In the Working Notes of the 2012 Workshop on Computational Models of Narrative at the International Conference of Language Resources and Evaluation, pages 34-43, 2012. (Best Student Paper on a Cognitive Science Topic)

Indexter: A Computational Model of the Event-Indexing Situation Model for Characterizing Narratives

Rogelio E. Cardona-Rivera, Bradley A. Cassell, Stephen G. Ware, R. Michael Young

Liquid Narrative Research Group
North Carolina State University
{recardon, bacassel, sgware}@ncsu.edu, young@csc.ncsu.edu

Abstract

Previous approaches to computational models of narrative have successfully considered the internal coherence of the narrative's structure. However, narratives are also externally focused and authors often design their stories to affect users in specific ways. In order to better characterize the audience in the process of modeling narrative, we introduce *Indexter*: a computational model of the Event-Indexing Situation Model, a cognitive framework which predicts the salience of previously experienced events in memory based on the current event the audience is experiencing. We approach computational models of narrative from a foundational perspective, and feel that salience is at the core of comprehension. If a particular narrative phenomenon can be expressed in terms of salience in a person's memory, the phenomenon, in principle, is representable in our model. This paper provides the fundamental bases of our approach as a springboard for future work which will use this model to reason about the audience's mental state, and to generate narrative fabula and discourse intended to achieve a specific narrative effect.

Keywords: Narrative understanding and generation, representations, retrieval and indexing, artificial intelligence, cognitive psychology

1. Introduction

Historically, computational models of narrative have focused on representation of the diverse structural properties of narratives (Lebowitz, 1985; Cavazza et al., 2001; Riedl et al., 2003; Szilas, 2003). These models consider only the internal properties of the narrative. Authors, however, intentionally design stories to affect their audience in specific ways (Bordwell, 1989; Holland, 1989). As Szilas (2010) has suggested, a computational model of narrative must go beyond simple story structure and account for how the experiencer receives the narrative.

In this paper, we provide initial steps toward a computational model that accounts for a user's comprehension process during the experience of a narrative. This model, which we call *Indexter*, explicitly reasons about the salience of narrative events in a person's memory as they experience an unfolding story. The salience of a narrative event indicates how recallable the event is in a person's mind. An author's manipulation of the salience of events during a narrative experience is a key means used to affect a reader's comprehension of the story's structure. Salience enables the drawing of connections between new material and earlier parts of the story. Salience prompts expectations about upcoming action. Lack of salience obscures predictions and facilitates surprise or misdirection. A model of narrative that accounts for salience could be linked to existing models that build off of salience to account for a reader's inference-making process (Niehaus and Young, 2010), her feelings of suspense (Cheong and Young, 2006), and her level of surprise (Bae and Young, 2009), along with many other narrative phenomena.

Though our current model focuses on the manipulation of salience in narrative, salience alone is not sufficient for the modeling or creation of effective stories. A story's internal structure clearly plays a role in how a reader understands it (Graesser et al., 2002). Thus, the computational model that we present extends an existing planning-based approach to narrative (Young, 2007), which models coherent story structure (Riedl and Young, 2010). We augment this plan-based approach with information that allows us to model the updates being made to a reader's mental model of the story during *online comprehension*, that is, during the process of experiencing the narrative. To do this, we incorporate elements into the planning model drawn from an empirically verified cognitive model of online comprehension called the *Event-Indexing Situation Model* (Zwaan et al., 1995a; Zwaan and Radvansky, 1998). While we are basing our work on a planning-based knowledge representation previously developed to generate stories, our discussion here does not describe a system that uses this representation in a generative fashion. The work we describe here is preliminary. It is the first step of a four-part research agenda involving:

1. Development of a plan-based knowledge representation for narratives and an algorithm that characterizes the reader's construction of event-indexing situation models.

2. Validation of the predictive power of the algorithm and representation.

3. Integration of the computational model into a generative system.

4. Validation of the generative system in an online comprehension scenario.

A generative system which uses a computational model that characterizes both the internal structure of a narrative and its effects on a reader during online comprehension will lead to the creation of more engaging, effective and understandable stories.

2. Theoretical Bases of our Implementation

Our work has two fundamental bases. The first basis is a cognitive model of online story comprehension, the Event-Indexing Situation Model (Zwaan et al., 1995a; Zwaan and Radvansky, 1998). The second basis is an AI plan-based model of narrative, which follows directly from IPOCL (intentional, partial order, causal link) plans (Riedl and Young, 2010).

2.1. The Event-Indexing Situation Model

The Event-Indexing Situation Model (EISM) is a cognitive model of online narrative comprehension. Cognitive psychologists studying narrative comprehension define a *situation model* as an integrated mental representation of a particular situation in the story world. Situation models are formed by a reader from an amalgamation of information explicitly stated in a narrative and inferred by the reader (see McNamara and Magliano (2009) for a review of several variants of situation models). In particular, the EISM posits that, as we perceive a narrative, we discretize the narrative into *events*, or chunks of narratively important action (Zwaan et al., 1995a). This event segmentation centers around verb phrases in text and character actions in film (Zacks et al., 2009). Each event is indexed by the reader relative to a number of key factors or dimensions including:

- *time index* - the time frame in which the event occurs
- *space index* - the space in which the event takes place
- *protagonist index* - whether or not the event involves the protagonist
- *causal index* - the event's causal status with regards to previous events
- *intention index* - the event's relatedness to the intentions of a character

The EISM makes predictions about the salience of events based on these indices.

