

A Knowledge Distribution Model to Support an Author in Narrative Creation

Hochang Kwon¹, Sukhwan Jung¹, Hyuk Tae Kwon², and Wan Chul Yoon¹

¹ Dept. of Knowledge Service Engineering, KAIST,

² Dept. of Industrial Systems Engineering, KAIST,
291 Daehak-ro, Yuseong-gu, Daejeon, 305-701, Korea
{hochang,raphael,htkwon,wcyoon}@kaist.ac.kr

Abstract. Adjusting the knowledge of characters and the reader is a critical task for an author in narrative creation. Throughout a narrative, both characters and the reader experience events according to their own timelines and perspectives. They interpret information accumulated through their experience and update knowledge to the narrative-world which the author constructed. In this paper, we present a *Knowledge Distribution Model* which supports an author in finely controlling the knowledge of characters and the reader. Within the model, the *Knowledge Structure* is constructed by connecting event, information, and knowledge. The *Knowledge State* is evaluated as the *degree of belief* under the knowledge structure. We adopted a probabilistic reasoning model to calculate the knowledge state. The change in knowledge state, defined as *Knowledge Flow*, is visually presented to the author. We designed a GUI prototype to implement the proposed modeling process, and demonstrated the knowledge flow with an actual cinematic narrative.

Keywords: knowledge distribution, knowledge structure, knowledge flow, narrative creation, authoring tool

1 Introduction

Narrative is an effective media to convey information and knowledge [1]. In a narrative, explained by a communication model [2], an author (sender) constructs a particular story-world composed of characters, background (space-time), and events to convey his/her message; the reader (recipient) understands this message by reconstructing it. Throughout this process, the author distributes various information/knowledge related to a message (theme) over the narrative, then the reader and characters (agents in narrative) comprehend them while they experience events. The author uses a variety of techniques and rhetorical devices to organize information/knowledge and control the knowledge of each agent in the narrative. This work is essential to attaining the completeness and a valuable narrative for several reasons:

First, the author's message is expressed by the information arrangement in the narrative. Adjusting the arrangement and the flow of the information is

important to maintain the semantic consistency of the whole narrative. Second, it is also an important issue for the author to design coherent characters. The characters knowledge is closely related to their actions because they must have corresponding information/knowledge to act coherently. Third, the reader's knowledge is about when and to what extent the reader conceives the changes in information. This is directly connected with the reader's understanding of the narrative. Fourth, adjusting differences in knowledge between characters and the reader is a powerful narrative method to make the reader participate in the narrative and feel sympathy [3]. It is strongly related to reader's emotional responses, including suspense, surprise, and curiosity.

In spite of the importance of this process, it has been manually done in general practice, depending solely on the ability and knowledge of the author. In this paper, we propose a novel model to both represent the structure of information/knowledge and control the knowledge of the agents in a narrative. We designed this model to support the author's narrative creation.

2 Related Work

The literature on constructing a model for the knowledge of the agents in narrative writing can be divided into three categories.

The first group is about modeling the knowledge of characters in a story-world. Riedl et al. [4] constructed a knowledge structure of characters to maintain *character believability* using IPOCL algorithm. In their system, the knowledge of a character is regarded as constraints and guidelines for action. The Belief-Desire-Intention (BDI) model, which is well-known in the field of software agents, was applied to character's behavior [5]. The character designed by the BDI model updates knowledge, modifies goals, and executes plans in the process of narrative events. The next approach utilizes modeling a reader's inference in the narrative comprehension. The most representative research related to this topic is the work of Graesser et al. [6,7], which is based on the constructionist theory in psychology. They described how a reader constructs a knowledge-base and explain why actions, events, and states are mentioned in the narrative text. Based on former studies, an algorithm for generating discourse plans that prompt a reader's inference was suggested [8]. The last category is the research on the suspense model using the disparity of knowledge between character and the reader. Many studies on the suspense model are based on the Structural Affect Theory (SAT) suggested by Brewer et al. [9], in which a specific discourse structure is set up by manipulating the order of events. Based on the SAT, plan-based models that evoke a sense of suspense in readers are also suggested [10,11].

