Methodological remarks about comparing formal frameworks for narratives

Benedikt Löwe*

Institute for Logic, Language and Computation, Universiteit van Amsterdam, Postbus 94242, 1090 GE Amsterdam, The Netherlands

Fachbereich Mathematik, Universität Hamburg, Bundesstrasse 55, 20146 Hamburg, Germany

Vlaams Academisch Centrum, Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Hertogstraat 1, 1000 Brussel, Belgium

1 Introduction

1.1 Motivation

In conversations about narratives, we often make judgments about whether two stories are the same. Judgments of this type range from overgeneralizations ("All Bollywood movies have the same story") via popular attempts to classify stories in various finite lists (e.g., Polti's famous *36 dramatic situations* [21]) to detailed discussions of whether the story told in a movie is the same as the story told in the book from which the movie was made.

In some of these analyses we base our judgment on superficial features of the narrative, in others we deal with structural features that are independent of the exact representation of the story. For instance, without reflection, a person might be under the impression that there is some similarity between the movies *The Talented Mr. Ripley* (1999) and *Ocean's Eleven* (2001), but upon further investigation it turns out that this was just because the actor Matt Damon features in both of these movies. On the other hand, the similarity between *Ocean's Eleven* and *Ocean's Twelve* (2004) is more structural, as some of the plot elements occur in both movies. This is of course even more the case in the comparison of *Ocean's Eleven* with the 1960 movie *Ocean's 11* upon which it was based. These examples describe the extremes in a spectrum ranging from the most superficial similarity to the most structural similarity.

* The author would like to thank the *Vlaams Academisch Centrum* at the Royal Flemish Academy of Belgium for a fellowship in 2010/11. He would also like to thank his students Irma Cornelisse, Stefan Eichinger, Kevin van Andel, Noortje Venhuizen who did papers and an undergraduate thesis in various projects related to the topic. The discussion of the comparison between DPF and PUF goes back to a discussion with Eric Mueller in Washington DC in September 2009. The final section of the paper describes a joint project with Rens Bod and Sanchit Saraf. Last but not least, the author would like to thank the organizers of 3WPI for inviting him as a speaker, and his commentator at the workshop, Lorenz Demey, for valuable input.

When forced to focus on *structural similarity*, human audiences tend to be quite adept at identifying which of the features are consequences of the presentation of a narrative (choice of actors, choice of exact words, cinematography, etc.) and which are structural. The fact that we can meaningfully debate whether the 1960 and 2011 versions of the *Ocean's Eleven* narrative tell the same story witnesses our ability to distill structural content from highly complex narratives.

Even though human beings, when prompted for judgments whether stories are similar, might at first be focussing solely on superficial features, they are able with some effort to abstract from these superficial features to reach a level on which relatively objective debates can be performed.

Apart from studying an interesting feature of how the human mind works, structural analysis of narratives has had an influence on other areas of scholarly dispute:

> There are several obvious reasons for the long-standing and apparently unending fascination with the flood myth. First of all, the flood myth is one of the most widely diffused narratives known. ... It is therefore a narrative of interest to most peoples of the earth. ... The comparative study of flood myths was invoked as documentary "proof" that the flood had indeed been a worldwide historical event. [7, p. 2–3]

The classification of motifs in folk literature in the *Motif-Index* [28] (where we find flood myths as "A1010: *Deluge: Inundation of whole world or section*") is one attempt to represent the spectrum of narratives at hand in a formal system that removes certain features of the presentation (e.g., the names of the protagonists, the cultural setting, etc.).

Probably the most famous and influential such attempt was Propp's paradigmatic 1928 study *Morphology of the Folktale* by Vladimir Propp [22] in which he identifies seven *dramatis personae* and 31 functions that allow him to formally analyse a corpus of Russian folktales. Propp's analysis started the structuralist school of narratology, and still is the most widely received study in this field. Most of the formal approaches we shall be discussing in the paper derived in one way or another from Propp.

In this paper, we are working under the assumption that there is a *structural core* of narratives, or, to be more precise, there are various structural cores, depending on the level of granularity (cf. § 4.3) of your analysis. Based on this assumption, we shall ask the question:

Question. *When are narratives N and N' structurally the same?* (1)

Our analysis will reveal a number of issues that make any attempt to answer this question difficult. First of all, we cannot expect a context-independent answer to (1): the components of the structure of a narrative depend on what we are interested in. But even with a contextualist reading of the question, the methodological difficulties remind us that it is not at all obvious that such a thing as the *structural core* of a narrative is a well-defined object. Whether there is a stable phenomenon underlying human judgments about *structural similarity of* *narratives* is not yet known, and this question is central to the entire enterprise (and will not be answered by this paper).

1.2 Structure of this paper

We connect the discussion of the methodological issues raised by (1) to the debates of the study of analogy, in particular *Structure Mapping Theory* (§ 2), and the framework of computational models of narrative (§ 4). We propose a notion of comparison between formal frameworks of narrative (§ 5) and give an example of how to compare two formal frameworks (in this case, Lehnert's *Plot Units* [15] and the *Doxastic Preference Framework* from [19]; § 5.2). Our discussion is structured by a three-step approach to answering (1) given in (5). A number of these issues require empirical grounding of the formal frameworks. In § 6, we discuss first steps towards developing this empirical grounding.

The work described in §§ 4 and 5 is an extended version of the material described in [16]. The approach described in § 6 is joint work with Bod and Saraf and based on [2].