2.1.1. The EISM and Memory

Zwaan and Radvansky (1998) discuss the interplay between the EISM and memory in the context of Ericsson and Kintsch's (1995) conceptualization of Short-Term Working Memory (STWM) and Long-Term Working Memory (LTWM). Zwaan and Radvansky point out that,

> It is possible in highly practiced and skilled activities, such as language comprehension, to extend the fixed capacity of the general short-term working memory (STWM) system by efficiently storing information in long-term memory and keeping this information accessible for further processing. This expansion of STWM is called long-term working memory (LTWM) and corresponds to the accessible parts of a previously constructed mental representation in long-term memory.
>
> (Zwaan and Radvansky, 1998)

The STWM is represented in the EISM by a structure known as the *current situation model*.

Definition 1 (Current Situation Model) *The* current situation model *refers to the model of the event that is currently being perceived. This is the model at time t_n, for a given event e_n.*

When an event is perceived, a situation model of that event is built to identify what it's situation model indices are with respect to all previously perceived events. All previously perceived events represent Ericsson and Kintsch's idea of LTWM, which is represented in the EISM by a structure known as the *integrated situation model*.

Definition 2 (Integrated Situation Model) *The* integrated situation model *refers to the model of the events that have been perceived up until right before the event currently being perceived. This is the model for times t_1 through t_{n-1}, for events e_1 through e_{n-1}.*

The STWM maintains retrieval cues to information in LTWM to help with information storage and retrieval. The metaphor of a hash map is useful here: The STWM can be thought of as a set of keys to the values that are held in LTWM. For the EISM, the keys are all the unique situation model indices that exist in the development of a story. Each value in this memory hash map is a list of events that share a particular situation model index of a story. Online comprehension in the EISM is modeled as follows: each incoming event is analyzed (by the audience[1]) to determine which situation model indices it contains. The audience tries to match the incoming event to the most recently *foregrounded* events, or the events that are currently most salient.

The matching between events is done by verifying if there is any overlap between the incoming event's situation model indices and the most recently foregrounded events' indices. If the incoming event does not share any indices with the most recently foregrounded events, then a lookup is done to the memory hash map. If the lookup is successful (meaning that the situation model indices have been encountered before), the corresponding values (the previous events) become foregrounded. The incoming event is then inserted in the memory hash map and associated to the events that have been foregrounded. If the lookup is unsuccessful (meaning that we have encountered a completely novel situation), a new key is created with the new indices, and the key is mapped to the current event in the memory hash map.

2.1.2. Example Interaction Between the EISM and Memory

Consider a story which is perceived by the audience as a sequence of events $e = \langle e_1, e_2, ..., e_{10} \rangle$. In this story, only events e_1 and e_{10} have the same causal index (i.e.,

[1]In this paper, we refer to an individual experiencing a narrative as *the audience*. This term is intended to make no commitment to the medium through which the narrative is experienced, in contrast to terms like *reader*, *viewer* or *player*, which might imply restriction to a specific storytelling context.

they form part of the same causal chain). Recall that, as each individual event is perceived, a current situation model is created for it and it is subsequently integrated with the integrated situation model before the next event is perceived. According to the EISM, when event e_{10} is perceived, it acts like a retrieval cue to event e_1 due to their common causal index. Thus, the EISM will predict that, after having perceived event e_{10}, event e_1 will be more salient in memory than events e_2 through e_9.

The EISM does not make a commitment to determining which indices prove to be stronger predictors of recall. The strength of recall is operationalized in the cognitive psychology literature through various means including word association tasks (Zwaan et al., 1995a), question-answering tasks (Graesser and Franklin, 1990), timed reading tasks (Zwaan and Radvansky, 1998), and narrative summarization tasks (Graesser and Clark, 1985).

2.1.3. Building the indices in the EISM

The EISM establishes a set a criteria for assigning a situation model index to an event, with one criterion for each situational dimension: time, space, protagonist, causation and intention. Since the EISM makes predictions on salience relative to how many indices are shared between events, the criteria for indices is best expressed in terms of when events share an index.[2] The criteria for assigning situation model indices is succinctly described by Zwaan et al. (1995a) and we paraphrase and expand upon it here:

- Two events share a *time* index if they occur in the same time frame. This time frame is identifiable using the criteria employed by Zwaan (1996): two events are assumed to share a time index if they are perceived by the audience in sequential order and neither event contains an explicit discontinuity in time.
- Two events share a *space* index if they occur in the same spatial region.
- Two events share a *protagonist* index if they both involve the story's protagonist. The protagonist index is special in that it contributes to an event's salience, regardless of whether the event has been foregrounded or not. The authors of the EISM distinguish a single character as the protagonist of a story, and the model predicts that any event that deals with the protagonist is more likely to be salient than events that do not deal with the protagonist.
- Two events share a *causation* index if they are related causally. A *direct causal relation* is directed, from one event to another. A direct causal relation from event e_1 to e_2 exists, as specified by Trabasso and Sperry (1985), if it meets the logical criteria of necessity and if the events pass a counterfactual test of the form: if event e_1 had not occurred, then in the context of the story, event e_2 would not have occurred. An *indirect causal relation* between two events e_i to e_n exists if there is a path in the transitive closure of the causal relation from e_i to e_n. Trabasso and Sperry (1985) reference four types of causal relations that can exist between events:
  - *Enablement* is a causal relation that involves events which are necessary but not sufficient to cause other events.
  - *Motivation* and *Psychological Causation* are causal relations that are similar in that they both purposefully effect a change in the world, with the difference that *motivation* is goal-directed whereas *psychological causation* is not.
  - *Physical Causation* involves a naive interpretation of the physical world or of mechanical causality between objects and/or people.
- Two events share an *intention* index if they are part of the same plan to achieve a goal. Goal structures are derived from General Knowledge Structures as identified by Graesser and Clark (1985).

