Using Unexpected Simplicity to Control Moral Judgments and Interest in Narratives*

Antoine Saillenfest and Jean-Louis Dessalles

Department of Computer Science and Networks (INFRES)
Telecom ParisTech
Paris, France
`{antoine.saillenfest,jean-louis.dessalles}@telecom-paristech.fr`

Abstract

The challenge of narrative automatic generation is to produce not only coherent, but interesting stories. This study considers the problem within the Simplicity Theory framework. According to this theory, interesting situations must be unexpectedly simple, either because they should have required complex circumstances to be produced, or because they are abnormally simple, as in coincidences. Here we consider the special case of narratives in which characters perform actions with emotional consequences. We show, using the simplicity framework, how notions such as intentions, believability, responsibility and moral judgments are linked to narrative interest.

1   Introduction

Whatever a story is used for (whether to entertain or to teach), to address the question of what makes a story interesting is of major importance. The field of computational generation of narratives has explored many ways to generate narratives. A well-formed and understandable story can make a *good* story but not necessarily an interesting one.

Many studies have addressed the question of what makes a story interesting and different factors have been listed. One of them is the reference to major life themes that elicit strong affective reactions such as death, sex, religion or politics [16, 28]. Another factor is the occurrence of unexpected or unusual events [16, 28].

Beyond these universal determining factors, narrative interest may also crucially depend on the audience's personal values, personal experiences, personal emotions or pre-existing knowledge [29]. Taking into account personal values or personal experiences is a major challenge, but it also offers opportunities for automated scenario generation, especially for the replay options they offer.

Our review of previous works on narrative interest revealed that moral considerations towards the characters play a crucial role. We investigated how interest and moral judgments are related, in the specific context of moral dilemma stories [27]. We identified a set of factors that can be manipulated to control both interest and moral response. Using a framework

---

based on the notion of Kolmogorov complexity, we proposed a unified model of interest and moral judgment that can be applied to moral dilemma stories [27]. The present paper extends these previous results. It proposes new ways of controlling narrative interest by manipulating moral judgments, intentions and responsibility in narratives.

2 Some ingredients of interestingness

Humans are narrative beings. From childhood to adulthood, they are surrounded by narratives. They use them to make sense of the world, to order events and assimilate them. This process has been called *narrative intelligence* [21]. Narratives are a cognitive tool to organize experience and understand encountered situations. They play a major role in many forms of entertainment media including theater, novels or video games. They are also widely used in educational contexts.

The process of generating narratives has been the subject of intense study in Artificial Intelligence over the last century (see [11] for a review of existing systems). Different approaches to narrative generation have been explored. The standard approach consists in pre-scripting a specific story; characters perform the same actions without any variation each time the program is run. Systems using this approach present a limited number of stories (or permutations of a single story), they show little adaptation to the user and offer few opportunities of user interaction. An alternative approach consists in generating the narrative dynamically. Systems using this latter approach are able to adapt to individual preferences, needs, abilities or values. Such systems have extended replay capabilities.

A narrative is classically defined as the recounting of a sequence of events that have a continuant subject and constitute a whole [26]. A basic form of narrative is the *fabula*, a temporally ordered sequence of events from the time the story begins to the time the story ends [1]. The fabula is only one component of a typical narrative; the other one is the *discourse*, which refers to the way the narrative is told. The process of telling a narrative consists in selecting a subset of the fabula (the *sjužet*) and in structuring the outline of the main events in order to elicit interest in the audience [1, 11].

Generating a narrative consists in producing a sequence of events which brings a world from an initial state into a final state. This narrative content is then put into words to give the discourse. This process must meet some requirements to form acceptable narratives. One of them is to be well-structured, in the sense that it respects Freytag's triangle: the narrative should have a beginning, where some conflict is introduced, a middle, containing the climax, and an ending [9]. Another requirement for the generation of an acceptable narrative is to be understandable by the audience. In particular the events should respect the causal rules of the (possibly imaginary) world and the audience must be able to infer the characters' intentionality [3].

This generation process, however, does not ensure that these well-formed and under-standable narratives make interesting stories. Producing interesting stories represents a major challenge, because of the crucial role of interest in learning or entertaining contexts [14]. Some of the factors controlling narrative interest have been investigated. They include emotional responses to fundamentals such as sex, religion or death [28, 16]. Surprising or unexpected events are also essential to interest (see below).