2 Lessons from the study of analogy

The search for the structural core of narratives is closely related to questions in the field of analogy study:

> Similarity-based remindings range from the sublime to the stupid. On one extreme, seeing the periodic table of elements reminds one of octaves in music. At the other, a bicycle reminds one of a pair of eyeglasses. Often, remindings are neither brilliant nor superficial, but simply mundane ... Theoretical attention is inevitably drawn to spontaneous analogy: That is, to structural similarity unsupported by surface similarity. [10, pp. 141–142]

The spectrum of similarities has been explored by cognitive psychologists with short narratives of the type of *Karla the Hawk* to study analogical reasoning. The fact that the human notion of similarity of stories is used for studies on analogy stresses once more the relation between the two research areas. *Karla the Hawk* and its true and false analogies go back to [23] and have been used in many studies since:

> Karla, an old hawk, lived at the top of a tall oak tree. One afternoon, she saw a hunter on the ground with a bow and some crude arrows that had no feathers. The hunter took aim and shot at the hawk but missed. Karla knew the hunter wanted her feathers so she glided down to the hunter and offered to give him a few. The hunter was so grateful that he pledged never to shoot at a hawk again. He went off and shot deer instead. [12, p. 533] (2)

> Once there was an eagle named Zerdia who donated a few of her tailfeathers to a sportsman and he promised never to attack eagles. One day Zerdia was nesting high on a rocky cliff when she saw the sportsman coming with a crossbow. Zerdia flew down to meet the man, but he attacked and felled her with a single bolt. As she fluttered to the ground Zerdia realized that the bolt had her own tailfeathers on it. [12, p. 533] (3)

> Once there was a small country called Zerdia that learned to make the worlds smartest computer. One day Zerdia was attacked by its warlike neighbor, Gagrach. But the missiles were badly aimed and the attack failed. The Zerdian government realized that Gagrach wanted Zerdian computers so it offered to sell some of its computers to the country. The government of Gagrach was very pleased. It promised never to attack Zerdia again. [12, p. 533] (4)

Most test subjects recognize after some reflection that the similarities between the hawk story and the eagle story are superficial (both about birds and shooting), whereas the similarities between the hawk story and the countries-at-war story are structural.

The main methodological point of Gentner's *Structure Mapping Theory* [11] is that the structure of these short narratives can be expressed in terms of relations between the agents of the narrative, and thus analogy or structural similarity corresponding to structure-preserving mappings between the relational structures representing the narratives, or, in mathematical terms, homomorphisms.

Following this methodology, we can give a first attempt of answering (1):

Question. *When are narratives N and N' structurally the same?*

Answer.

Step 1. Develop a formal description language with mathematical structures S corresponding to narratives and a notion of isomorphism $\simeq$ between structures. (5)

Step 2. Formalize the narratives N and N' to obtain structures S and S' faithfully representing N and N'.

Step 3. The narratives N and N' are structurally the same if and only if $S \simeq S'$.

Structure Mapping Theory has been criticized by some authors as being focused too much on the physical relations between the agents in the narrative. Gentner's original examples are physical relations [11, p. 157]:

$$\mathsf{large}(x)$$
$$\mathsf{collide}(x, y)$$
$$\mathsf{strike}(x, y)$$
$$\mathsf{cause}[\mathsf{collide}(x, y), \mathsf{strike}(y, z)]$$

In [13], Lam shows in an experiment that the *Structure Mapping Engine* (SME), the implemented version of Gentner's *Structure Mapping Theory* (cf. [9]), is not able to explain that humans in their similarity judgments tend to prefer false analogies with similar emotive content over false analogies with different emotive content.[1] Thagard and Shelley have lamented the fact that "analogy researchers have paid remarkably little attention to emotion" [27, p. 335] and discuss analogies such as "Love is a rose and you better not pick it".

There are two separate affective components of narrative which we shall call *emotion* (referring to the emotional relation between events and agents in the story) and *sympathy* (referring to the emotional relation between the story or its characters and the audience). In [6], Cornelissen and Venhuizen have investigated human similarity judgments for stories that differ in emotive and sympathic content to find out that even for structurally similar stories, human similarity judgments depend on emotive and sympathic components of the story.[2]

We should stress that this criticism is directed towards the concrete implementation of *Structure Mapping Theory*, not towards the general methodology as it underlies (5). Emotive and sympathic content of the type investigated by Cornelissen and Venhuizen could be incorporated in a formal framework as a relation: emotive content would be a relation between the agents of the narrative and the events; sympathic content would be a relation between the events and the audience (which would then have to be represented in the framework). We shall make this clearer in a toy example framework in the following section.

3 Illustration: A toy example

In order to show how emotive content can be incorporated into a formal framework, let us give a toy example. As a disclaimer, we stress that the toy languages developed in this section are by no means intended to serve as a formal framework for formalizing narratives: they serve purely as illustration.

Consider a language $\mathbf{TL}_1$ (for "toy language") with variables $A = \{a_0, a_1, ...\}$ for agents and $O = \{x_0, x_1, ...\}$ for objects. If a is an agent and o is an object, then both $\mathsf{own}(a, x)$ and $\mathsf{not\text{-}own}(a, x)$ are *states*. Furthermore, if a, b are agents, o is an object, s is a state, and e is an event, then

$$\begin{aligned}&\mathsf{desire}(a, s),\\&\mathsf{attack}(a, b),\\&\mathsf{failure}(e),\\&\mathsf{give}(a, b, x), \text{ and}\\&\mathsf{promise}(a)\end{aligned}$$

[1] "We have shown that [the] lack of inclusion of emotive content [in Gentner's *Structure Mapping Engine*] has made it psychologically implausible." [13, p. 38].

[2] "[A] story [with] different emotional content [and a] story ... imply[ing] a different feeling of sympathy ... are both [rated] significantly ... less similar to the *Base Story* than the *True Analogy*." [6, p. 13]

are *events*. A $\mathbf{TL}_1$-structure is a finite sequence $\langle p_0, ..., p_n \rangle$ such that all p_i are either states or events.

We can now represent *Karla the Hawk* by the following $\mathbf{TL}_1$-structure $K = \langle p_0, p_1, ..., p_6 \rangle$:

$$\begin{array}{l} \mathsf{not\text{-}own}(a, x) \\ \mathsf{desire}(a, \mathsf{own}(a, x)) \\ \mathsf{attack}(a, b) \\ \mathsf{failure}(\mathsf{attack}(a, b)) \\ \mathsf{give}(b, a, x) \\ \mathsf{own}(a, x) \\ \mathsf{promise}(a) \end{array}$$

The toy language $\mathbf{TL}_1$ comes with a natural notion of isomorphism of structures: If $P = \langle p_0, ..., p_n \rangle$ and $Q = \langle q_0, ..., q_n \rangle$ are $\mathbf{TL}_1$ structures, they are *isomorphic* if there is are permutations π_A and π_O of the agent and object variables, respectively, such that for any i, ${p_i}^{\pi_A, \pi_O}$ is q_i.