The EISM situational indices are coded dichotomously; that is, two events can either share an index, or not. Zwaan notes that the model may be extended in future work.

2.2. The IPOCL Planning Model

Intentional Partial Order Causal Link (or IPOCL) plans are a data structure for representing stories that explicitly model the events of a story along with the casual, temporal, and intentional relationships between them (Young, 1999; Riedl and Young, 2010). Here we introduce the IPOCL Planning Model, and give a brief formal description of what an IPOCL plan looks like.

A plan is a sequence of steps that describes how a world transitions from its beginning, or initial state, to its end, or goal state (Newell and Simon, 1961). In narrative terms, it describes how the plot of a story causes the story world to transition from beginning to end.

Definition 3 (State) *A* state *is a single function-free ground predicate literal or a conjunction of literals describing what is true and false in a story world. The* initial state *completely describes the world before the start of a plan. The* goal state *is a conjunction of literals which must be true at the end.*

Definition 4 (Planning Problem) *The initial and goal states together make up the* planning problem *to which a particular IPOCL plan is the solution.*

Characters, items, and places in the story are represented as logical constants. The actions which materialize between the initial and goal states make up the plan. Actions are created from templates.

Definition 5 (Operator) *An* operator *is a template for an action which can occur in the world. It is a three-tuple $\langle P, E, A\rangle$ where P is a set of* preconditions*, literals which must be true before the action can be executed, E is a set of* effects*, literals which are made true by the execution*

[2] A situation model index is a property of the event, independent of other events. In other words, each event has an individual time, space, protagonist, causation and intention index.

of the action (Fikes and Nilsson, 1971)*, and A is a set of characters which must consent to the execution of that action* (Riedl and Young, 2010). *For generality, P, E, and A can have variable terms to convey ideas such as "creature x steals item y." An operator for which $A = \emptyset$ is called a* happening*; these actions represent accidents or the forces of nature which are not intended by anyone.*

Definition 6 (Planning Domain) *The set of all available operators is called the* planning domain*. A domain describes all the possible kinds of actions that can occur.*

An instance of an operator, called a step, represents an actual action that will take place in a story.

Definition 7 (Step) *A* step *is a three-tuple $\langle P, E, A\rangle$, where P, E, and A are the preconditions, effects, and consenting characters from the step's operator. If any literals in P or E contain variables, or if any symbols in A are variables, those variables must each be bound to a single constant. The step "the dragon steals the treasure" is an instance of the operator "creature x steals item y."*

Plan steps are partially ordered with respect to time (Sacerdoti, 1975).

Definition 8 (Ordering) *An* ordering *over two steps is denoted $s < u$, where s and u are steps in the plan and s must be executed before u.*

A plan must guarantee that, for each step, all of the step's preconditions are true before it is executed (McAllester and Rosenblitt, 1991). A precondition can be true in the initial state or made true by the effect of an earlier step.

Definition 9 (Causal Link) *A* causal link *is denoted $s \xrightarrow{p} u$, where s is a step with some effect p and u is a step with some precondition p. A causal link $s \xrightarrow{p} u$ implies the ordering $s < u$. A casual link explains how a precondition of a step is met. In other words, p is true for u because s made it so. Step u's* causal parents *are all steps s such that there exists a causal link $s \xrightarrow{p} u$. A step's* causal ancestors *are its causal parents in the transitive closure of the parent relation.*

IPOCL plans contain structures called frames of commitment to explain a character's actions in terms of individual goals (Riedl and Young, 2010).

Definition 10 (Intention) *An* intention *is a modal predicate of the form $intends(a, g_a)$ where a is an actor and g_a is a literal that actor a wishes to be true. A* motivating step *is a step which causes an actor to adopt a goal. It has as one of its effects an intention—a modal predicate of the form $intends(a, g_a)$. A* final step *is a step which achieves some actor goal. It must have g_a as one of its effects.*

The steps which materialize between a motivating and final step make up a frame of commitment.

Definition 11 (Frame of Commitment) *A* frame of commitment *is a five-tuple $\langle S', P, a, g_a, s_f\rangle$ where S' is a subset of steps in some plan P, a is a character, g_a is some goal of character a, and s_f is a final step which has effect g_a. The steps in S' are all the steps which character a takes in order to achieve goal g_a. All steps in S' must be causal ancestors of s_f, and all steps in S' must be ordered before s_f.*

Simply put, a frame of commitment describes the steps an actor takes to achieve some goal, and the step which finally achieves the goal.

The artifact produced by a planner is a plan:

Definition 12 (Plan) *A* plan *is a five-tuple $\langle S, B, O, L, I\rangle$ where S is a set of steps, B a set of variable bindings, O a set of orderings, L a set of causal links, and I a set of frames of commitment. A* complete plan *is guaranteed to achieve the goal from the initial state. A plan is complete if and only if:*

- *For every precondition p of every step $u \in S$, there exists a causal link $s \xrightarrow{p} u \in L$. This means that every precondition of every step is satisfied.*
- *For every step $s = \langle P, E, A\rangle \in S$, and for every character $c \in A$, there exists a frame of commitment $i = \langle S', P, a, g_a, s_f\rangle$ such that $s \in S'$ and $c = a$. This means that every step which is not a happening is a member of some frame of commitment that explains why the characters who carry out that step choose to carry it out. In short, every action is taken for a reason.*
- *For every causal link $s \xrightarrow{p} u \in L$, there is no step $t \in S$ which has effect $\neg p$ such that $s < t < u$ is a valid ordering according to the constraints in O. In other words, it is not possible that a causal link gets undone before it is needed.*

IPOCL plans are formal data structures which can be manipulated by planning algorithms (Riedl and Young, 2010). They model important information about stories which, we claim in this paper, can be modified to operationalize the EISM to predict how well humans remember certain steps.