Few models of narrative interest have been proposed. Simplicity Theory [7] is one of them. It offers a purely cognitive account of interest, based on an extension of the notion of Kolmogorov complexity to the cognitive domain [4]. The main claim of Simplicity Theory is that unexpectedness, defined as a complexity drop, plays a major role in eliciting

narrative interest. It covers various situations humans usually regard as interesting, including coincidences, fortuitous encounters or rare events [6]. Simplicity Theory has also been applied to the study of moral judgment in the context of moral dilemma stories [27].

The observation that moral dilemmas often make good stories led us to investigate which factors control both narrative interest and moral judgments. Previous studies had pointed out the crucial role played by emotions and causal reasoning in the formation of moral judgments [12]. By applying Simplicity Theory to moral dilemma stories, we could show that factors such as the length of causal chains or the interpersonal relations between the characters can be manipulated in order to control interest and moral (dis)approval [27].

The formation of judgments on characters is an integral part of the narrative experience and is involved in the perception of character believability. Believable characters must be first perceived as intentional agents by the audience, who will then approve or disapprove of their actions. Simplicity Theory offers a new way to define and control, not only the interest of a generated story, but also the believability of its characters.

We will first review the main factors that influence moral judgments in narratives and point out the importance of these judgments for the believability of its characters. Then we will briefly introduce Simplicity theory and show how it can be applied to model intentions and moral judgments. We will illustrate how some parameters in the model can be used to control narrative interest and moral judgments. Finally, we will discuss the implications for narrative generation.

3   Interesting stories with moral characters

3.1   Believability and interestingness

Believable characters in a story should be perceived as clever and their actions as intentional [5], or else the audience's suspension of disbelief will be negatively affected [3]. But being perceived as intentional is not enough. The audience also forms various moral judgments including moral approval and responsibility. Approving or disapproving of others' actions and attributing responsibility is daily routine. We do it also when experiencing narratives. Let us consider the following story fragment.

The [**Wine Story**]. *Mary is looking through the window of her manor, waiting for a taxi. John enters the room, carrying a bottle of vintage red wine. He takes two crystal glasses in the cupboard and fills them. He discretely adds a pinch of black powder in one of them and brings it to Mary. The taxi will arrive in only a few minutes. Mary starts talking about her work, waving the hand and the glass. They hear the taxi's horn. She takes a sip from the glass, grabs her handbag, drinks the glass in one gulp and goes out.*

At this point, no one knows for sure what will happen next. One assumes that John performed his act (add powder to the drink) with some specific goal in mind. Imagine that Mary dies. If Mary is perceived as a nice person by the audience, John's act will be condemned. But the context can be manipulated to change this judgment, for example if John has been threatened to be killed if he did not kill Mary. This piece of information may also change our judgment about John's responsibility.

In any case, John appears as a believable agent. But this is only one ingredient of interest. Will the story be more or less interesting if Mary does not die? If John had selfish or altruistic motives (inherit Mary's fortune vs. save many from Mary's dark plans)? John's intentions and responsibility, in the eye of the audience, are a key factor in determining narrative interest.

To generate interesting stories automatically, one must first understand which factors influence our moral perception of the characters' actions. We will first review some previous works on morality and narrative interest before proposing our own approach.

3.2 Moral Judgments, Intentions and Responsibility

During the 1950s and the 1960s, studies on moral psychology focused on reasoning. Cognitive models of information processing were the preferred framework for the study of the formation of moral judgments [17]. However, in the 1980s, the idea that moral emotions played a significant role has been put forward. Recent evidence suggests that moral judgment is a matter of emotion *and* of reasoning, though automatic emotional processes tend to dominate [12, 13].

The formation of moral judgments occurs as soon as a character in a narrative performs an action that has good or bad consequences. The character's causal reasoning before the action is an essential determining factor that affects moral judgment. Jones (1991) proposed that the certainty or probability that the effect will follow from the action highly influences the intensity of our moral judgment [15]. In a previous study [27], we could observe that the *unexpectedness* of the outcome has a major influence on the moral approval of a character's action.

Causal reasoning is of major importance not only for the approval of others' actions, but also in the process of attributing intentions and responsibility. People are generally considered to be responsible for at least some consequences of their actions [30, 19], because they are supposed to be able to anticipate these consequences. The influence of causal analysis in the way responsibility is attributed to agents has been investigated [22, 23]. For an action to appear intentional, the agent must believe that this action can be performed and will lead to a specific outcome [20].

Moral judgment depends not only on causal reasoning, but also on the affective or emotional attitude that the characters or the audience have towards the consequences of the actions performed. Jones (1991) calls it the magnitude of the consequences [15]. The emotional response to a state of affairs determines what is desired or not and to what extent. It is essential to determine the character's intention to reach some specific goal [20]. Phares and Wilson (1972) also showed that it is of major importance in the attribution of responsibility [24].