The following story is analogous to *Karla the Hawk* (within the constraints of expressivity of $\mathbf{TL}_1$):

> Argutt, a wise owl, watched a merchant with a bow with crude arrows that had no feathers. The merchant tried to shoot Argutt, but the shot missed. Argutt realized that the merchant needed the feathers for his arrows, approached him and offered a single owl feather. The merchant accepted the gift and was so surprised by seeing a talking owl, that he made a promise to take his own life so that he could never harm animals again. (6)

But we feel that there is a noticeable difference between *Karla the Hawk* and this story. In *Karla the Hawk*, the hunter "went off and shot deer instead". The promise that the hunter made is positive for Karla, but has no major negative effect on the hunter. In the above story, the promise obviously has a strongly negative effect for the merchant. This difference (not expressible in $\mathbf{TL}_1$) makes the human reader decide that the stories are not similar and thus shows that the language $\mathbf{TL}_1$ is deficient.

In order to fix this, we now say that a sequence $\langle p_0, ..., p_n, V \rangle$ is a $\mathbf{TL}_2$-structure if $\langle p_0, ..., p_n \rangle$ is a $\mathbf{TL}_1$ structure and $V : \{0, ..., n\} \times A \to \{+, \circ, -\}$ is a function, interpreting $V(i, a) = +/\circ/-$ as "p_i is positive/neutral/negative for agent a". These structures come with their natural notion of isomorphism: If $P = \langle p_0, ..., p_n, V \rangle$ and $Q = \langle q_0, ..., q_n, W \rangle$ are $\mathbf{TL}_2$ structures, they are *isomorphic* if there is are permutations π_A and π_O of the agent and object variables, respectively, such that for any i, ${p_i}^{\pi_A, \pi_O}$ is q_i and $V(i, a) = W(i, \pi_A(a))$ for all i and a.

The following $\mathbf{TL}_2$-structure $K_\circ = \langle p_0, p_1, ..., p_6, V_\circ \rangle$ represents *Karla the Hawk* (we represent the values of the function $V_\circ(i, x)$ in the column labeled x

behind p_i):

	a	b
$\mathsf{not\text{-}own}(a, x)$	$-$	$\circ$
$\mathsf{desire}(a, \mathsf{own}(a, x))$	$\circ$	$\circ$
$\mathsf{attack}(a, b)$	$\circ$	$-$
$\mathsf{failure}(\mathsf{attack}(a, b))$	$-$	$+$
$\mathsf{give}(b, a, x)$	$+$	$\circ$
$\mathsf{own}(a, x)$	$+$	$\circ$
$\mathsf{promise}(a)$	$\circ$	$+$

Since the promise at the end of *Argutt the Owl* is negative for the merchant, the representation of that story as a $\mathbf{TL}_2$-structure $K_- = \langle p_0, p_1, ..., p_6, V_- \rangle$ would be (again, the values of the function $V_-(i, x)$ are in the column labeled x behind p_i):

	a	b
$\mathsf{not\text{-}own}(a, x)$	$-$	$\circ$
$\mathsf{desire}(a, \mathsf{own}(a, x))$	$\circ$	$\circ$
$\mathsf{attack}(a, b)$	$\circ$	$-$
$\mathsf{failure}(\mathsf{attack}(a, b))$	$-$	$+$
$\mathsf{give}(b, a, x)$	$+$	$\circ$
$\mathsf{own}(a, x)$	$+$	$\circ$
$\mathsf{promise}(a)$	$-$	$+$

The structures $K_\circ$ and K_- are not isomorphic according to the above definition, and hence—following our methodological outline in (5)—we have shown that the two stories are not structurally the same.

This toy example shows us a great deal about what is involved in the procedure described in (5). First of all, the example highlights that whether two stories are the same according to this procedure depends on the choice of the formal framework: the more expressive a formal framework is, the fewer narratives will be seen as structurally the same. It also tells us something about the process of finding the right framework: we observed that the language $\mathbf{TL}_1$ forced us to conclude that *Karla the Hawk* and *Argutt the Owl* are structurally the same; this conclusion went against our intuitions, and forced us to refine $\mathbf{TL}_1$ to a richer language $\mathbf{TL}_2$. This process is a typical case of conceptual modelling as it is very common in philosophy (cf. [17, § 2] for a discussion of this general technique): we iterate the design of a formal framework and its testing against the phenomena (in this case our intuitions of story similarity) until we reach a reflective equilibrium. In [17], the authors stress that a good instance of conceptual modelling is testing the formal framework against *stable phenomena* rather than *idiosyncratic data* (this distinction is due to Bogen and Woodward [3]). We shall come back to this important point in § 6.

In our given case of conceptual modelling, we have two conflicting desiderata for our formal framework. It should be

1. simple enough so that we have no disagreement about whether a structure is the correct representation of the structure of a story (again, this will be taken up again in § 6), and
2. expressive enough to capture all features relevant for the notion of structural equivalence we are aiming for.

4 Formal Frameworks of Narrative

4.1 History

The formal study of narratives goes back to the Russian structuralist school, paradigmatically represented by the 1928 study *Morphology of the Folktale* by Vladimir Propp [22] in which he identifies seven *dramatis personae* and 31 functions that allow him to formally analyse a corpus of Russian folktales. Propp's motivation, as described in the introduction of his book, is not too far from ours:

> Since [narratives are] exceptionally diverse, and evidently cannot be studied at once in [their] full extent, the material must be divided into sections, i.e., it must be classified. Correct classification is one of the first steps in a scientific description. The accuracy of all further study depends upon the accuracy of classification. [22, p. 5]

The work of Propp initiated at least two broad directions of research. First of all, modern narratology started with the Russian structuralist school, represented by researchers such as Tzvetan Todorov (who coined the word "narratology"), Gérard Genette, and Roland Barthes. The classification of mythical motifs [28] mentioned in the introduction has to be understood as part of this tradition. In this paper, we are not so much concerned with narratology proper as it is studied in literature departments, but rather with its computational reflection.