3. Indexter: a model that characterizes Situation Models using Plan Structures

Indexter is realized by expanding the IPOCL plan representation with information regarding EISM relevant data. As defined, the IPOCL plan representation already captures many of the features needed to represent EISM structures, and the enhancements we outline are straightforward to introduce. By extending an existing knowledge representation used to characterize the structural properties of a narrative, Indexter can characterize both proper narrative structure and the online mental state of the audience which experiences the narrative. This characterization is a foundational approach to a computational model of narrative; our model does not currently characterize specific narrative phenomena such as tension, suspense, expectations, humor, character development, etc. Instead, our model is intended to provide a foundation for prompting the experience of these narrative phenomena thorough the facilitation of the manipulation of salience.

Events in the EISM framework map to steps in an IPOCL plan, assuming that the steps center around verbs (in text) or actions (in film). In the remaining discussion, we use the EISM term *event* and the IPOCL term *step* interchangeably. Recall that the IPOCL model represents elements of the *fabula*, the set of events, characters, locations and the other entities within the story world. The EISM deals with online narrative comprehension which occurs when the audience perceives the narrative's *discourse*, or the telling of the events in a story.

To extend IPOCL with the EISM knowledge representation, we leverage IPOCL for use as a *discourse plan*, a structure which contains the elements from the story that will be included in the story's telling to the audience (Young, 2007). In general, discourse plans do not have to preserve the ordering of events as they occur in the *fabula* (for instance, as in cases of foreshadowing or flashback (Bae and Young, 2009)), nor do they have to contain all the events that occur in the story (e.g., as in the case of temporal ellipsis). However, in the discussion here, we use a fairly straightforward set of extensions to the IPOCL plan representation already used to model *fabula* in order to characterize just those elements of the narrative discourse relevant to the current work. The way we represent all the EISM indices is explained in the following sections.

3.1. Time

Time is implicitly represented in the current IPOCL model. Steps are modeled as executing instantaneously and IPOCL's temporal representation provides a partial ordering over all of a plan's steps' times of occurrence. Rather than extending this base representation with more complex models of time as has been done with temporal planning approaches (e.g., that of Penberthy and Weld (1994)), we approximate this by requiring that each operator in an IPOCL planning domain contain a distinguished variable called the *time frame*. In any IPOCL plan, each step's time frame variable must be bound to one of a list of constants that refer to time frames in the given narrative. These constants, enumerated as part of a planning problem's initial state description, can be defined either by the domain creator or automatically, for instance, by a temporal cluster analysis of the steps in the plan.

3.2. Space

In the current IPOCL model, spatial properties of steps are represented only to the extent that the writer of an operator includes spatial relations in its preconditions and effects. To model where an event occurs, we require that each operator in an IPOCL plan domain contain a distinguished variable called the *location*. In any IPOCL plan, each step's location variable must be bound to one of a list of constants that refer to locations in the given narrative. These constants, enumerated in a planning problem's initial state description, can be defined by the domain creator or automatically, for instance, by inferring a step's location from the bindings of variables that appear in the step's preconditions or effects.

3.3. Protagonist

In the EISM model, the protagonist is single designated character that fulfills the role for an entire story. To model the protagonist, we require that an IPOCL initial state description contain an entry designating a single character as the protagonist for the given story plan. For any given step in a plan, the protagonist index captures whether or not events involve a story's protagonist. The IPOCL plan representation already contains elements that can be used to characterize the protagonist index of each step. Each step designates a set A of consenting characters. A contains zero or more variables that are bound to the agents that consent to the execution of that step. By comparing the designated protagonist for a story with the members of A, we can easily check whether or not the protagonist is involved in the execution of the step.

3.4. Causation

The causation index captures whether or not events have a causal relation. The IPOCL plan representation contains several elements that can be used to represent the causation index. Recall that Trabasso and Sperry (1985) used four types of causal relations in their analysis of causal structure. One of these causal relations (physical causation) is currently not represented in our initial work. The other three causal relation types are represented as follows:

- *Enablement*

  An IPOCL causal link represents an enablement causal relation. In an IPOCL plan, a causal link's originating step s_i is a necessary but not sufficient condition for the subsequent step s_j through $s_i \xrightarrow{p} s_j$, because it is possible for any other step s_k to establish p for s_j, creating a causal link $s_k \xrightarrow{p} s_j$.

- *Motivation* and *Psychological Causation*

  A causal link $s_i \xrightarrow{p} s_j$, where s_i is an IPOCL motivating step represents a motivation causal relation. Recall that a motivating step is a step that causes an actor to adopt a goal. For all the steps taken by that actor in service to that goal, two types of psychological causal relations could occur:

  1. All effects of steps taken in service of the goal that do not establish a causal link to any other step that is executed by the actor are said to be psychologically caused by the action that establishes the effects.
  2. All effects that are made true by the final step (that achieves the effects that define the goal) which are not part of the goal condition are said to be psychologically caused by the final step.

