirl`, `Wedding`, `DifficultTask`, `Rescue`. This is what makes system outputs be sketches of plot structures rather than actual narratives. An interesting line of research to explore would be to use the plot structure obtained in this way in combination with a mechanism for expanding the descriptions of these scenes into more detailed elaborations, possibly including elements like: sequences of actions by the characters to achieve the described overall result, descriptions of characters and locations, and elaboration on character feelings. As each scene in the input would be sandwiched between some preceding scene and the scene following it, this might a particularly suitable task for narrative generators based on planning [37], which generally take as input a description of an initial situation and a desired goal state. The procedure described here to generated single plot stories could also be combined with a procedure for combining single plotlines into multi-plotline stories such as the one proposed by [6].

Following the approach taken by Gervás [14], in the present paper we have assumed that the relative chronological order of the scenes in each AoI is respected in the final discourse. Consideration of cases of altered chronology (flashbacks, flashforwards) are left to be addressed in further work.

**Acknowledgements** The authors want to acknowledge the assistance of the volunteer evaluators and the helpful comments of the reviewers which have helped improve the paper significantly.

**Author contributions** All authors contributed to the study conception and design. Pablo Gervás wrote the code for the proposed solution. Gonzalo Méndez designed and managed the evaluation and processed the results. Eugenio Concepción and Pablo Gervás carried out the knowledge engineering of the required resources. The first draft of the manuscript was written by Pablo Gervás and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. Pablo Gervás and Gonzalo Méndez revised the paper in response to the reviewers' comments.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This paper has been partially funded by the project CANTOR: Automated Composition of Personal Narratives as an aid for Occupational Therapy based on Reminiscence, Grant. No. PID2019-108927RB-I00 (Spanish Ministry of Science and Innovation) and the ADARVE (Análisis de Datos de Realidad Virtual para Emergencias Radiológicas) Project funded by the Spanish Consejo de Seguridad Nuclear (CSN).

**Data availability and materials** The data are at the present time not publicly available.

## Declarations

**Conflict of interest** The authors are not aware of any competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

1. C. Booker, The seven basic plots: Why we tell stories. *The Seven Basic Plots: Why We Tell Stories*, Continuum (2004). <http://books.google.pt/books?id=qHj9gO10j8C>
2. E. Brehob, Online Academics: The Wiki TV Tropes as a Community of Pseudo-Academic Producers. PhD thesis (University of Michigan, Michigan, 2013)
3. J. Campbell, P. Cousineau, S. Brown, The Hero's Journey: Joseph Campbell on His Life and Work. *Collected works of Joseph Campbell*, New World Library (1990). <https://books.google.co.uk/books?id=0LxpikJraoC>
4. N. Chambers, D. Jurafsky, Unsupervised learning of narrative schemas and their participants, in *Annual Meeting of the Association for Computational Linguistics* (2009)
5. S. Colton, J.W. Charnley, A. Pease, Computational creativity theory: the face and idea descriptive models, in *ICCC, Mexico City*, pp 90–95 (2011)
6. E. Concepción, P. Gervás, G. Méndez G, Exploring baselines for combining full plots into multiple-plot stories. *New Generation Computing* pp 1–41 (2020). <https://doi.org/10.1007/s00354-020-00115-x>
7. W. Cook, *Plotto: The Master Book of All Plots*. Tin House Books (2011). <https://books.google.es/books?id=HDOPLT0Wer8C>
8. T. Cour, C. Jordan, E. Miltsakaki et al., Movie/script: alignment and parsing of video and text transcription, in *Computer Vision—ECCV 2008*. ed. by D. Forsyth, P. Torr, A. Zisserman (Springer, Berlin Heidelberg, Berlin, Heidelberg, 2008), pp.158–171
9. C. Fairclough, Story games and the opiate system: Using case based planning for structuring plots with an expert story director agent and enacting them in a socially simulated game world. PhD thesis, Trinity College Dublin (2004)
10. M.P. Fay, Driving story generation with learnable character models. PhD thesis, Massachusetts Institute of Technology (2014)
11. E.M. Fredericks, B. DeVries, (genetically) improving novelty in procedural story generation. arXiv preprint (2021). [arXiv:2103.06935](https://arxiv.org/abs/2103.06935)
12. Y. Gabriel, 169 Studying emotion through stories, in *Storytelling in Organizations: Facts, Fictions, and Fantasies* (Oxford University Press, Oxford, 2000). <https://doi.org/10.1093/acprof:oso/9780198290957.003.0009>
13. P. Gervás, Generating a search space of acceptable narrative plots. In: 10th International Conference on Computational Creativity (ICCC 2019), UNC Charlotte, North Carolina, USA (2019)
14. P. Gervás, E. Concepción, G. Méndez, Evolutionary construction of stories that combine several plot lines. In: *Computational Intelligence in Music, Sound, Art and Design—11th International Conference, EvoMUSART 2022*, Springer. Springer, Madrid, Spain (2022). [https://doi.org/10.1007/978-3-031-03789-4\\_5](https://doi.org/10.1007/978-3-031-03789-4_5)
15. P. Gomes, A. Paiva, C. Martinho et al., Metrics for character believability in interactive narrative, in *Interactive Storytelling*. ed. by H. Koenitz, T.I. Sezen, G. Ferri et al. (Springer International Publishing, Cham, 2013), pp.223–228
16. P. Gorinski, M. Lapata, Movie script summarization as graph-based scene extraction. In: *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pp 1066–1076 (2015)
17. D. Grasbon, N. Braun, A morphological approach to interactive storytelling. In: *Conference on Artistic, Cultural and Scientific Aspects of Experimental Media Spaces (CAST)*, 2001, St. Augustin (2001)
18. M. Guzdial, B. Harrison, B. Li et al., Crowdsourcing open interactive narrative. In: *Proceedings of the 10th International Conference on the Foundations of Digital Games, FDG 2015*, Pacific Grove, CA, USA, June 22–25, 2015 (2015). [http://www.fdg2015.org/papers/fdg2015\\_paper\\_06.pdf](http://www.fdg2015.org/papers/fdg2015_paper_06.pdf)
19. B. Kartal, J. Koenig, S.J. Guy, User-driven narrative variation in large story domains using monte carlo tree search. In: *Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, AAMAS '14, p 69–76 (2014)
20. J. Lehman, K.O. Stanley, Exploiting open-endedness to solve problems through the search for novelty, in *Proceedings of the Eleventh International Conference on Artificial Life*. MIT Press (2004)
21. E.S. de Lima, B. Feijó, A.L. Furtado, Procedural generation of quests for games using genetic algorithms and automated planning, in *18th Brazilian Symposium on Computer Games and Digital*