Linguistics typically deals with the question how smaller units are composed to form larger units: sounds to words, words to sentences, sentences to discourses. In this hierarchy, narratives are the natural next step, consisting of several discourses. In the early days of Artificial Intelligence, researchers were fascinated by the structuralist ideas and by the vision of generalizing the analysis of linguistics to the level of narratives. Rumelhart's *Story Grammars* [24] are the paradigmatic case for this type of formal representation of narratives. Early computational models of narrative were close to the original structuralist ideas: systems of high-level descriptions, avoiding the details of the narrative process (we shall discuss the example of Lehnert's *Plot Units* [15] below). While systems like this are adequate for a structural analysis of narratives, they do not work very well for automated systems for story understanding or generation. A programme that extracts story structure from a given text will first have to learn which of the details to omit, and in order to do so, cannot ignore the details altogether:

> Scott Turner, like many before and since, first became interested in story generation after running upon Vladimir Propp's analysis of Russian folktales. Propp provides a grammar that describes the structure of many folktales. As linguists and computer scientists know, grammars can be used for describing the structure of given things—and also for generating new things. But, as Turner soon discovered, this task is not easily accomplished with Propp's grammar. Its elements are rather abstract, making them workable for analysis but insufficient for generation. [30, pp. 185–186]

Thus, the next generation of formal frameworks for narratives paid more attention to the details of the narratives and the systems became increasingly complex. Examples of these systems are Schank's *Thematic Organization Points* (TOPs) [25], Dyer's *Thematic Abstraction Units* (TAUs) [8], or Turner's *Planning Advice Themes* (PATs) [29]. These more detailed systems have been used with great success in the research of the last decades:

> There is now a considerable body of work in artificial intelligence and multi-agent systems addressing the many research challenges raised by such applications, including modeling engaging virtual characters ... that have personality ..., that act emotionally ..., and that can interact with users using spoken natural language. [26, p. 21]

We believe that formal frameworks adequate for capturing the informal notion of *structural equivalence of stories* are likely to be closer to the early coarse frameworks than to the very elaborate modern models. As mentioned at the end of § 3, one of the requirements of our enterprise is that we need a framework that is simple enough to allow human audiences to agree whether a formalization properly and adequately represents a narrative.

We can now return to (5) and become more specific about Step 1 asking for the development of a formal description language; due to the mentioned requirements of simplicity, such a development should start from very simple frameworks and add features after there is evidence that we cannot ignore them in our analysis of the notion of *structural equivalence*.

A *formal framework* is a mathematical or logical entity, given by a *syntax* and a corresponding *semantics*. The syntax determines a formal language which in turn determines a type of mathematical model for the formal language together with a notion of *satisfaction* in the usual sense of mathematical logic. There is a natural notion of *isomorphism* between models of the right type for the given formal language (denoted by $\simeq$): as usual, bijections preserving all of the relevant structure.

4.2 The plot unit framework

We shall describe Lehnert's *Plot Units* [15] as an example for such a formal framework. In her framework, we represent the narrative as a grid of events.

Each *agent* in the narrative is represented by a column in the grid; all events occurring in that column are events affecting that particular agent. There are three types of events: *mental events*, and *positive* and *negative events*; here, "positive" ("negative") means "positively (negatively) affecting the agent corresponding to the column where the event is listed". Events in the same column can be linked by *causal links* of which there are three types: motivation, actualization, termination and equivalence. Events in different columns can be linked by *interactive links.* There are a number of rules as to which links are allowed, but we do not go into detail here (cf. [15] for details.) The reason for the name "plot unit" is that Lehnert gives a list of basic constituents in the form of plot unit structures that can then be used to generate more elaborate narratives.

The plot unit structures can now be represented graphically as labelled graphs where $+$, $-$ and $\mathbf{M}$ represent the three types of events (positive, negative and mental, respectively), and `m`, `a`, `t`, and `e` label the causal links as "motivation", "actualization", "termination", and "equivalence", respectively. In Figure 1, we give an example of a plot unit structure with two agents: we read it from top to bottom, thinking of time flowing downwards; the second agent (right column) has a mental state representing the desire to perform an action of mutual benefit to both agents; this action in turn motivates the first agent (left column) to reciprocate in kind. This plot unit structure could be the representation of the following narrative:

> John liked Adam, and invited him for a pleasant dinner in a fancy restaurant one evening. Adam wanted to reciprocate and realized that John likes the ballet. Adam bought two ballet tickets and invited John to join him. (7)

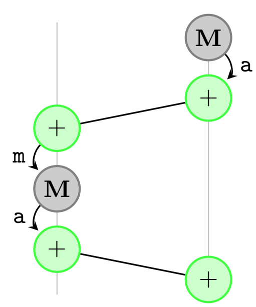


Fig. 1. An example of a plot unit structure representing the narrative (7); the left column represents Adam, the right one represents John.

We notice that (7) is hardly more than a verbalization of the plot unit structure, and thus very far from what actual narratives look like. In novels or even short stories, we would expect more information that is not necessarily relevant for the structure of the narrative: the type of restaurant they went to, the food they had, whether Adam also likes the ballet etc. In narratology, this relates to the distinction between "story" and "discourse" (alternatively, "histoire"/"récit") [5]. We would expect to find the structural elements that determine whether stories are structurally the same in the *story* rather than the *discourse*. But not all information that is given as part of the *discourse* is irrelevant for the structure of the narrative. This phenomenon gives rise to the crucial notion of *granularity* in the formalization process that we'll describe in the next section.

4.3 Granularity

In order to understand granularity of formalizations, let us give a simple example:

> Simon came home from an exhausting day at work. While he was standing in front of his apartment door, he heard his phone ringing. He had been waiting for an important call all week, and he was hoping that this would be it. He fumbled for his key and tried to open the door as quickly as possible. First he grabbed his car key, obviously with no success. Finally, he managed to rush into his apartment, but when he reached the phone, the caller had hung up. (8)

This narrative is about failing to reach the phone before it stops ringing. It is thus faithfully represented by the plot unit standing for failure and termination (the left hand structure of Figure 2). On the other hand, there is a second action embedded in the narrative: Simon fumbles for his key and is unsuccessful in the first attempt. This could be represented by the plot unit for failure embedded in the one for failure and termination (the right hand structure of Figure 2).

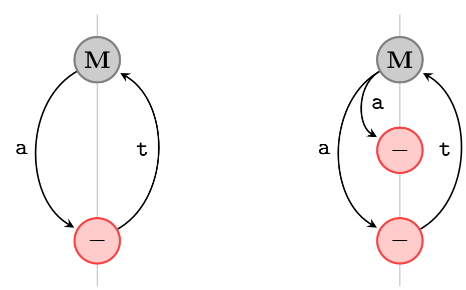


Fig. 2. Two plot unit structures, both representing (8) at different levels of granularity.