  Psychological causal relations are contextualized by a specific goal state and an actor that intends to achieve it. Informally, they can be thought of as unwanted/unplanned side-effects of actions taken in service of a goal.

3.5. Intention

The intention index characterizes the role that an event plays in a character's plan to achieve a single goal. The IPOCL plan representation currently contains elements that can represent the intention index through the IPOCL frames of commitment. A frame of commitment describes the steps an actor takes to achieve a specific, designated goal condition. We say that two steps share an intention index just when they are part of the same IPOCL frame of commitment.

4. Indexter in action: using the model

Indexter allows for calculating the saliency of any previously experienced event with respect to the current event being perceived. Being able to estimate salience at any point by using our model will allow an AI planner to generate a narrative which can directly operate on the salience of events in the audience's mind. This generative system would be able to actively track and manipulate the audience's mental model to achieve specific narrative phenomena that arise, in part, from the dynamics of salience.

4.1. Calculating Salience

Once an IPOCL plan has been augmented to keep track of EISM information, one way to compute salience is by taking a majority-vote of the indices that are referenced by the event that is currently being perceived. Recall that indices are dichotomously tracked for events. If two events share an index, the value for that index is 1; otherwise, it is 0. Assume that salience can be represented with a real-numbered value between 0 and 1, where 0 represents no salience whatsoever and 1 represents maximum salience. Salience is calculated from the parameter event e_i with respect to the current event being perceived e_n. Each EISM index is assigned a weight coefficient such that the total salience will be between 0 and 1. Under these constraints, an equation to calculate the salience of any event e_i is:

$$\begin{aligned} salience(e_i, e_n) &= w_1 t_{e_n} + w_2 s_{e_n} + w_3 p_{e_n} \\ &\quad + w_4 c_{e_n} + w_5 i_{e_n} \end{aligned} \tag{1}$$

Where t_{e_n} is the time index, s_{e_n} is the space index, p_{e_n} is the protagonist index, c_{e_n} is the causality index, i_{e_n} is the intentionality index for the *event that is currently being perceived* e_n. Each index represents the overlap on that index between any event e_i and the current event e_n. For any situation model index of the current event x_{e_n}, $x_{e_n} = 1$ just when event e_i shares the x index with the current event e_n and $x_{e_n} = 0$ otherwise.

The coefficient w_j represents the contribution (weight) of its respective index to the saliency of the parameter event e_i. The coefficients are restricted to sum to 1, that is, $\sum_{j=1}^{n=5} w_j = 1$.

Clearly, assigning specific weights to the various indices will affect the salience of events in a significant way. Unfortunately, the authors of the EISM do not specify which indices are stronger predictors of recall. In future work, we will seek to determine the values for these indices empirically, through experimental evaluation. For the current discussion, however, a straightforward way to weigh the indices is to assign each an equal value. For five indices, this implies $w_j = 0.2$. Equation 1 then becomes:

$$\begin{aligned} salience(e_i, e_n) &= 0.2 t_{e_n} + 0.2 s_{e_n} + 0.2 p_{e_n} \\ &\quad + 0.2 c_{e_n} + 0.2 i_{e_n} \end{aligned} \tag{2}$$

We use equation 2 to calculate the salience of events in the following example.

Example salience calculation: The Knight's Quest

Consider the following story, in which each event is tagged with an event marker that illustrates the order in which the audience (in this case, the reader) perceives the events:

> A dragon flies to a castle (e_1), steals the treasure in the castle (e_2), and flies off to a cave (e_3). A couple of hours later the knight smiths a sword at the castle (e_4) to prepare for his quest. The following day, the knight travels to the cave (e_5), slays the dragon (e_6), reclaims the treasure (e_7), and returns to the castle (e_8).

In this story, we designate the `Knight` as the protagonist. This story is a totally ordered, text-based realization of a plan that IPOCL can produce (Riedl and Young, 2010). This plan is illustrated in Figure 1, augmented with our extended knowledge representation elements. Time, space, and protagonist indices are indicated in the steps. The steps are grouped into their specific frames of commitment, and are connected by causal links shown as arrows.

To calculate the salience of a given step, we use Equation 2. For example, to calculate the salience of step e_2 = `(steal Dragon Treasure)` at the step e_6 = `(slay Knight Dragon)`, we determine how many indices overlap between event e_2 and event e_6. Event e_2 is not connected in space, time, or protagonist to event e_6. The events are also not in the same frame of commitment. They are, however, connected causally. Using Equation 2:

$$\begin{aligned} salience(e_2, e_6) &= (0.2 \cdot t_{e_6}) + (0.2 \cdot s_{e_6}) + (0.2 \cdot p_{e_6}) \\ &\quad + (0.2 \cdot c_{e_6}) + (0.2 \cdot i_{e_6}) \\ salience(e_2, e_6) &= (0.2 \cdot 0) + (0.2 \cdot 0) + (0.2 \cdot 0) \\ &\quad + (0.2 \cdot 1) + (0.2 \cdot 0) \\ salience(e_2, e_6) &= 0.2 \end{aligned}$$

The salience of all steps relative to e_6 can be calculated in the same manner. For comparison, consider the step e_5 = `(walk Knight Cave)`. Event e_5 shares the time, space, protagonist, causation and intention indices with event e_6, such that the salience of event e_5 at event e_6 is:

$$\begin{aligned} salience(e_5, e_6) &= (0.2 \cdot 1) + (0.2 \cdot 1) + (0.2 \cdot 1) \\ &\quad + (0.2 \cdot 1) + (0.2 \cdot 1) \\ salience(e_5, e_6) &= 1 \end{aligned}$$