- Entertainment, SBGames 2019, Rio de Janeiro, Brazil, October 28-31, 2019. IEEE*, pp 144–153 (2019). <https://doi.org/10.1109/SBGames.2019.00028>
22. N. McIntyre, M. Lapata Plot induction and evolutionary search for story generation, in *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. Association for Computational Linguistics, Uppsala, Sweden*, pp 1562–1572 (2010). <https://aclanthology.org/P10-1158>
  23. M. McKenzie, A. Kilayko, J.C. Macbeth et al., Script combination for enhanced story understanding and story generation systems, in *Proceedings of the Tenth Annual Conference on Advances in Cognitive Systems, ACS 2022* (2022)
  24. B. O'Neill, M.O. Riedl, Applying qualitative research methods to narrative knowledge engineering, in *2014 Workshop on Computational Models of Narrative, CMN 2014, July 31 - August 2, 2014, Quebec City, Canada*, pp 139–153 (2014). <https://doi.org/10.4230/OASICS.CMN.2014.139>
  25. A. Paiva, I. Machado, R. Prada, Heroes, villains, magicians, & hellip:: Dramatis personae in a virtual story creation environment, in *Proceedings of the 6th International Conference on Intelligent User Interfaces. ACM, New York, NY, USA, IUI '01*, pp 129–136 (2001). <https://doi.org/10.1145/359784.360314>
  26. P. Papalampidi, F. Keller, M. Lapata, Movie plot analysis via turning point identification, in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). Association for Computational Linguistics, Hong Kong, China*, pp 1707–1717 (2019). <https://doi.org/10.18653/v1/D19-1180>, <https://aclanthology.org/D19-1180>
  27. P. Papalampidi, K. Cao, T. Kocisky, Towards coherent and consistent use of entities in narrative generation (2022). <https://doi.org/10.48550/ARXIV.2202.01709>
  28. RPÉY Pérez, M. Sharples M, MEXICA: a computer model of a cognitive account of creative writing. *J. Exp. Theor. Artif. Intell.* **13**(2), 119–139 (2001)
  29. G. Polti, L. Ray, *The Thirty-six Dramatic Situations*. Editor Company (1916). [https://books.google.es/books?id=AeB\\_yAEACAAJ](https://books.google.es/books?id=AeB_yAEACAAJ)
  30. J. Porteous, F. Charles, M. Cavazza, Plan-based narrative generation with coordinated subplots, in *European Conference on Artificial Intelligence (ECAI 2016)*, IOS Press, pp 846–854 (2016)
  31. V. Propp, *Morphology of the Folktale* (University of Texas Press, Austin, 1968)
  32. M. Sap, A. Jafarpour, Y. Choi et al., Quantifying the narrative flow of imagined versus autobiographical stories. *Proc. Natl. Acad. Sci.* **119**(45), e2211715 (2022). <https://doi.org/10.1073/pnas.2211715119>
  33. R. Schank, R. Abelson, *Scripts, Plans, Goals, and Understanding: An Inquiry Into Human Knowledge Structures*. Artificial intelligence series, L. Erlbaum Associates (1977). <https://books.google.es/books?id=YZ99AAAAMAAJ>
  34. A.G. Silva Garza de, *An evolutionary approach to design case adaptation*. PhD thesis. (The University of Sydney, Australia, 2000)
  35. A.G. Silva Garza de, R.P. y Pérez, Towards evolutionary story generation, in *Proceedings of the Fifth International Conference on Computational Creativity, ICC3 2014*, ed. by S. Colton, D. Ventura, N. Lavrac et al. Ljubljana, Slovenia, June 10-13, 2014, pp 332–335 (2014). [http://computationalcreativity.net/icc32014/wp-content/uploads/2014/06/15.3\\_Garza.pdf](http://computationalcreativity.net/icc32014/wp-content/uploads/2014/06/15.3_Garza.pdf)
  36. R. Tobias R, *20 Master pots: and how to build them*. F+W Media (2012). <http://books.google.pt/books?id=EZgEqCYIBwsc>
  37. R.M. Young, S. Ware, B. Cassell et al., Plans and planning in narrative generation: a review of plan-based approaches to the generation of story, discourse and interactivity in narratives. *Sprache und Datenverarbeitung. Special Issue on Formal and Computational Models of Narrative* **37**(1–2), 41–64 (2013)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.