From Episodic Memory to Narrative in a Cognitive Architecture

Tory S. Anderson

Georgia Institute of Technology
North Avenue, Atlanta, GA 30332, USA
`torys.anderson@gatech.edu`

Abstract

Human experiences are stored in episodic memory and are the basis for developing semantic narrative structures and many of the narratives we continually compose. Episodic memory has only recently been recognized as a necessary module in general cognitive architectures and little work has been done to examine how the data stored by these modules may be formulated as narrative structures. This paper regards episodic memory as fundamental to narrative intelligence and considers the gap between simple episodic memory representations and narrative structures, and proposes an approach to generating basic narratives from episodic sequences. An approach is outlined considering the Soar general cognitive architecture and Zacks' Event Segmentation Theory.

1998 ACM Subject Classification I.2.0 Cognitive Simulation, J.4 Psychology, J.5 Literature

Keywords and phrases Narrative, Episodic Memory, Cognitive Architecture, Event Segmentation

Digital Object Identifier 10.4230/OASIcs.CMN.2015.2

1 Introduction

Since Tulving's pioneering work on episodic memory [33] it has become apparent that any general model of human cognition must account for memory for temporally and causally situated data just as well as memory for the general facts of semantic memory. It has been observed that we perform extensive narrative sense-making over the data we experience in an effort to gather meaning from our raw experiences [9]; this activity is central to our lives. This ability to cast our experience in narrative terms has been referred to as narrative intelligence [20, 3] and develops through our formative years. Sharing features of both narrative comprehension and narrative generation, narrative intelligence is important to our planning, social interaction, and coping with challenges [23]. This has led to a surge of interest in narrative processes for artificial intelligence [20]; nonetheless, cognitive architectures aimed at modeling human intelligence have been slow to implement support for episodic memory and have as-yet showed few signs of approaching narrative cognition.

1.1 Narrative Intelligence, Comprehension, and Generation

Mateas' definition of narrative intelligence has already been invoked as a guiding concept: the ability to cast our experience in narrative terms. We are here concerned with this sophisticated process, which simultaneously draws from and defies frameworks that attempt to delineate story comprehension from story generation. The input to our model is a stream of experiential data; the process of parsing and selecting from this data, for which Event Segmentation

6th Workshop on Computational Models of Narrative (CMN'15).
Editors: Mark A. Finlayson, Ben Miller, Antonio Lieto, and Remi Ronfard; pp. 2–11
OpenAccess Series in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

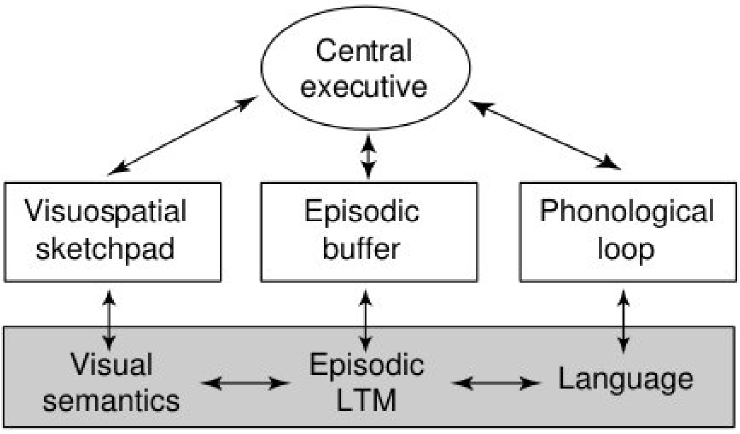

Figure 1 Baddeley's revised working memory model, including the episodic buffer [2].

Theory (EST) will be applied, can be seen as narrative comprehension insomuch as top-down processing occurs to recognize matching narrative patterns. Inasmuch as bottom-up processing is performed upon the received data, a process central to the gating mechanisms of EST, it is similar to some plan-based narrative generation systems which receive a repertoire of actions and use that repertoire to generate a sequence of states as a narrative (e.g. [29]). This reciprocation between narrative comprehension and narrative generation bears striking similarity to the driving tension of cognitive narrative pointed out by Ochs and Capps in their landmark study of personal narratives, described as "the oscillation between narrators' yearning for coherence of life experience and their yearning for authenticity" [23, p. 24]. For cognitive narrative the distinction between narrative comprehension and narrative generation, principle to some notions of intelligence for narrative [17], may need reevaluation.