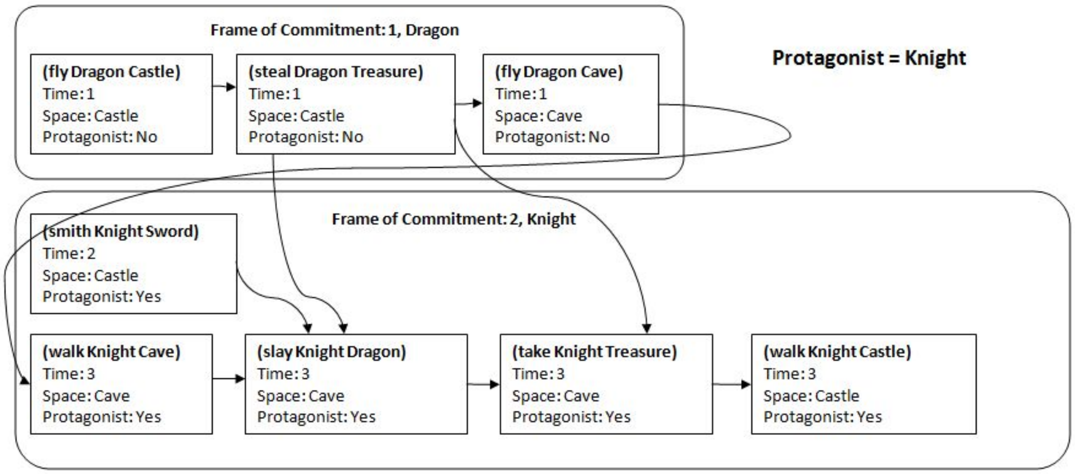

Figure 1: Example IPOCL Plan for *The Knight's Quest*, with extensions indicating aspects of our knowledge representation.

Thus, our computational model would predict that the `(walk Knight Cave)` step is more salient in the reader's mind than the step `(steal Dragon Treasure)` when the reader is reading the narrative at step `(slay Knight Dragon)`.

5. Potential Applications

The computational model we have presented here could be useful as a generative model for both *fabula* and discourse planning. We outline below some of the applications that a generative model would enable. Generally, if a particular narrative phenomenon can be expressed in terms of salience in a person's memory, the phenomenon, in principle, is representable in our model.

5.1. Fabula Planning

Using EISM information, an AI planner could construct the story in a way that it produces a plan according to the expected salience of events, which (if executed) could optimize a person's feeling of suspense (Cheong and Young, 2006) or surprise (Bae and Young, 2009). Alternatively, we could use an AI planner to dynamically construct a story that manipulates salience online, in order to actively affect a reader's inference-making process (Niehaus and Young, 2010). The latter manipulation could be used in educational contexts, with intelligent tutoring systems (Thomas and Young, 2011), as well as entertainment contexts, to affect which future events are narratively afforded (Young and Cardona-Rivera, 2011) by the events perceived in the story at a given point.

5.2. Cinematic Discourse Planning

Most work in cinematic generation reasons about low-level frame-by-frame placement of a camera in a virtual scene (Bares et al., 2000; Christianson et al., 1996). However, in film, cinematographers either explicitly or implicitly frame shot sequences to manipulate the mental state of the viewer (Branigan, 1992). Initial work on a system which reasons about high level narrative goals has been done by Jhala and Young (2010). However, this work does not make extensive use of a knowledge representation to characterize a cinematic's effect on the mental state of a viewer. EISM has been shown to accurately model cognition for film (Magliano et al., 2001). One way for cinematic generators to achieve a higher level of communicative capability is for them to reason about the cognitive effects of shots on viewers. Our model could provide this information. Given a plan step's saliency prediction, an planning system could construct a visual discourse specifically to bring certain information into focus.

Manipulation of focus would prove useful for many application that relate to visual discourse. For example, a character may deliberate on a course of action and decide to change his or her plan. In these cases a cinematographer may want to make events salient which help the viewer infer what a character is currently thinking before the character makes a drastic change in their intentions. Using our model, a discourse planner would be able to reason directly about the effects of shots on the viewer's mental state, allowing for the design of cinematic action from a narrative standpoint.

6. Limitations and Future Work

Our computational model was designed to follow the cognitive EISM very closely. Our intent is to increase the likelihood that our model will demonstrate a similar effectiveness at predicting salience. While the EISM is an empirically verified and very useful framework for characterizing the mental state of an audience during online comprehension, it does not track certain information which would be useful in a generative computational model of narrative. Also, there are details of our implementation which are subject to refinement. We identify some of the limitations from a computational perspective of the EISM and Indexter in the subsections that follow.

6.1. Limitations of the EISM

Previous work on situation models have focused on a single protagonist, and our model is restricted to one protagonist as well. However, stories often include multiple important characters beyond the protagonist (e.g. the antagonist) which may prove to be important indicators of salience. In these cases, it may be useful to extend our model to more than one character. Future work will involve determining the need for extending the protagonist index to account for multiple characters beyond the protagonist.

Space is a very complicated phenomenon. The EISM model of space is a simplification, in that it does not require or provide representations of spatial hierarchies (e.g., rooms within buildings), spaces within spaces, movable spaces (e.g., shipping containers), adjacent spaces, etc. We are interested in improving the representational capacity of our computational model to capture these types of potentially complex spatial relations in a narrative.