This example shows that the framework alone does not determine a unique structure that is *the formalization* of a narrative. The *discourse* of typical narratives will consist of many micro-actions that could be embedded into the formal representation as successful or unsuccessful actions: grabbing for keys, opening doors, sitting down, standing up, etc. In many standard circumstances, these actions will be of no relevance for the question of structural similarity, and thus we would decide to formalize the narrative at a level of granularity that will ensure that these are not represented in the formal structure.

While formal frameworks and their structures are mathematical objects, the notion of *granularity* lives in the world between mathematical objects and the narratives: it tells us how much of the informal information that is contained in the narrative will be represented in the mathematical structure. Therefore, we cannot give a mathematical definition of *level of granularity.* Instead, the level of granularity would be given as part of the guidelines given to human formalizers in the process of formalization. It is important to note that these guidelines should be independent of the formal framework: when we compare formal frameworks, we wish to keep the level of granularity constant between the two formal frameworks.

4.4 The doxastic preference framework

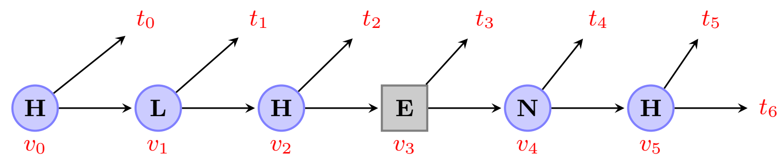
We shall now describe a formal framework developed by the author and Pacuit in [18] which is quite different in its set-up from the plot unit framework. This framework, called the *doxastic preference framework*, considers narratives as game-theoretic (perfect information) decision trees where each node of the tree represents either a decision of one of the agents or an event. The terminal nodes of the tree are the possible *outcomes* of the narrative, and the agents of the story have a preference concerning those outcomes, represented by a linear order of the set of outcomes. In addition, we have layers of belief about these preferences: at the first level of these layers, agent $\mathbf{X}$ has a belief about what he or she thinks is the preference relation of agent $\mathbf{Y}$ for each point in time (i.e., a node of the decision tree). At the next level, we have the belief about what agent $\mathbf{X}$ thinks what agent $\mathbf{Y}$ believes are the preferences of agent $\mathbf{Z}$ for each point in time.

Formally, this is represented as follows. For each sequence of agents $\vec{\mathbf{P}} = (\mathbf{P}_0, ..., \mathbf{P}_n)$ of agents, every agent $\mathbf{X}$, and every node v of the decision tree, we write

$$S(v, \vec{\mathbf{P}})(\mathbf{X})$$

for the belief of $\mathbf{P}_0$ about the belief of $\mathbf{P}_1$ about ... about the belief of $\mathbf{P}_n$ about the preference of $\mathbf{X}$. If $\vec{\mathbf{P}} = \varnothing$, then $S(v, \varnothing)(\mathbf{X})$ stands for the true preference of agent $\mathbf{X}$ at node v. We represent preferences as a sequence of terminal nodes, i.e., (t_1, t_2, t_0) stands for "t_1 is preferred over t_2 and t_2 is preferred over t_0. If v is a non-terminal node and t is a terminal node, we write (v, t) to mean "all terminal nodes succeeding v are preferred over t", and similarly for (t, v). Figure 3 gives an example of a relatively typical doxastic preference structure representing a narrative. Details can be found in [18].

Doxastic preference structures have been used in [19] to study actual narratives from the TV crime series *CSI: Crime Scene Investigation*™. We shall use this formal framework in the next section in order to explain another phenomenon that we encounter naturally in the formalization process.



$S(v_0,\varnothing)(\mathbf{H}) = (t_3,t_0)$; $S(v_1,\varnothing)(\mathbf{L}) = (t_2,t_1)$; $S(v_1,\mathbf{L})(\mathbf{H}) = (t_2,v_3)$;
$S(v_1,\varnothing)(\mathbf{H}) = (t_3,t_2)$; $S(v_2,\varnothing)(\mathbf{H}) = (t_3,t_2)$; $S(v_2,\mathbf{H})(\mathbf{E}) = (t_3,v_4)$;
$S(v_3,\varnothing)(\mathbf{E}) = (v_4,t_3)$; $S(v_4,\varnothing)(\mathbf{N}) = (t_6,t_4)$; $S(v_4,\mathbf{N})(\mathbf{H}) = (t_6,t_5)$;
$S(v_5,\varnothing)(\mathbf{H}) = (t_6,t_5)$

Fig. 3. An example of a doxastic preference structure.

4.5 Multiple representations

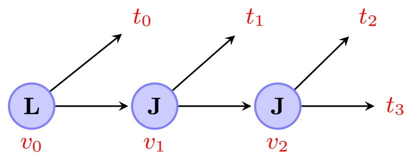
Even if a level of granularity is fixed by objective rules, it is possible to run into situations where there is no unique formal representation of a part of a narrative. In [19, § 3.3], the authors discuss a difficult case in the formalization of *CSI: Crime Scene Investigation*™ narratives where the formalization will depend on the interpretation of the human formalizer. In the following, we shall discuss a simpler case:

> Jeff and Linda agreed to go on a vacation to China together. Each of them was only interested in doing this trip with the other person. Both put considerable effort into the preparations for this trip. But when the day of departure came, Jeff got cold feet and cancelled his flight. Linda was not interested in going without him and was terribly disappointed. (9)

In order to represent the decision structure of the narrative (9), we wish to express that originally, Linda believes that Jeff will commit to do the trip and vice versa. Each of these should be decision nodes in our doxastic preference structure. However, the tree nature of our structures requires us to decide on a temporal order, so we obtain two doxastic preference structures, given in Figures 4 and 5. These structures are not isomorphic in the natural notion of isomorphism for doxastic preference structures. And yet, we would say that they represent the narrative equally faithfully. It was a formal constraint of the framework that forced the us to make a representational decision for which we cannot give any rationale on the basis of the narrative.