Determining if a character in a story intentionally produced a situation or is responsible for it requires that we access its mental state at the time of the action [2]. Accessing mental states means inferring the character's knowledge and emotional state. In believable stories, sufficient information should be provided for the audience to assess the characters' mental states and to form moral judgments.

3.3 From moral judgments to interest

Our review of previous works on narrative interest revealed that moral considerations towards characters play a crucial role. We decided to investigate how interest and moral judgment are related, in the specific context of moral dilemma stories [27]. We especially pointed out the different but complementary roles of unexpectedness and emotional intensity in moral judgment and narrative interest. For example, we found that readers are more likely to disapprove of an action leading to undesired outcomes when the action appears more unexpected. But at the same time, the narrative interest of the sequence increases.

Using the framework of Simplicity Theory (see next section), we proposed a purely cognitive and unified model of interest and moral judgment that can be applied to moral dilemma stories. This model, based on the notion of Kolmogorov complexity, highlights the role of unexpectedness in the elicitation of emotional responses to events. The model proposed in our previous work explains our observations and makes correct predictions in the context of moral dilemma stories, taking into account the personal values of the audience and the estimated personal values of the characters of the story.

The present article extends these previous results and proposes new way of controlling narrative interest by manipulating judgments of intentionality and responsibility.

4   The key role of unexpectedness

4.1   Unexpected simplicity

As Livia Polanyi observed thirty years ago: "the question of what [people] tell stories about has remained not only unanswered, but largely unasked." [25, p. 207]. Polanyi, as the few authors who addressed the issue, acknowledged the importance of *unexpectedness* in spontaneously reported stories "stories have as their point that something 'odd' or 'unexpected' happened." [25, p. 212]. William Labov makes a similar observation: "[If an] event becomes common enough, it is no longer a violation of an expected rule of behaviour, and it is not reportable." When reporting an event, people say "[this] was strange, uncommon, or unusual – that is, worth reporting. It was not ordinary, plain, humdrum, everyday, or run-of-the-mill." [18]. Teun van Dijk also points the importance of unexpectedness: "interestingness is usually obtained by the account of events or actions that are unexpected, deviant, extra-ordinary, or unpredictable, given the knowledge and beliefs of the audience." [31, p. 123]. The same law of unexpectedness applies to newsworthiness: "Events have to be unexpected or rare, or preferably both, to become good news." [10, p. 67].

Unexpectedness stands out as a key feature of interestingness in spontaneous conversation. This is especially true for events leading to moral evaluation. Situations of deviant behaviour, leading to positive or negative moral judgments, constitute by definition exceptions to an expected norm, and are therefore unexpected. This may contribute to explaining why conversations are replete with moral evaluation and gossip [8].

What is true of spontaneous event reports must be true of fiction, at least up to some point. People are likely to use the same cognitive dispositions when they enjoy stories about factual events and about fictitious ones. Our hypothesis is that most parameters that control interest, both in conversation and in fiction, do it by controlling unexpectedness. It is thus important to reach a formal definition of that notion.

4.2   Defining unexpectedness

Unexpectedness is intuitively associated with an impression of low probability. Rare events, remarkable happenings, oddities, exceptions, deviations from norms are all extra-ordinary. They contrast with ordinary situation which, almost by definition, conform to expectations and are regarded as probable.

Probability unfortunately does not correctly capture the notion of unexpectedness. For instance, people may assess the unexpected nature of events that they would have previously regarded as impossible, such as the fact that two nuns start running on a jogging trail. Moreover, events of equal probability, such as lottery draws, are not regarded as equally

expected. A draw like 1–2–3–4–5–6 would be regarded as highly unexpected, whereas a "normal" draw like 13–23–24–31–35–44 is just boring news for those who did not play.

Simplicity Theory [7] is an attempt to capture the notion of unexpectedness. Unexpected (and therefore interesting) situations share the property of being less complex than expected. The word 'complexity' is used here in its technical sense, meaning 'size of the most concise description'. Unexpectedness results from a contrast between expected complexity (size of the most concise explanation of the event) and observed complexity (size of the most concise unambiguous designation of the event). In other words:

Unexpected situations are more complex to generate than to describe.