As we noted in Section 2.1.3., the current EISM treats situation model indices as binary (Zwaan et al., 1995a): two events are either connected by an index or not. In an effort to accurately depict the cognitive psychology research, our model treats indices in the same binary fashion. We are interested in relaxing this constraint to be able to represent events that are moderately (as opposed to directly) linked via situation models. We hypothesize that moderate situational relations will have a significant, but more gradual effect on saliency. We can think of two potential ways to approach this limitation:

1. *Introduce a distance penalty on the salience score for distant events*. This would be useful when considering events that happened (in the discourse) relatively early in comparison to the event that is currently being perceived. For example, an event further back on a causal chain should be slightly more difficult to recall than one which is closer to the step that is currently being perceived in the causal chain.

2. *Allow for a non-dichotomous indexing of events*. This would be useful when considering events that have a close relationship along one index, but that would be ignored because there is no strict overlap. For the space dimension, Zwaan and Radvansky (1998) have shown that the spatial representation in terms of distance between objects in an environment does affect response time of readers when probed with questions regarding the environment. For example, consider a spatial relationship of two adjacent rooms A and B. If the mental representation of the audience captures this adjacency, mentioning room A in the discourse will elicit transient memory saliency for room B.

6.2. Limitations of Indexter

Indexter calculates the salience between two events. Future work would extend the capabilities of our model to calculate saliency between an event and an object. Experiments have shown the abilities of people to recall objects or rooms, not necessarily a specific event alone (Zwaan et al., 1995b). However, it is not clear how the EISM handles this. Future work will determine how to expand Indexter to handle salience for elements other than events.

Indexter also depends on weights for each index to determine saliency. The weights we used are arbitrary and we allow these to be set manually. Knowing what coefficients accurately represent the predictive power of each index would increase the accuracy of salience calculations. Preliminary research has been done to determine the index weights (Zwaan et al., 1995a), and results suggest that these indices could be narrative or genre-dependent.

In Indexter, salience is calculated by computing a real-numbered value between 0 and 1 and is always calculated with respect to another event. Future work would determine at what point an event becomes salient and if it is dependent on any specific aspect of the story.

7. Conclusion

Narratives are an important part of the human experience, and they are used in diverse contexts well beyond entertainment. Psychologists (e.g., Bruner (1991)) suggest that narratives are key for explaining the ways that humans understand and reason about the world around them; these narrative psychologists posit that people perceive and interpret activities and behaviors by structuring them into a narrative. While many approaches to the development of a computational model of narrative have focused on the models' uses when generating stories, the foundational role that stories play in our cognition suggests that these models are significant also for the insight they provide to us about our own intelligence.

Previous approaches to computational models of narrative have been successful in capturing the diverse structural aspects of narrative. We propose that reasoning about the effects of narratives on their audience is the next step on the path of developing an artificial intelligence system capable of communicating narratives to humans. To this end, we have presented Indexter: a computational operationalization of the Event-Indexing Situation Model, drawn from the field of cognitive psychology. Our model extends previous work in computational models of narrative which uses AI planning constructs. Specifically, we have modified IPOCL plan structures to be capable of tracking situation model indices as a narrative is experienced. This paper presented the foundation of future work, which will leverage our model to predict the salience in memory of previously experienced events in a narrative, use that information to reason about the audience's mental state, and generate narrative *fabula* and discourse to achieve a specific narrative phenomenon.

8. Acknowledgements

This material is based upon work supported in part by U.S. DOE Computational Science Graduate Fellowship No. DE-FG02-97ER25308 and NSF Grants Nos. IIS-0915598 and IIS-0941421. The opinions, findings, and conclusions expressed here are those of the authors and do not necessarily reflect the views of the sponsoring agencies.

9. References

Byung-Chull Bae and R. Michael Young. 2009. Suspense? Surprise! or How to Generate Stories with Surprise Endings by Exploiting the Disparity of Knowledge between a Story's Reader and Its Characters. *Interactive Storytelling*, pages 304–307.

William Bares, Scott McDermott, Christina Boudreaux, and Somying Thainimit. 2000. Virtual 3D Camera Composition from Frame Constraints. *Proceedings of the Eighth ACM International Conference on Multimedia*, pages 177–186.

David Bordwell. 1989. *Making Meaning: Inference and Rhetoric in the Interpretation of Cinema*. Cambridge: Harvard University Press.

Edward Branigan. 1992. *Narrative Comprehension and Film*. Routledge.

Jerome Bruner. 1991. The Narrative Construction of Reality. *Critical Inquiry*, 18(1):1–21.

Marc Cavazza, Fred Charles, and Steven J. Mead. 2001. *Characters in search of an author: AI-based virtual storytelling*. Springer.

Yun-Gyung Cheong and R. Michael Young. 2006. A Computational Model of Narrative Generation for Suspense. *Working Notes of the AAAI Workshop on Computational Aesthetics*, pages 8–14.

D.B. Christianson, S.E. Anderson, L.W. He, D.H. Salesin, D.S. Weld, and M.F. Cohen. 1996. Declarative Camera Control for Automatic Cinematography. In *Proceedings of the National Conference on Artificial Intelligence*, pages 148–155.

David E. Copeland, Joseph P. Magliano, and Gabriel A. Radvansky. 2006. Situation Models in Comprehension, Memory, and Augmented Cognition. In Chris Forsythe, Michael L. Bernard, and Timothy E. Goldstein, editors, *Cognitive Systems: Human Cognitive Models in Systems Design*, chapter 2, pages 37–66. Lawrence Erlbaum Associates, Mahwah, New Jersey.