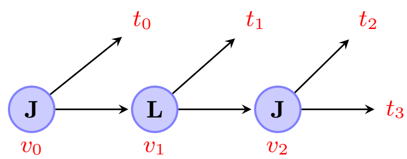
4.6 Formalizations

With the examples of §§ 4.3 and 4.5 in mind, we can now give a semi-formal definition of the notion of *formalization*. A formalization (despite its name) is necessarily a semi-formal process: it links the narratives—informal objects—to mathematical objects. If we fix a formal framework Σ and a level of granularity $\mathcal{G}$, we say that $(\Sigma,\mathcal{G})$ determines a (semi-formal) formalization operation $F_{\Sigma,\mathcal{G}}$



$S(v_0,\varnothing)(\mathbf{L}) = (t_3,t_0,t_1,t_2)$; $S(v_0,\varnothing)(\mathbf{J}) = (t_3,t_0,t_1,t_2)$; $S(v_2,\varnothing)(\mathbf{J}) = (t_0,t_1,t_2,t_3)$;
$S(v_2,\mathbf{L})(\mathbf{J}) = (t_3,t_0,t_1,t_2)$

Fig. 4. A representation of (9) giving Linda's decision before Jeff's.



$S(v_0,\varnothing)(\mathbf{L}) = (t_3,t_0,t_1,t_2)$; $S(v_0,\varnothing)(\mathbf{J}) = (t_3,t_0,t_1,t_2)$; $S(v_2,\varnothing)(\mathbf{J}) = (t_0,t_1,t_2,t_3)$;
$S(v_2,\mathbf{L})(\mathbf{J}) = (t_3,t_0,t_1,t_2)$

Fig. 5. A representation of (9) giving Jeff's decision before Linda's.

that assigns to each narrative N the set of all Σ-structures that represent N faithfully given the level of granularity $\mathcal{G}$, written as $F_{\Sigma,\mathcal{G}}(N)$. Then, a formal framework together with a level of granularity generates a relation $\equiv_{\Sigma,\mathcal{G}}$ between narratives by

$$N \equiv_{\Sigma,\mathcal{G}} N^* :\Leftrightarrow \forall M \in F_{\Sigma,\mathcal{G}}(N) \exists M^* \in F_{\Sigma,\mathcal{G}}(N^*)(M \simeq M^*)$$
$$\wedge \forall M^* \in F_{\Sigma,\mathcal{G}}(N^*) \exists M \in F_{\Sigma,\mathcal{G}}(N)(M \simeq M^*).$$

This semi-formal relation is the notion of structural equivalence that is used in our described methodology in (5). Under the assumption that we have found an adequate formal framework Σ and level of granularity $\mathcal{G}$, the formal relation $\equiv_{\Sigma,\mathcal{G}}$ represents the informal relation of *structural similarity* and is thus our object of study. This lets us now turn to the task of finding the adequate Σ and $\mathcal{G}$.

5 Comparison of formal frameworks

5.1 Formal and semi-formal definitions

Assume that we have fixed a level of granularity $\mathcal{G}$. We shall now compare formal frameworks Σ and Σ^* by studying the relations $\equiv_{\Sigma,\mathcal{G}}$ and $\equiv_{\Sigma^*,\mathcal{G}}$. There are three cases:

Case 1 Σ is a refinement of Σ^*. This means that for any two narratives N and N^*, if $N \equiv_{\Sigma^*,\mathcal{G}} N^*$, then $N \equiv_{\Sigma,\mathcal{G}} N^*$.

Case 2 Σ^* is a refinement of Σ. This means that for any two narratives N and N^*, if $N \equiv_{\Sigma,\mathcal{G}} N^*$, then $N \equiv_{\Sigma^*,\mathcal{G}} N^*$.

Case 3 The frameworks are incomparable. This means that there are narratives N_0, N_1, N_2, and N_4 such that $N_0 \equiv_{\Sigma,\mathcal{G}} N_1$, $N_0 \not\equiv_{\Sigma^*,\mathcal{G}} N_1$. $N_2 \equiv_{\Sigma^*,\mathcal{G}} N_3$, and $N_2 \not\equiv_{\Sigma,\mathcal{G}} N_3$.

Given two formal frameworks Σ and Σ^*, we will test their expressive power against our intuition of structural equivalence. For instance, if Σ is a proper refinement of Σ^*, i.e., there are narratives N and N^* such that $N \not\equiv_{\Sigma,\mathcal{G}} N^*$, but $N \equiv_{\Sigma^*,\mathcal{G}} N^*$, we relate this situation to our intuitions: are the narratives N and N^* structurally the same? If so, then $\equiv_{\Sigma^*,\mathcal{G}}$ represents our notion of structural equivalence better.

We are searching for features that can be expressed in Σ, but not in Σ^* that are a relevant part of our notion of structural equivalence. It is features like this that need to be included in our formal framework. In the following section, we shall explain how such a comparison works in the example of the plot unit framework and the doxastic preference framework.

5.2 The plot unit framework and the doxastic preference framework are incomparable

The doxastic preference framework is very good in expressing expectations of agents since you can calculate the current belief of an agent about future actions from the states. This feature seems to be missing in the plot unit framework. On the other hand, the causal links of the plot unit framework allow us to express precisely which actions in the past are causal for actions in the future, whereas in the doxastic preference framework it is always the entire history of the game played so far that is taken into account in the agents' decisions. We shall now make this observation precise. In order to do so, consider the following two narratives:

> Andrea loved to play the cello and thought that her neighbours were grateful for the beautiful evening music. Unbeknowst to her, her neighbour Bart disagreed. If she had known how much he hated it, she would have played in the music room at her university. But she had no clue, and thus, one evening, the music annoyed Bart so much that he rushed into her apartment, crushing the cello with a large hammer. (10)

> Abel was fully aware that Barbara hated it when he called her "dear", but he couldn't stop himself from doing so. One day, he was particularly annoying and Barbara slapped him in the face. Abel was not really surprised, but stopped calling her "dear" nevertheless. (11)

Both narratives (10) and (11) are represented by the same plot unit structure, given in Figure 6. The only structural difference between the narratives is that

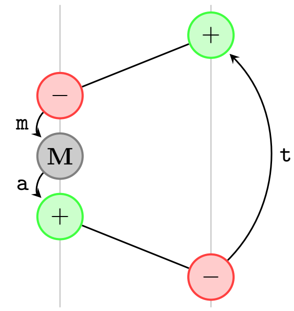
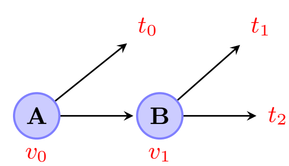


Fig. 6. The plot unit structure for (10) and (11).