Many of the previous researches have paid attention to applying AI methods to the narrative generation system. The quality of automatically generated narratives, however, are still limited despite of computational advances. It is necessary to construct more flexible model that guarantees an author's participation in narrative creation. The desirable model has to provide the author with intuitive interaction for seamless creation work and to visualize the quantified changes of knowledge in narrative.

3 Narrative and Knowledge

3.1 Narrative Structure and Agents' Experience

The structuralism model pointed out that a narrative consists of story (content) and discourse (expression), and the basic unit of content is event [12]. The basic attributes of an event can be revealed by four-W questions: *Who does what? When and where?* At the early stage in the conception of a narrative, a number of loosely connected events are created in the author's mind. The author elaborates a set of events and composes a narrative by configuring those events with reference to causal and temporal relations. The events are ordered along two different time axes (timelines): a chronological *story-time* in a story-world along which characters experience the events, and a plotted *discourse-time* along which readers follow. The reader and characters jointly experience the events in a narrative; however, throughout the narrative, they accumulate experience according to their own timelines and perspectives. Each of them interprets the meaning of accumulated information based on successive experience and updates knowledge to the narrative-world which the author constructed. Fig. 1 shows how the author configures the events and makes the disparity of knowledge in a narrative.

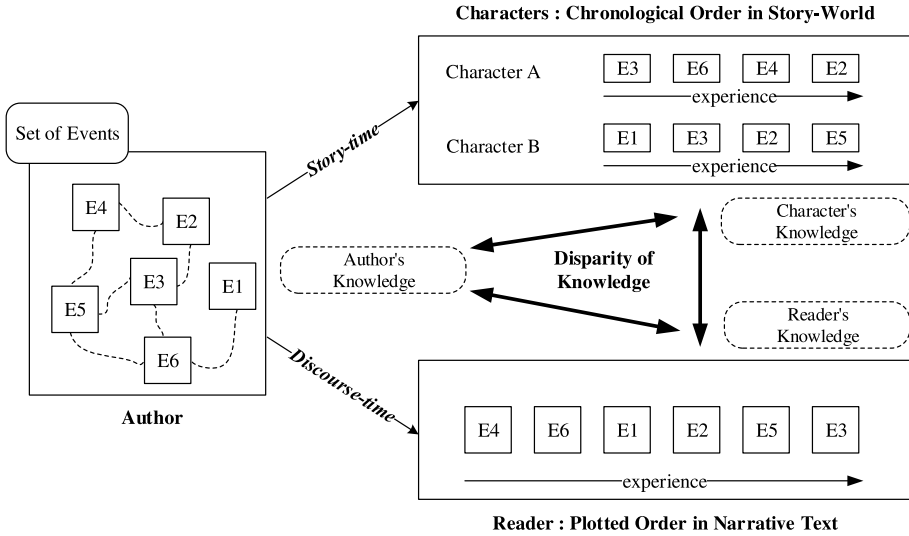


Fig. 1. Event configuration and the disparity of knowledge

3.2 Narrative Creation and Knowledge Distribution

In a narrative, only the author has all the information and knowledge. The author can adjust when, what and to what extent each agent gets to know by organizing events and information. We define this organizing process as *knowledge distribution* task, which must be carried out to create a valuable narrative. Simultaneously controlling a variety of knowledge of each agent, however, has fairly

heavy cognitive workload, hence we devised a model to support the author's knowledge distribution work; the proposed model consists of three functional parts as follows.

The first part is the *knowledge structure*, where the events are arranged and the information and knowledge are set to reveal to each of agents. Under this structure, agents experience the events to acquire information and knowledge along their own perspectives and timelines. The second part is a reasoning model for computational agent modeling. At any particular time-point in a narrative, agents reason about the *author-settled* knowledge based on their accumulated information. The reasoning model performs this inference of agents using the computational methods. This model is independent of the knowledge structure, thus any kind of inference method can be applied. In this paper, we adopt a Bayesian method to demonstrate the probabilistic reasoning about the knowledge of agents. The last part is the *knowledge state* and the *knowledge flow*. The knowledge state is defined as the *degree of belief* about the knowledge specified by the author, which is calculated by the reasoning model. Sequentially experiencing the events, the knowledge state of agents is subject to change between *belief* and *disbelief*. The knowledge flow is defined by the trajectory of change in the knowledge state.