This definition can be applied to the lottery example: all draws are equally complex to produce (if the machine is not biased). Most of them are complex to describe as well: their best description merely enumerates the numbers. There are therefore not unexpected. The remarkable draws are those which can be concisely designated. If a draw is a natural sequence such as 1–2–3–4–5–6, numbers need not be mentioned one by one and the designation, measured in bits of information, is much more concise. Similarly, exceptions are unexpected if what makes them exceptional is simple. A running nun is exceptional if all other nuns walk at a slow pace, because 'run' is a simple characteristic. A nun whose rosary has a missing bead may be unique as well, but as far as the 'rosary-with-missing-bead' characteristic is more complex than 'running', the event is less likely to be perceived as unexpected, and therefore as interesting.

Complexity is a concept of theoretical computer science. The corresponding quantity is measured in bits. Its abstract definition is the size of the shortest computer program, expressed in binary form, that outputs the object under consideration. We must distinguish two aspects of the notion.

Generation complexity: Size of the minimal instruction set that leads to a causal generation of the event.

Description complexity: Size of the minimal set of characteristics that leads to an unambiguous designation of the event.

Unexpectedness is defined by the difference (see also the Appendix):

$$
\begin{array}{rl}
 & \text{generation complexity} \\
- & \text{description complexity} \\
\hline
= & \text{unexpectedness}
\end{array}
\tag{1}
$$

When some feature, such as the fact of running, makes a situation (running nuns) unique, describing the situation once the feature is known requires no additional information. This explains why simple unique features (running, as opposed to missing-bead-in-rosary) make the situation more interesting.

4.3 Unexpectedness and emotions

Simplicity Theory has much to say, not about emotional experience *per se*, but about emotional *intensity*. The basic idea is that emotional intensity is controlled by unexpectedness. In what follows, any reference to 'emotion' has to be understood as meaning 'emotional intensity'.

Unexpectedness plays a decisive role in the **Wine Story**. If Mary suddenly dies, her death will be totally unexpected to all witnesses but John. Of course, any one may die

any time, but Mary's spontaneous death is unlikely. The difference between actual emotion (when Mary suddenly dies) and hypothetical emotion (considering Mary's possible death) is unexpectedness (see also the Appendix).

$$
\begin{array}{ll}
 & \text{actual emotion} \\
- & \text{hypothetical emotion} \\
\hline
= & \text{unexpectedness}
\end{array}
\tag{2}
$$

From this equation, we can deduce that the following terms are interdependent:

- the observer's emotional reaction to Mary's death (actual emotion)
- the observer's concern about Mary's possible death (hypothetical emotion)
- the complexity of the simplest causal scenario leading to Mary's death that the observer may think of (generation complexity).

If the observer never anticipated Mary's death, the reference may change. If the observer compares Mary's death with her/his own, the equation will involve the following terms:

- the observer's emotional reaction to Mary's death (actual emotion)
- the observer's concern about her/his possible own death (hypothetical emotion)
- the complexity of the simplest causal scenario leading to Mary's death that the observer may think of (generation complexity)
- the minimal description of Mary for the observer (description complexity)

This explains why Mary's death will be perceived differently by different people, depending on their closeness to the victim.

4.4   Emotions and actions

If the black powder is poison, Mary's death is not unexpected from John's perspective. When Mary dies, John is not surprised and his emotion does not change... unless the causal link from his action (adding the powder) to the effect (Mary's death) is not perfect. In that case, the preceding equation can be rewritten as:

$$
\begin{array}{ll}
 & \text{anticipated emotion} \\
- & \text{causal unexpectedness} \\
\hline
= & \text{hypothetical emotion}
\end{array}
\tag{3}
$$

Causal unexpectedness is just an instance of the general notion of unexpectedness. It is evaluated when an agent performs an action. If the causal link to the emotional outcome involves many intermediary steps, the agent will perceive the outcome as more complex to generate. The outcome will be more unexpected and the gap between hypothetical and actual emotion will be larger. In the **Wine Story**, adding poison to the drink is not enough to provoke Mary's death. John's act must remain unnoticed, Mary should drink the glass of wine, the poison has to be effective, at a time when Mary is far from a hospital, and so on. The occurrence of Mary's death is still unexpected, and thus emotional, from John's perspective.

4.5   Intentionality, responsibility and moral judgment

The preceding results follow from Simplicity Theory's basic principles. They already provide theoretical tools to control emotional intensity when events are told in a narrative, from anyone's perspective (readers or audience, but also other characters). When it comes to

anticipating the emotional effects of characters' actions, some additional notions such as intentionality, responsibility and moral judgments must be given precise definitions as well.

Our previous works on moral judgment suggest that narrative interest and moral approval of one's action can be controlled by various factors such as the causal unexpectedness or the emotional intensity of the outcomes [27]. We will now use the framework of Simplicity Theory to define the notions of intention and responsibility.