Andrea is surprised by Bart's action, whereas Abel expected Barbara's reaction; this difference is not expressible in the plot unit framework.

In Figure 7, we represent "Andrea doesn't play the cello" by t_0, "Andrea plays the cello and Bart does not destroy it" by t_1, and "Andrea plays the cello and Bart destroys it" by t_2. Andrea's surprise is expressed by the fact that her belief about Bart's preferences coincides with her own preferences: doing the backwards induction reasoning on the basis of Andrea's subjective belief state would give t_1 as Andrea's predicted (and preferred) outcome. The fact that Andrea's predicted outcome does not coincide with the actual outcome t_2 models Andrea's surprise.

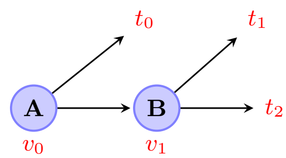


$S(v_0, \varnothing)(\mathbf{A}) = (t_1, t_0, t_2)$; $S(v_1, \varnothing)(\mathbf{B}) = (t_0, t_2, t_1)$; $S(v_0, \mathbf{A})(\mathbf{B}) = (t_1, t_0, t_2)$.

Fig. 7. The doxastic preference structure for (10).

In Figure 8, we now give the doxastic preference structure for the narrative (11). We notice that Abel's belief about Barbara's preferences is correct. Abel's subjective prediction of the outcome is t_2, coinciding with the actual outcome and thus modelling Abel's lack of surprise.

We see that the doxastic preference structures given in Figures 7 and 8 are not isomorphic, and hence the narratives (10) and (11) are not structurally the same with respect to the doxastic preference framework. Together, we showed that there are narratives that can be distinguished by the doxastic preference framework, but not by the plot unit framework.



$S(v_0, \varnothing)(\mathbf{A}) = (t_1, t_2, t_0)$; $S(v_1, \varnothing)(\mathbf{B}) = (t_0, t_2, t_1)$; $S(v_0, \mathbf{A})(\mathbf{B}) = (t_0, t_2, t_1)$.

Fig. 8. The doxastic preference structures for (11).

To see the converse, consider a third narrative that is almost the same as (10) except for the fact that there is no causal connection between the fact that agent **B** dislikes the actions of agent **A** and the action of agent **B**:

> Arnold loved to play his electric guitar in the common room and thought that his roommates were grateful for the exciting evening music. Unbeknowst to him, his roommate Beatrix disagreed, but never dared to do anything about it. If he had known, he would have used the music room at his university for his guitar sessions. One evening, Arnold came to the common room, but Beatrix was preparing supper, using the only power outlet for her toaster, so that Arnold could not play his usual evening music. Beatrix enjoyed her toast in silence. (12)

The doxastic preference framework analysis for (12) is exactly the same as for (10) and results in the same doxastic preference structure (as given in Figure 7). However, the plot unit structure is different from the one given in Figure 6: in (10), the mental state created by playing the cello is causal for Bart's action; in (12), making supper is not causally related to the fact that Beatrix doesn't like the music and just has the positive side effect of silencing Arnold. The resulting plot unit structure can be seen in Figure 9.

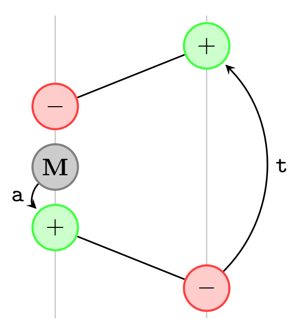


Fig. 9. The plot unit structure for (12).

We see that (10) and (12) have the same representation in the doxastic preference framework, but different representations in the plot unit framework.

5.3 Lessions for the design of frameworks

In the last section, we saw that the feature of "expectation" cannot be represented in the plot unit framework but is naturally represented in the doxastic preference framework; and similarly, the feature of "causality" cannot be represented in the doxastic preference framework but in the plot unit framework. This way of expressing our result is more informative than just saying that the two frameworks are incomparable: in terms of our methodology in (5), we have just witnessed an important part of the design of formal frameworks. We have observed that the two frameworks are incomparable, and furthermore, we know features that cannot be expressed properly. We will now have to make a decision whether these features (in this case, "expectations" and "causality") are relevant for the structural core of a narrative or not. If they are, we will have to add them to our frameworks. Such a judgment will depend on empirical evidence: asking human readers of narratives whether they agree on the structural isomorphism statements made by particular formal systems or not. We shall discuss this empirical grounding of the formal frameworks in the final section of the paper.

6 Empirical grounding of the choice of formal framework

Reconsidering the methodology given in (5) and its three steps, we observe that § 4 provided us with a general setting in which we formulate the frameworks that are candidates for Step 1 as well as a precise method for dealing with Step 3. In § 5, we discussed how we can compare different candidates for Step 1: if a particular relevant feature of narratives is expressible in some formal framework, but not in our currently preferred one, we will have to supplement our preferred framework with the means to express this feature. However, in order to determine whether a feature is relevant, we have to resort to empirical work and base our decisions in human judgments of whether a feature is structurally relevant or not.

Another empirical question concerns Step 2 of our procedure: the actual process of transforming a natural language narrative into a formal structure. Of course, we would like that this process is objective and independent of idiosyncratic decisions of the person who performs it.

Both of these issues relate to the fundamental question raised at the end of § 1.1: is there a stable phenomenon underlying human judgments about structural similarities of narratives?

Questions of this type are quite common in any research that links formal languages to natural phenomena: To name but one example, in linguistics, corpora (i.e., natural language) are being annotated to be represented in formal grammatical structures (cf., e.g., [14]). The rules of the annotation are typically given by a list of examples in annotation guidelines; the process of annotation is performed by human annotators who could disagree about the correct annotation or also be just wrong in some of their judgments. In sentence- or discourse-level annotation, the quality of annotation is typically studied as *inter-annotator agreement* (cf., e.g., [4,20]).