The author can selectively adjust the knowledge structure to alter the knowledge state of agents in this model, which is immediately reflected by the change of their knowledge flow. It helps authors to elaborate the knowledge distribution work in an interactive manner as depicted in Fig. 2.

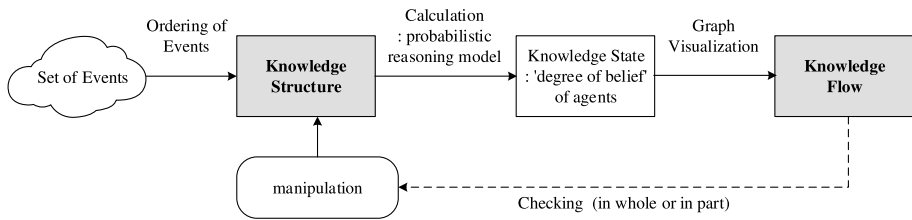


Fig. 2. Overall process of knowledge distribution

4 Framework of Knowledge Distribution Modeling

In this section, we provide a detailed description of the model and our approach. During the modeling process, the knowledge structure is constructed by relating the knowledge elements. The knowledge structure is the foundation of the knowledge state and flow calculated by a reasoning model on the assumption that a reader and characters are rational agents within a narrative.

4.1 Knowledge Structure

The proposed *knowledge structure* is modeled by connecting event, information, knowledge, and meta-knowledge. This structure is similar to the Data-Information-Knowledge-Wisdom (DIKW) hierarchy model [13] proposed in the

field of information science. Fig. 3 shows the framework of knowledge structure and the author conducts the *knowledge distribution* on this workbench.

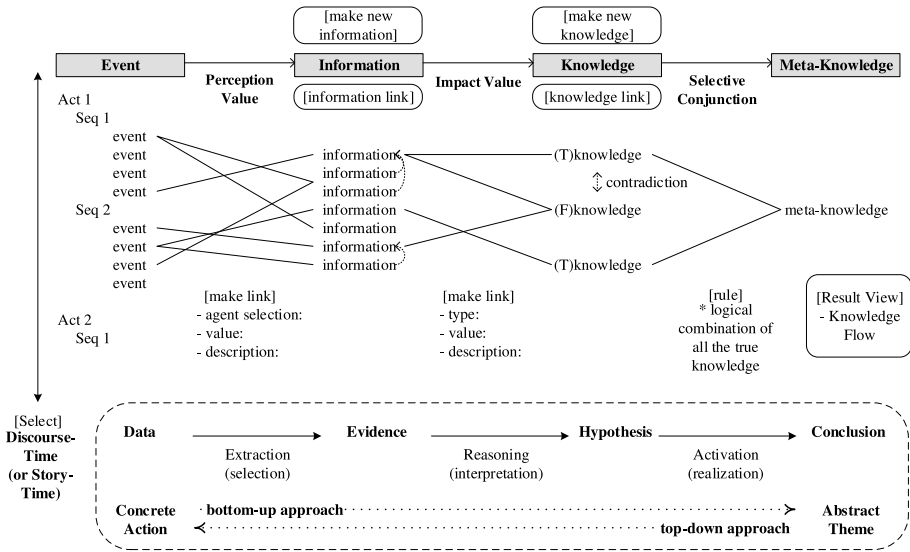


Fig. 3. Framework of knowledge structure and description

Event and Information. Events can be regarded as a set of data. Information that influences the knowledge state are extracted from a list of events. Agents obtain and accumulative the relevant information while experiencing the events. Each of agents has different perspective to events, hence information is obtained at a different level even for the same event. It is reflected by the *perception value* on the links between events and information. The author can finely express the difference between perspectives of agents in the narrative by assigning proper perception values for each agent.

Information and Knowledge. Information can be regarded as evidence of knowledge that the author set up within the narrative. The type and strength of the evidence are expressed by the *impact value* on the links connecting information and knowledge. Each item of information either concedes or rebuts the linked knowledge. Agents rationally reason about the knowledge based on the accumulated information in order to formulate and verify a hypothesis.