Intentions are a matter of desires. One intends to perform an action if one desires some resulting situation to happen. The degree of intention to perform an action in order to produce some desired outcome depends on the anticipated emotional benefit. In the **Wine Story**, suppose that Mary dies because John added some poison in her glass. From John's perspective, adding poison into Mary's glass makes her death highly expected. John's degree of intention to kill Mary by adding poison depends on his expected emotional response to her death. In other words, the more John desires that Mary dies, the more his act of adding poison is judged intentional. Intention can be defined using the following equation (see also the Appendix).

$$
\begin{array}{rl}
 & \text{agent's anticipated emotional response to the outcome} \\
- & \text{causal unexpectedness (from the agent's perspective)} \\
\hline
= & \text{intention of performing the action to produce the outcome}
\end{array}
\tag{4}
$$

Some anticipated consequences of one's action may not be desired. These negative anticipated consequences tend to decrease the global intention to perform an action. In such cases, these negative outcomes are brought up knowingly. We must suppose that the action has been performed for some other purpose.

When an action has several emotional effects, as in moral dilemma, the various intentional values add up, with negative signs when the agent's emotion attached to the outcome is negative.

$$
\begin{array}{rl}
 & \text{intention of performing the action to produce the outcome 1} \\
+/- & \text{intention of performing the action to produce the outcome 2} \\
+/- & \dots \\
\hline
= & \text{intention of performing the action}
\end{array}
\tag{5}
$$

Responsibility is defined in a similar way. There is, however, a crucial difference. The observer's emotions (or sometimes standard social emotions) must be substituted for the agent's ones. The responsibility attributed by an observer to an agent can be defined using the following equation (see also the Appendix).

$$
\begin{array}{rl}
 & \text{observer's anticipated emotional response to the outcome} \\
- & \text{causal unexpectedness (from the agent's perspective)} \\
\hline
= & \text{agent's responsibility for the outcome}
\end{array}
\tag{6}
$$

The observer's moral judgment about the agent's action is a cumulative responsibility (taking signs into account). The signs depend on the desirability or the undesirability of the outcomes from the observer's perspective.

$$
\begin{array}{rl}
 & \text{responsibility for outcome 1} \\
+/- & \text{responsibility for outcome 2} \\
+/- & \dots \\
\hline
= & \text{moral judgment about the agent's action}
\end{array}
\tag{7}
$$

Using these notions of intention, responsibility and moral judgment based on Simplicity Theory, we can not only anticipate moral evaluations in a narrative, but also control the characters' credibility. Rational characters will perform actions that realize the best compromise between the different judgments they can anticipate. The same notions can also be used to control interest, by maximizing the observer's emotion.

5   Controlling interest and moral judgment in a narrative context

Unexpected simplicity is our best candidate to control narrative interest and moral judgments in narratives. In this section, we will see that, by modifying the context of the **Wine Story**, we can influence not only the moral judgments experienced when reading the story, but also its interestingness.

5.1   Characters' desires

In the **Wine Story**, John may attempt to poison Mary in order to inherit her fortune. If so, John's anticipated emotion concerning Mary's death depends on various factors including her wealth and John's personal feelings for her. John's desires therefore modify his intention to act (see equations 4 and 5). For John to appear as a believable character, his action must make sense, which means that he should only perform acts that are significantly intentional. Equation 4 explains why actions should be consistent with the agent's desires.

Readers' emotional reaction to Mary's death is crucial to know whether John will be blamed or praised (see equations 6 and 7). This reaction depends on various factors, such as their empathy for Mary or their personal identification with her. These effects are captured by Mary's simplicity in the reader's eyes (see equation 2), whether she is for instance a main or a peripheral character.

Anticipated emotions are crucial in the case of moral dilemmas. Imagine that John loves Mary but is forced to kill her or else he will be killed himself. John anticipates two emotions, one for Mary's death, another one for his own. If John chooses to kill Mary, then the higher his anticipated emotional response to Mary's death, the more interesting the story. In such a situation, John kills Mary with a lower degree of intentionality (equation 5). His emotional reaction to the outcome is compensated by the fact that he will not get killed himself. The dilemma also affects the observers' attribution of responsibility if they have similar emotions (equation 7). Observers (readers or other characters of the story) who are ignorant of the death threat hanging over John may form a rather different judgment. To them, John's act will be highly unexpected (see equation 2). This example highlights the importance of the way information is provided to readers and to the characters of the story.