In § 4.1, we already emphasized that the formalization of narratives can be seen as the next level after the formalization of discourses. The influence of ambiguity, personal decisions of the formalizer and judgment errors increases as the complexity of the formalized units grows:

> Ever since the mid-[1990s], increasing effort has gone into putting semantics and discourse research on the same empirical footing as other areas of Computational Linguistics. This soon led to worries about the subjectivity of the judgments required to create annotated resources, much greater for semantics and pragmatics than for [other areas of linguistics]. [1, p. 555]

At the level of narratives, we should be even more worried about this. Interestingly enough, for the annotation or formalization of narratives, no inter-annotator analysis has ever been done, not even with the oldest and best-known formal approach to narrative structure, the Proppian narratemes [22]. In 2011, an annotation study has been set up at the *Universiteit van Amsterdam* (jointly by Bod, Saraf and the author) to train annotators to do annotations in the Proppian system and study inter-annotator agreement for this formalization task. A formal framework that does not allow for a stable inter-annotator agreement is not fit for the task at hand and would violate the requirement of simplicity discussed at the end of § 3. The Proppian study will serve as a template for future empirical grounding of decisions for or against particular formal systems.

References

1. R. Artstein and M. Poesio. Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4):555–596, 2008.
2. R. Bod, B. Löwe, and S. Saraf. How much do formal narrative annotations differ? A Proppian case study, 2011. Abstract of a presentation at IACAP 2011, Aarhus, 4–6 July 2011.
3. J. Bogen and J. Woodward. Saving the phenomena. *Philosophical Review*, 97(3):303–352, 1988.
4. J. C. Carletta, A. Isard, S. Isard, J. Kowtko, G. Doherty-Sneddon, and A. Anderson. The reliability of a dialogue structure coding scheme. *Computational Linguistics*, 23(1):13–31, 1997.
5. S. B. Chatman. *Story and Discourse: Narrative Structure in Fiction and Film.* Cornell University Press, 1980.
6. I. Cornelisse and N. Venhuizen. The influence of emotion and sympathy on the evaluation of story similarity, 2010. Student paper.
7. A. Dundes. Introduction. In A. Dundes, editor, *The Flood Myth*, pages 1–5. University of California Press, 1988.

8. M. G. Dyer. *In-depth understanding: A computer model of integrated processing for narrative comprehension*. Artificial Intelligence Series. MIT Press, 1983.
9. B. Falkenhainer, K. Forbus, and D. Gentner. The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 20:1–63, 1989.
10. K. D. Forbus, D. Gentner, and K. Law. Mac/fac: A model of similarity-based retrieval. *Cognitive Science*, 19:141–205, 1995.
11. D. Gentner. Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7(2):155–170, 1983.
12. D. Gentner, M. J. Rattermann, and K. D. Forbus. The roles of similarity in transfer: Separating retrievability from inferential soundness. *Cognitive Psychology*, 25:524–575, 1993.
13. S. Lam. Affective analogical learning and reasoning. Master's thesis, School of Informatics, University of Edinburgh, 2008.
14. G. Leech. Adding linguistic annotation. In M. Wynne, editor, *Developing Linguistic Corpora: a Guide to Good Practice*, AHDS Guides to Good Practice. Oxbow Books, 2005.
15. W. G. Lehnert. Plot units and narrative summarization. *Cognitive Science*, 4:293–331, 1981.
16. B. Löwe. Comparing formal frameworks of narrative structures. In M. Finlayson, editor, *Computational Models of Narrative. Papers from the 2010 AAAI Fall Symposium*, volume FS-10-04 of *AAAI Technical Reports*, pages 45–46.
17. B. Löwe and T. Müller. Data and phenomena in conceptual modelling. *Synthese*, 2011. to appear.
18. B. Löwe and E. Pacuit. An abstract approach to reasoning about games with mistaken and changing beliefs. *Australasian Journal of Logic*, 6:162–181, 2008.
19. B. Löwe, E. Pacuit, and S. Saraf. Identifying the structure of a narrative via an agent-based logic of preferences and beliefs: Formalizations of episodes from CSI: Crime Scene Investigation™. In M. Duvigneau and D. Moldt, editors, *Proceedings of the Fifth International Workshop on Modelling of Objects, Components and Agents. MOCA'09*, FBI-HH-B-290/09, pages 45–63, 2009.
20. D. Marcu, M. Romera, and E. A. Amorrortu. Experiments in constructing a corpus of discourse trees: Problems, annotation choices, issues. In *Workshop on Levels of Representation in Discourse*, pages 71–78, 1999.
21. G. Polti. *Les trente-six situations dramatique*. Mercure de France, 1895.
22. V. Propp. *Morfologiya skazki*. Akademija, Leningrad, 1928.
23. M. J. Rattermann and D. Gentner. Analogy and similarity: Determinants of accessibility and inferential soundness. In *Proceedings of the Ninth Annual Conference of the Cognitive Science Society*, pages 23–35, Hillsdale NJ, 1987. Lawrence Erlbaum.
24. D. E. Rumelhart. On evaluating story grammars. *Cognitive Science*, 4:313–316, 1980.
25. R. C. Schank. *Dynamic memory: A theory of reminding and learning in computers and people*. Cambridge University Press, 1982.
26. M. Si, S. C. Marsella, and D. V. Pynadath. Thespian: Using multi-agent fitting to craft interactive drama. In M. Pechoucek, D. Steiner, and S. Thompson, editors, *International Conference on Autonomous Agents. Proceedings of the fourth international joint conference on Autonomous agents and multiagent systems. Utrecht, The Netherlands, July 25–29, 2005*, pages 21–28, 2005.
27. P. Thagard and C. P. Shelley. Emotional analogies and analogical inference. In D. Gentner, K. H. Holyoak, and B. K. Kokinov, editors, *The analogical mind: Perspectives from cognitive science*, pages 335–362. MIT Press, 2001.

28. S. Thompson. *Motif-index of folk-literature: a classification of narrative elements in folktales, ballads, myths, fables, medieval romances, exempla, fabliaux, jest-books, and local legends.* Indiana University Press, 1955-1958.
29. S. Turner. *The creative process. A computer model of storytelling.* Lawrence Erlbaum Associates, 1994.
30. N. Wardrip-Fruin. *Expressive Processing. Digital Fictions, Computer Games and Software Studies.* MIT Press, 2009.