Knowledge and Meta-knowledge. Knowledge is expressed in a form of proposition, each with a value of *true* or *false*; the knowledge set up as false is a distractor that aims to confuse the agents. Meta-knowledge is the most important knowledge in the narrative because it is directly connected with the theme (message) that the author wants to convey. Meta-knowledge is a true

proposition which can be deduced by logically combining all the true knowledge. Agents realize the meta-knowledge when all the true knowledge are known.

In summary, agents have limited knowledge. As they experience events along their own timelines and perspectives, they accumulate information and reach a level of knowledge state by rational reasoning. Agents may or may not eventually reach a literary epiphany (meta-knowledge).

4.2 Knowledge State and Flow

The author configures and attributes events and information to affect the knowledge by constructing the knowledge structure, and manipulates the knowledge of agents by making some information overt or covert to them through a sequence of events in the narrative. In the knowledge distribution model, the author can verify the knowledge state and flow of agents by ordering events and assigning appropriate perception and impact values in the knowledge structure. The knowledge flow of an agent on specific knowledge is represented as a form of line graph, in which the x-axis represents the order of events in the narrative, and the y-axis represents the knowledge state after the agent experiences each event. The knowledge state expressed by the degree of belief ranges from -1 to 1, and its detailed meaning is shown in Table 1. Note that the degree of belief can be a qualitative value such as *doubtful*, *neutral*, *definitely positive* or *negative*, etc.

Table 1. Detailed meaning of degree-of-belief values of knowledge state

Degree-of-Belief	1	0.5	0	-0.5	-1
Level of certainty	maximum	middle	minimum	middle	maximum
Interpretation	inevitable	probable		improbable	impossible
Dominance of information link	Information linked to the knowledge is positive		★	Information linked to the knowledge is negative	
On true knowledge	understanding			misunderstanding	
On false knowledge	misunderstanding			understanding	

★ Degree of belief = 0 can be interpreted in two ways:
1) Ignorance: No information is linked to the knowledge.
2) Uncertainty: Information linked to the knowledge contradicts, rendering it unjustifiable.

The knowledge flows of agents differ in most narratives as they experience events on different timelines with individual perspectives. This causes each agent to experience events in a different sequence and magnitude. Note that the proposed model assumes a rational agent and that the reasoning processes and capabilities of characters and the reader are identical.

4.3 Reasoning Model: Analytic Approach

In this paper, we took an analytic approach to calculate the knowledge flow from the knowledge structure. Among diverse methods that can be utilized, a

probabilistic reasoning model based on Bayesian theory, one of the oldest and most traditional methods for dealing with uncertainty, was applied to calculate the inference process of narrative agents. For a hypothesis H , which is knowledge in our model, an agent's degree of belief in H is denoted as a *subject probability* of the agent, $P(H)$, that the agent believes H is true. Given a related event E to H , the degree of belief is denoted by $P(H|E)$.

The base model that we adopted was from the Abell's application of Bayesian probability to narrative inference [14]. It was initially intended to analyze narratives in social situations and can also be used to model the degree of belief in a narrative; only the agents are changed, from social individuals to characters and readers. The method uses the logarithm of the odds ratio of a hypothesis to show the degree of belief of an agent. With the use of odd ratio, the degree of belief changes more drastically as the magnitude of event increases. Applying the Bayes' rule, the agent's prior and posterior odds ratio are

$$Odds(H : \neg H) = \frac{P(H)}{P(\neg H)}, Odds(H : \neg H|E) = \frac{P(H|E)}{P(\neg H|E)} = \frac{P(E|H)P(H)}{P(\neg E|H)P(\neg H)}, \quad (1)$$

where the posterior odds ratio can also be shown as

$$Odds(H : \neg H|E) = L_E \times Odds(H : \neg H), \text{ if } L_E = \frac{P(E|H)}{P(E|\neg H)}. \quad (2)$$

To measure how event E changes the odds of H against $\neg H$, a logarithmic function is used, where the sign of $\log L_E$ shows how E changes the degree of belief in H [15].

$$\log Odds(H : \neg H|E) = \log L_E + \log Odds(H : \neg H). \quad (3)$$

Note that an appropriate clamping rule is needed since the logarithm cannot be evaluated in two extreme cases, $P(H) = 1$ and $P(H) = 0$.