5.2   Unexpectedness in causal chains

When reading the **Wine Story**, one can easily imagine that John is attempting to poison Mary. Now imagine that the poison turns out to have no effect on her. This introduces surprise in the story because the end of the sequence of events is unexpected. There are different ways to introduce unexpectedness in a causal sequence during the process of story generation. The action itself can be unexpected for the reader, for the acting character or for any other character of the story (see the Appendix for a commentary about inadvertence); the expected consequence may fail to happen for some unanticipated reason; or the action provokes not only its expected outcomes, but other unanticipated consequences as well.

5.2.1   Action unexpectedness

In a written story, a character's action may be unexpected for the reader. Action unexpectedness should be manipulated carefully, as it may negatively affect the reader's "suspension of disbelief". A way to make an action unexpected is to keep the agent's goals untold. However, sufficient information has to be disclosed for readers to be able to infer a causal explanation for the unexpected action. Not only readers, but also other characters in the story (as readers imagine them) may also be surprised by a character's action. Even the acting character may be surprised by her/his own action, when it has been performed inadvertently. In such case, inadvertence may considerably reduce blame or praise (equation 7 and the Appendix).

5.2.2   Causal unexpectedness

Causal unexpectedness refers to the unexpectedness of the anticipated consequence once the action has been performed. If the causal chain of events that goes from the action to its anticipated consequence consists in many intermediary events, the occurrence of the consequence will appear less expected. More causal unexpectedness decreases the attribution of intentions (equation 4) and of responsibility (equation 6), and the amplitude of the blame or praise will diminish as well (equation 7). But more causal unexpectedness increases the gap between hypothetical emotion and actual or anticipated emotion, both for the acting character and for observers (equations 2 and 3). Note that in all cases, causal unexpectedness is supposed to be assessed by the agents, with the knowledge and inference capabilities that observers grant them.

In the **Wine Story**, we suppose that John knows that the poison will be effective. The only unexpected step, from John's perspective, is to know whether Mary will drink up her glass of poisoned wine, as the unexpectedness of Mary's death once she has drunk it is nearly zero. There is room for suspense in the first phase, because of the time pressure (the taxi will arrive soon). The occurrence of "Mary drinks the wine" becomes increasingly unexpected as the time is running out.

Causal unexpectedness should also be manipulated with care, as it may negatively affect readers' suspension of disbelief. There is a risk that agents be no longer believable if their intentionality drops too much (equation 4).

5.2.3   Unexpected consequences

Of course, the course of events may differ from what characters and readers anticipate. The simplest case is when the outcome is the exact opposite of the anticipated one. In the **Wine Story**, Mary may not die nor even get sick from having drunk the wine. This turn of events is unexpected, because observers must imagine complex circumstances that may have produced the alternate ending (equation 1). Such unexpectedness must be resolved, at least for readers. This means that some explanation must be found or must be provided that provides a simpler generation process for the event.

Any action may have various consequences besides its expected consequences. These side-effects are interesting, because they are unexpected. But as for deceived anticipations, their unexpectedness must be resolved, at least for readers. Unexpected consequences are not intentional since the agent did not anticipate them (equation 4), but the agent may be judged responsible for them (equation 6).

6   Conclusion

We have shown various ways through which Simplicity Theory can be used to manipulate emotional responses and interest in narratives. The same parameters also affect the believability of characters. The main parameter underlying narrative experience is unexpectedness. It is defined as the difference between the complexity of the circumstances that brought the situation about and the amount of information required to describe it. Unexpectedness is a key factor to predict interestingness in conversational narratives. It is also involved in fictitious stories.

When actions with emotional consequences are performed in a story, unexpectedness is involved three times. Once when the consequences are unexpected, independently from the action that caused them (e.g., when a character dies). A second time if the action itself is unexpected (e.g., if the action has been performed inadvertently, as in an accident). And a third time when the causal link is regarded as unexpected by the agent (e.g., when many steps separate the action from its emotional consequences). The various roles of unexpectedness offer as many possibilities to control interest in narratives.

This paper investigated these possibilities from a theoretical point of view. We are currently planning to implement them in a narrative generation program. The main difficulty comes from the fact that the precise values of complexity are sometimes difficult to compute. Various strategies can be used to overcome the problem. Generation complexity relies on the addition of choice points (see www.simplicitytheory.org). Description complexity for imaginary objects or characters is easy to compute by ranking those objects or characters by importance. For objects or people of real life like the Eiffel Tower or Barack Obama, a rough estimate can be deduced from the number of hits on a Web search engine.

Our conviction is that computational models of narrative cannot ignore the various parameters that control unexpectedness, emotions and moral judgments. This paper was a first attempt to list these parameters, to provide tentative definitions and to consider their logical relations.

A   Appendix

Here is a formal summary of the notions presented in this article. Corresponding equations in the article and in this appendix share the same number.