K. Anders Ericsson and Walter Kintsch. 1995. Long-term working memory. *Psychological Review*, 102:211–245.

Richard E. Fikes and Nils J. Nilsson. 1971. STRIPS: A new approach to the application of theorem proving to problem solving. *Artificial intelligence*, 2(3-4):189–208.

Arthur C. Graesser and Leslie F. Clark. 1985. *Structures and Procedures of Implicit Knowledge: Advances in Discourse Processes*. Ablex Publishing.

Arthur C. Graesser and Stanley P. Franklin. 1990. QUEST: A cognitive model of question answering. *Discourse Processes*, 13(3):279–303, July.

Arthur C. Graesser, Brent Olde, and Bianca Klettke. 2002. How Does the Mind Construct and Represent Stories? *Narrative Impact: Social and Cognitive Foundations*, pages 231–263.

David Herman. 2000. Narratology as a cognitive science. *Image & Narrative (Online Magazine of the Visual Narrative)*, 1(1).

Norman N. Holland. 1989. *The Dynamics of Literary Response*. New Work, NY: Columbia University Press.

Arnav Jhala and R Michael Young. 2010. Cinematic Visual Discourse: Representation, Generation, and Evaluation. *IEEE Transactions on Computational Intelligence and AI in Games*, 2(2):69–81, June.

Michael Lebowitz. 1985. Story-telling as planning and learning. *Poetics*, 14(6):483–502.

Joseph P. Magliano, Jason Miller, and Rolf A. Zwaan. 2001. Indexing space and time in film understanding. *Applied Cognitive Psychology*, 15(5):533–545.

David McAllester and David Rosenblitt. 1991. Systematic Nonlinear Planning. In *Proc. of AAAI*, pages 634–639. Massachusetts Institute of Technology Press.

Danielle S. McNamara and Joseph P. Magliano. 2009. Toward a Comprehensive Model of Comprehension. *Psychology of Learning and Motivation*, 51:297–384.

A. Newell and H. A. Simon. 1961. GPS, a program that simulates human thought. *Computers and thought*, pages 279–293.

James Niehaus and R. Michael Young. 2010. A method for generating narrative discourse to prompt inferences. In *Proceedings of the 3rd Workshop on Intelligent Narrative Technologies*, pages 1–8.

J. Scott Penberthy and Daniel S. Weld. 1994. Temporal Planning with Continuous Change. In *Proceedings of the National Conference on Artificial Intelligence*, pages 1010–1015.

Mark O. Riedl and R. Michael Young. 2010. Narrative planning: Balancing plot and character. *Journal of Artificial Intelligence Research*, pages 164–7.

Mark O. Riedl, C.J. Saretto, and R. Michael Young. 2003. Managing Interaction Between Users and Agents in a Multi-agent Storytelling Environment. In *Proceedings of the Second International Joint Conference on Autonomous Agents and Multi-Agent Systems*, pages 741–748.

Earl D. Sacerdoti. 1975. The nonlinear nature of plans. In *Proceedings of the 4th International Joint Conference on Artificial Intelligence*, pages 206–214.

Nicolas Szilas. 2003. IDtension: a narrative engine for Interactive Drama. *Proceedings of the 1st International Conference on Technologies for Interactive Digital Storytelling and Entertainment Conference*, pages 187–203.

Nicolas Szilas. 2010. Requirements for Computational Models of Interactive Narrative. *AAAI Fall Symposium on Computational Models of Narrative*.

James M. Thomas and R. Michael Young. 2011. Dynamic Guidance for Task-based Exploratory Learning. In *Artificial Intelligence in Education*, pages 369–376. Springer.

Tom Trabasso and Linda L. Sperry. 1985. Causal relatedness and importance of story events. *Journal of Memory and Language*, 24:595–611.

R. Michael Young and Rogelio E. Cardona-Rivera. 2011. Approaching a Player Model of Game Story Comprehension Through Affordance in Interactive Narrative. In *Proceedings of the 4th Workshop on Intelligent Narrative Technologies*, pages 123–130.

R. Michael Young. 1999. Notes on the use of Plan Structures in the Creation of Interactive Plot. In

Proceedings of the AAAI Fall Symposium on Narrative Intelligence, pages 164–167.

R. Michael Young. 2007. Story and discourse: A bipartite model of narrative generation in virtual worlds. *Interaction Studies: Social Behaviour and Communication in Biological and Artificial Systems*, 8:177–208.

Jeffrey M. Zacks and Khena M. Swallow. 2007. Event Segmentation. *Current Directions in Psychological Science*, 16(2):80–84, October.

Jeffrey M. Zacks, Nicole K. Speer, and Jeremy R. Reynolds. 2009. Segmentation in Reading and Film Comprehension. *Journal of Experimental Psychology: General*, 138(2):307–27, May.

Rolf A. Zwaan and Gabriel A. Radvansky. 1998. Situation models in language comprehension and memory. *Psychological Bulletin*, 123(2):162–85, March.

Rolf A. Zwaan, Mark C. Langston, and Arthur C. Graesser. 1995a. The Construction of Situation Models in Narrative Comprehension: An Event-Indexing Model. *Psychological Science*.

Rolf A. Zwaan, Joseph P. Magliano, and Arthur C. Graesser. 1995b. Dimensions of situation model construction in narrative comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(2):386.

Rolf A. Zwaan. 1996. Processing Narrative Time Shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(5):1196–1207.