There are two main differences between the assumptions of above Bayesian probability and our model. The one is that the values of the knowledge state varies from uncertainty to certainty with respect to two types of information (t): true or false, coded by 1 or -1, respectively. We adopted a linear mapping so that the closer $P(H)$ is to zero or one, the more certainly H is false or true, respectively, while the truth of H is uncertain where $P(H)$ is around 0.5. The other difference is that our model introduces the perception value as a measure of how much of the information is actually perceived by an agent. The *effective* impact value of information (v) can be defined as the product of the impact value and the perception value in the knowledge structure. Therefore, if each event carries single information, the equations must change as follows.

$$P(E|H) = \frac{1 + v \times t}{2}, P(\neg E|H) = \frac{1 - v \times t}{2}, \text{ where } 0 \leq v \leq 1, t = \begin{cases} -1 & \text{false} \\ 1 & \text{true} \end{cases} \quad (4)$$

$$Odds(H : \neg H|E_{i+1}) = L_{E_i} \times Odds(H : \neg H|E_i),$$

where $L_{E_i} = \frac{P(E_i|H)}{P(E_i|\neg H)} = \frac{1 + v_i \times t_i}{1 - v_i \times t_i}, Odds(H : \neg H|E_0) = 1.0. \quad (5)$

5 Case Study

In this section, the knowledge structure and flow of an actual cinematic narrative was analyzed using the proposed model. We chose a film *Incendies* (2010, Denis Villeneuve) as an example. In this film, the effect of war stained history on individuals is portrayed by a journey of twins following their mother's last will. This film would be suitable for validating the effectiveness of the proposed model for following reasons: First, it is a famous, multiple award-winning film with numerous reviews and critiques from which emotional responses for this film can be found. The dominant emotional responses of its viewers are found to be suspense and surprise by running keyword analysis on the reviews and critiques from *IMDB.com* and *Metacritic.com*. These emotional responses are closely related to the change in knowledge of agents as many researches on the SAT showed. Second, the main plot of the film is built around a certain secret created at the beginning of the film. The knowledge state of agents are constantly changing as the events are unfolded in narrative. This shows the importance of knowledge distribution task in narrative creation to elaborately adjust the knowledge flow of agents. Third, this film can be used as an example to show that the proposed knowledge distribution model is not only restricted to typical thriller or mystery narratives, but also can be applied to more generic drama narratives. Fourth, the knowledge flow of agents greatly varies with a notable difference between the story and discourse timeline, while the disparity of knowledge between the viewers and the characters are closely related to the theme of the narrative.

As a result of this analysis, thirty-seven events, seventeen information, six knowledge and one meta-knowledge were found from this film. Fig. 4 shows the knowledge structure of *Incendies* with the relevant list of events and information shown in Table 2. Based on the analyzed knowledge structure, Abell's Bayesian method for narrative inference model was applied to calculate the changes in the degree of belief for each agent. Fig. 5 shows different knowledge flows of the viewers and *Nawal* (the protagonist) in both story and discourse time.

Fig.5(a) and Fig. 5(b) represent how the viewer's knowledge changes over the different timelines. Fig. 5(b) shows the viewer's knowledge flow as he/she watches the film; it captures the viewer's realization of the fact that the twins (*Jeanne* and *Simon*) are siblings of *Nihad* (K4) in a late scene. A brief misunderstanding about the nature of *Abu Tarek's* child (K2) and the gradual realization of *Nihad's* fate is also captured. Such knowledge flows of the viewers are consistent with their emotional responses while watching the film. Viewers reinterpret the meanings of accumulated information when a sudden change is occurred in knowledge state, evoking suspense or surprise. Authors typically try to create such a change at the narrative's climax scene. Fig. 5(a), which depicts a potential knowledge flow of the viewers if the narrative is unfolded without flashbacks, shows a different knowledge flow; the viewers know about the identity of the twins' sibling much sooner and starts to have doubts about *Abu Tarek's* child after learning the truth. Comparison between two timelines clearly shows that the knowledge flow of the viewers can be changed by reordering the events. This is a frequent activity in narrative creation, which grants the responsive controllability to the author.