Unexpected situations are more complex to generate than to describe

$$U(s) = C_w(s) - C_d(s) \tag{1}$$

$U(s)$ is the *unexpectedness* of situation s, as perceived by the observer. $C_w(s)$ measures the minimal amount of information that is necessary for the 'world' to produce s. This information evaluates the size of the minimal explication of the situation. It corresponds to the circumstances that the observer must imagine for s to happen. Description complexity $C_d(s)$ measures the quantity of information that the observer needs to unambiguously determine s.

The basic idea of Simplicity Theory is that emotional intensity is controlled by unexpectedness. It says that the expected or hypothetical emotional intensity $E_h(s)$ attached to the possible occurrence of an event s is related to the actual emotional intensity $E(s)$ attached to the effective occurrence of s:

$$E(s) - E_h(s) = U(s) \tag{2}$$

When some outcome s of an action a occurs, one can evaluate how necessary the action was, from any observer's point of view, to produce s. The previous equation can be rewritten as:

$$E(s) - U(s||a) = E_h(s) \tag{3}$$

The causal unexpectedness $U(s||a)$ is an instance of the general notion of unexpectedness. It evaluates the unexpectedness of a situation s once an action a has been performed. $E(s)$ is the emotional intensity attached to s as it has been anticipated. $E_h(s)$ is the hypothetical emotion attached to the occurrence of s.

The preceding results follow from Simplicity Theory's basic principles. They already provide theoretical tools to control emotional intensity when events are told in a narrative, from anyone's perspective (readers or audience, but also other characters).

Now we will define some additional notions such as intentionality, responsibility and moral judgment.

The degree of intention to perform an action a in order to produce some desired outcome s is the anticipated emotional benefit. For any observer O, the degree of intention $(Int^O(A, a, s))$ for an actor A is the hypothetical emotion attached to the occurrence of s from A's perspective $E_h^A(s)$ (the upper letter indicates that the calculation is calculated by O from A's perspective by inferring A's mental state).

$$Int^O(A, a, s) = E^A(s) - U^A(s||a) \tag{4}$$

Undesired but anticipated outcomes contributes to decrease the global intention to perform an action. They are produced knowingly. We must suppose that the action has been performed either accidentally $(Int^O < 0)$ or for some other purpose. Accidental actions are unexpected. This unexpectedness is a negative source of intentionality. Equation 4 should be rewritten: $Int^O(A, a, s) = E^A(s) - U^A(s||a) - U^A(a)$. For voluntary actions, $U^A(a) = 0$.

The global intention to perform an action a assigned by O to A results from the accumulation of the degrees of intention attached to all the outcomes s, with negative signs for undesired ones $(\epsilon(s) = -1)$, in the set S of all the outcomes that have been anticipated by A.

$$Int^O(A, a) = \sum_{s \in S} \epsilon(s) Int^O(A, a, s) \tag{5}$$

Responsibility is defined in a similar way. There is, however, a crucial difference. The observer's emotions $(E^O(s))$ must be substituted for the actor's ones $(E^A(s))$. An observer will find an actor more responsible if, in the same circumstances, the actor's actions elicit a more intense emotional response in the observer. Then the responsibility attributed to A by O for the occurrence of the outcome s which is a consequence of A's action a is:

$$Res^O(A, a, s) = E^O(s) - U^A(s||a) \tag{6}$$

O's moral judgment (or moral approval) MJ about A's action is a cumulative responsibility (taking signs into account). The signs depend on the desirability or the undesirability of the outcomes for the observer O.

$$MJ^O(A, a) = \sum_{s \in S} \epsilon(s) Res^O(A, a, s) \tag{7}$$

─── **References** ───

1   Mieke Bal. *Narratology: Introduction to the theory of narrative.* University of Toronto Press, 1997.

2   Simon Baron-Cohen. The evolution of a theory of mind. In Michael C. Corballis and Stephen E. G. Lea, editors, *The descent of mind: Psychological perspectives on hominid evolution*, pages 261–277. New York, NY: Oxford University Press, 1999.

3   Joseph Bates. The role of emotion in believable agents. *Communications of the ACM*, 37(7):122–125, 1994.

4   Nick Chater and Paul M. B. Vitányi. Simplicity: a unifying principle in cognitive science? *Trends in Cognitive Sciences*, 7(1):19–22, 2003.

5   Daniel C. Dennett. *The intentional stance.* MIT Press, Cambridge, MA, 1989.

6   Jean-Louis Dessalles. Coincidences and the encounter problem: A formal account. In B. C. Love, K. McRae, and V. M. Sloutsky, editors, *Proceedings of the 30th Annual Conference of the Cognitive Science Society*, pages 2134–2139, Austin, TX, 2008. Cognitive Science Society.