Table 2. List of events (a) and information (b) for *Incendies*

(a)			(b)
Discourse order	Story order	Event	ID Information
D7	S2	Nihad is born.	I1 Nawal is a mother of Nihad.
D13	S7	Orphanage is destroyed.	I2 Nawal is a mother of Jeanne and Simon.
D14	S8	Bus attacked by Christian terrorists.	I3 Abu Tarek raped Nawal.
D17	S9	Nawal joins Islamic extremists.	I4 Nawal gave birth to Abu Tarek's children.
D1	S10	Nihad is trained to become a soldier.	I5 Abu Tarek's children are twins.
D28	S11	Nihad becomes crazed with war.	I6 Nihad has a tattoo on his heel.
D19	S13	Nawal is sent to Kfar Ryat prison.	I7 Abu Tarek has a tattoo on his heel.
D22	S14	Nawal becomes <i>the woman who sings</i> .	I8 Nihad is sent to an orphanage.
D25	S15	Nawal is raped by Abu Tarek.	I9 National Party bombs the orphanage.
D26	S16	Nawal gives birth to Jeanne and Simon.	I10 National(N.) Party kills refugees.
D31	I17	Nawal meets Chamseddine after release.	I11 Refugee camp becomes a ruin.
D34	S18	Nawal meets Nihad at the pool.	I12 Nihad is sent to the refugee camp.
D2	S22	Jeanne and Simon hears Nawal's will.	I13 Nihad became a soldier of Islamic resistance.
D21	S28	Jeanne meets former guard of Kfar Pyat.	I14 Nihad is captured by N.Party.
D27	S30	Jeanne and Simon meets former nurse of Kfar Ryat.	I15 Nihad became an executioner for N.Party.
D29	S31	Simon searches for Nihad.	I16 Executioner for N.Party is Abu Tarek.
D32	S33	Simon meets Chamseddine.	I17 Abu Tarek exiled to Canada.
D33	S34	Janne and Simon get overwhelmed.	
D36	S36	Nihad reads letters given by Jeanne and Simon.	

(Unnecessary events are omitted.)

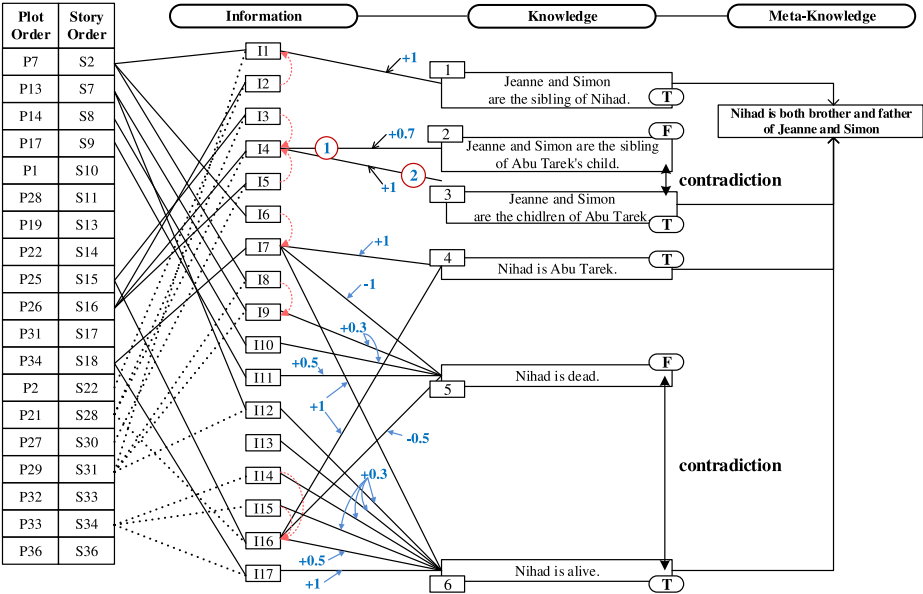


Fig. 4. Knowledge structure for *Incendies*. Event-information connection is show for Nawal (solid line) and Jeanne Simon (dotted line).

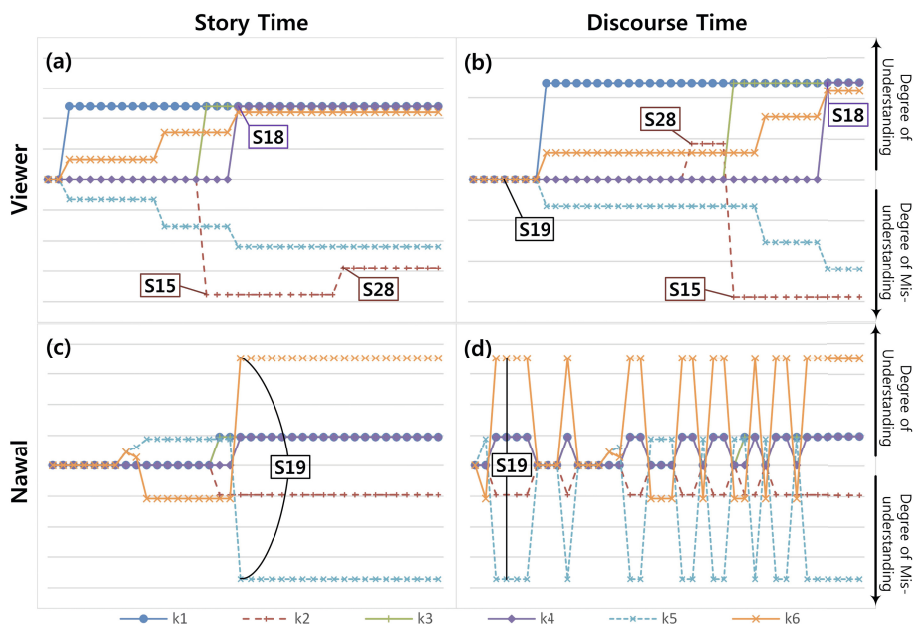


Fig. 5. Knowledge flow for viewer and *Nawal* with ordered event list as x-axis and degree of belief as y-axis, in story-time and in discourse-time

Fig. 5(c) depicts the knowledge flow of *Nawal* in the story-time. It captures the impact she experienced when she met *Nihad* in the pool (event 18), and realized that he is alive (K5, K6) as *Abu Tarek* (K4). Fig. 5(d) shows how the belief of *Nawal* is shifted throughout the film's discourse-time. *Nawal*'s knowledge of *Nihad*'s fate fluctuates over the graph, suggesting the existence of multiple flashbacks. Different knowledge flows between the viewers and *Nawal* explain why the viewers cannot understand *Nawal*'s shocking expressions at the pool (event 19), making the event more suspenseful. This disparity of knowledge is resolved in the later scene (event 18) when the viewers realize what previously happened at the pool, enhancing the viewer's involvement and sympathetic response.

As described so far, the proposed knowledge distribution model is shown to accurately express the complex knowledge flows and the corresponding emotional responses of agents in the actual film narrative. It is also shown that the knowledge flows of agents can interactively be altered by manipulating the knowledge structure in a simple way. Only the order of events is changed in this study, but more elaborative knowledge flow control can be done by manipulating relevant perception values and impact values.

6 Conclusion

We proposed a knowledge distribution model that quantifies information/knowledge in a narrative and simulated the knowledge changes according to the development of events. We designed the model to practically support an author to create narratives. In this modeling framework, we first constructed the *knowledge structure* by connecting event, information, knowledge, and meta-knowledge, then calculated the *knowledge state* and plotted the *knowledge flow* using a probabilistic reasoning model. We also conducted a concrete case study with an actual cinematic narrative to show that our model clearly expresses the complicated aspects related to the knowledge flow.

The model we proposed in this paper has three advantages as follows.