7   Jean-Louis Dessalles. *La pertinence et ses origines cognitives : nouvelles théories.* Hermes-Science Publications, Lavoisier, 2008.

8   Robin Dunbar. *Grooming, gossip, and the evolution of language.* Harvard University Press, 1998.

9   Gustav Freytag. *Freytag's technique of the drama: an exposition of dramatic composition and art.* Scott, Foresman, 1894.

10   Johan Galtung and Mari H. Ruge. The structure of foreign news, the presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research*, 2(1):64–90, 1965.

11   Pablo Gervás. Computational approaches to storytelling and creativity. *AI Magazine*, 30(3):49–62, 2009.

12   Joshua D. Greene and Jonathan Haidt. How (and where) does moral judgment work? *Trends in Cognitive Sciences*, 6(12):517–523, 2002.

13   Joshua D. Greene, Leigh E. Nystrom, Andrew D. Engell, John M. Darley, and Jonathan D. Cohen. The neural bases of cognitive conflict and control in moral judgment. *Neuron*, 44(2):389–400, 2004.

14   Suzanne Hidi. Interest and its contribution as a mental resource for learning. *Review of Educational Research*, 60(4):549–571, 1990.

15   Thomas M. Jones. Ethical decision making by individuals in organizations: An issue-contingent model. *The Academy of Management Review*, 16(2):366–395, 1991.

16   Walter Kintsch. Learning from text, levels of comprehension, or: Why anyone would read a story anyway. *Poetics*, 9(1–3):87–98, 1980.

17   Lawrence Kohlberg. *The development of modes of moral thinking and choice in the years 10 to 16.* PhD thesis, University of Chicago, 1958.

18   William Labov. The transformation of experience in narrative syntax. In William Labov, editor, *Language in the Inner city*, pages 354–96. Philadelphia: Pennsylvania University Press, 1972.

19   David A. Lagnado and Shelley Channon. Judgments of cause and blame: The effects of intentionality and foreseeability. *Cognition*, 108(3):754–770, 2008.

20   Bertram F. Malle and Joshua Knobe. The folk concept of intentionality. *Journal of Experimental Social Psychology*, 33(2):101–121, 1997.

21   Michael Mateas and Phoebe Sengers. Narrative intelligence. In M. Mateas and P. Sengers, editors, *Narrative Intelligence: Papers from the AAAI Fall Symposium*, number FS-99-01 in AAAI Technical Reports, pages 1–10, 1999.

22 Dale T. Miller and Saku Gunasegaram. Temporal order and the perceived mutability of events: Implications for blame assignment. *Journal of Personality and Social Psychology*, 59(6):1111–1118, 1990.

23 Ahogni N'gbala and Nyla R. Branscombe. Mental simulation and causal attribution: When simulating an event does not affect fault assignment. *Journal of Experimental Social Psychology*, 31(2):139–162, 1995.

24 E. Jerry Phares and Kenneth G. Wilson. Responsibility attribution: Role of outcome severity, situational ambiguity, and internal-external control. *Journal of Personality*, 40(3):392–406, 1972.

25 Livia Polanyi. So what's the point? *Semiotica*, 25(3-4):207–242, 1979.

26 Gerald Prince. *A Dictionary of Narratology.* University of Nebraska Press, Lincoln, NE, 1987.

27 Antoine Saillenfest and Jean-Louis Dessalles. Role of Kolmogorov complexity on interest in moral dilemma stories. In N. Miyake, D. Peebles, and R. Cooper, editors, *Building Bridges Across Cognitive Sciences Around the World, Proceedings of the 34th Annual Meeting of the Cognitive Science Society, Sapporo, Japan, August 1–4, 2012*, pages 947–952, Austin, TX, 2012. Cognitive Science Society.

28 Roger C. Schank. Interestingness: Controlling inferences. *Artificial Intelligence*, 12(3):273–297, 1979.

29 Gregory Schraw and Stephen Lehman. Situational interest: A review of the literature and directions for future research. *Educational Psychology Review*, 13(1):23–52, 2001.

30 Kelly G. Shaver. *The attribution of blame: causality, responsibility, and blameworthiness.* Springer-Verlag, 1985.

31 Teun Adrianus van Dijk. Stories and racism. In Dennis K. Mumby, editor, *Narrative and Social Control: Critical Perspectives*, pages 121–142, Newbury Park, CA, 1993. Sage Publications.