First, with this model, it is possible to quantitatively analyze a narrative. The knowledge state of agents is formalized as the degree of belief, thus, it can be calculated by a proper computational method such as probabilistic reasoning model. A graph that expresses the knowledge flow quantitatively shows when and to what extent an agent's knowledge state changes. Second, this model allows better control of the knowledge flow by the author in an interactive manner. During the creating process, an author frequently revises the configuration of events and the attributes of information within the knowledge structure. Events and knowledge within the knowledge structure can frequently be added, altered, rearranged or deleted to modify knowledge flows, helping the author to more effectively create narratives. Third, our model can be extended by combining it with other narrative system. The model we proposed is flexible in that it can be combined with the AI planning method or the suspense model mentioned above. It is also possible to be used as an analytic tool, analyzing and categorizing famous narratives according to the pattern of knowledge flow. This work can contribute to narrative studies on theme and genre.

Our current model has several limitations. We are planning to address these limitations in future studies.

The first limitation is ignoring the fact that characters should act differently according to their roles and personalities. Because we assume the rational agent in our model, the difference among characters cannot be reflected in the reasoning process. It also cannot reflect how bias or errors influence the reader's reasoning. Secondly, for practical purpose, automation is needed to relieve author's cognitive load in building a knowledge structure. Two functions can be considered: automatic extraction of important information from events and link generation between information and knowledge. NLP studies have to be combined with a knowledge flow model for those functionalities. Finally, the studies on the internal relationships of information are needed. Based on the prior information, the meaning of newly obtained information can be altered. In many cases, the meaning of information is clearly revealed only if multiple pieces of supporting information are assembled. The variation of meaning using the internal relationships of information makes a pattern of knowledge flow found in many narratives. It should be investigated by analyzing various narratives.

References

1. Oatley, K., Green, M.C., Strange, J.J., Brock, T.C.: Narrative impact: Social and cognitive foundations. *Narrative Impact: Social and Cognitive Foundations* (2002)
2. Baxter, L.A., Braithwaite, D.O.: Engaging theories in interpersonal communication: Multiple perspectives. SAGE Publications, Incorporated (2008)
3. Howard, D.: *How to Build a Great Screenplay: A Master Class in Storytelling for Film*, 1st edn. St. Martin's Press, New York (2004)
4. Riedl, M.O., Young, R.M.: Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research* 39(1), 217–268 (2010)
5. Peinado, F., Cavazza, M., Pizzi, D.: Revisiting character-based affective storytelling under a narrative BDI framework. In: Spierling, U., Szilas, N. (eds.) *ICIDS 2008*. LNCS, vol. 5334, pp. 83–88. Springer, Heidelberg (2008)
6. Graesser, A.C., Singer, M., Trabasso, T.: Constructing inferences during narrative text comprehension. *Psychological Review* 101(3), 371 (1994)
7. Graesser, A.C., Millis, K.K., Zwaan, R.A.: Discourse comprehension. *Annual Review of Psychology* 48(1), 163–189 (1997)
8. Niehaus, J., Young, R.M.: A method for generating narrative discourse to prompt inferences. In: *Proceedings of the Intelligent Narrative Technologies III Workshop*, p. 7. ACM (2010)
9. Brewer, W.F., Lichtenstein, E.H.: Stories are to entertain: A structural-affect theory of stories. *Journal of Pragmatics* 6(5), 473–486 (1982)
10. Cheong, Y.-G., Young, R.M.: Narrative generation for suspense: Modeling and evaluation. In: Spierling, U., Szilas, N. (eds.) *ICIDS 2008*. LNCS, vol. 5334, pp. 144–155. Springer, Heidelberg (2008)
11. Bae, B.-C., Young, R.M.: Suspense? Surprise! or how to generate stories with surprise endings by exploiting the disparity of knowledge between a story's reader and its characters. In: Iurgel, I.A., Zagalo, N., Petta, P. (eds.) *ICIDS 2009*. LNCS, vol. 5915, pp. 304–307. Springer, Heidelberg (2009)
12. Ryan, M.L., Herman, D., Jahn, M.: *Routledge Encyclopedia of Narrative Theory*. Routledge (2005)
13. Rowley, J.: The wisdom hierarchy: representations of the dikw hierarchy. *Journal of Information Science* 33(2), 163–180 (2007)
14. Abell, P.: Narratives, bayesian narratives and narrative actions. *Sociologica* 1(3) (2007)
15. Good, I.J.: *Good thinking: The foundations of probability and its applications*. Univ. of Minnesota Press (1983)