Importantly, while the joint pair of narrative comprehension and generation are of major relevance to this paper, the distinct process of *story telling*, by which narratives are prepared and committed via some media for purposes that include communication, falls beyond our consideration of cognitive narrative and can be regarded as an activity occurring subsequent to (and using the products of) the processes here proposed.

2 Memory, Segmentation, and Narrative

Narrative exists in the human mind as a particularly important form of mental technology. It's utilization includes experiential sense-making, imputing of causality, categorization and evaluation of events, complex communication, and planning [10]. Narrative cognition is inextricably involved with human memory, particularly the episodic and semantic long-term memory systems. Semantic memory supplies the scripts, schemas, and genres by which top-down processes influence narrative cognition [32, 27], and so plays a vital role in mature narrative intelligence. Evidence from developing narrative intelligence within children suggests that the acquisition of these semantic structures is one of the significant forms of progress as children grow [34][23, ch. 2]. However, the same evidence indicates that however poor, some degree of narrative ability precedes the significant acquisition of semantic narrative structures and that one of the functions of increasing experience is the construction of the scripts and schema that will allow for improved top-down contributions to narrative intelligence. This suggests that narrative intelligence may begin with episodic memory before being augmented with contributions from semantic memory.

Episodic memory is the system responsible for storage of both personal experiences and any other time-situated events attended to second-hand, for example through media or personally communicated stories. It is also implicated for prospective memory used to consider the future [31]. As a distinct memory system it was first proposed by Endel Tulving in 1972 [33]; since that time it has been widely researched. Of particular note is work by Baddeley, who augmented his 1970 model of working memory with an episodic buffer (Figure 1). This episodic buffer was proposed for use in short-term memory complementary to the conventionally understood episodic long-term memory [2]. The role of Baddeley's short-term episodic buffer is as a holding area for retrieved episodes to be integrated cross-modally with data from other sources, such as perception or semantic processing. From a narrative perspective, this may be where stories are constructed through blending with other elements in working and semantic memory, and is likely where narratives are manipulated for many of the afore-mentioned functions of narrative cognition.

The term "episode" excites a notion of scene, events, and change that would seem naturally compatible with most definitions of narrative. However, event recognition itself is an ongoing challenge in computer science. In practice, implementations of episodic memory usually operate as the storage and chronological indexing of system states. In essence, these systems take a snapshot of each state and give it a time label. While narratively intelligent humans are capable of looking at a photo (e.g. of a sport scene) and reconstructing a narrative situation to describe the events surrounding the scene, for these computational systems there has been no obvious way to produce from a life-long sequence of such snapshots a discrete set of narratives.

2.1 Event Segmentation Theory

Event Segmentation Theory (EST) [35, 13, 27] suggests an approach to the problem of dividing a non-delineated sequence of states into events that could become the constituents of narratives. In humans, event segmentation is an ongoing process occurring simultaneously at multiple time/action granularities. According to EST, event segmentation occurs as an effect of ongoing perceptual prediction. During the process of perception two structures participate in parsing the situation and forming predictions: long-term knowledge is brought to bear in the form of event schemata, which are similar to Schanks' and Abelson's scripts [32] and represent the way actions or events normally unfold in similar situations; and working-memory is brought to bear by event models, which are an interpretation of the specific situation at hand. In addition, behavioral models may be used so that predictions can be made based on the presumed goals of the actors in a situation, and world models that account for physical expectations (e.g. the trajectory of an object in free motion). The interplay between the semantic and episodic long-term memory systems in this process is cyclical: semantic memory provides the structures and models to help make episodes from experience, while these episodes are committed to episodic memory where, over time, they help distill further knowledge of semantic structures.

As perception occurs, the mind selects from its knowledge of usual event schemas and uses assumptions about the goals and processes at work in the attended situation to generate expectations of what will happen next. As long as these predictions are mostly fulfilled, the current event model is assumed to continue and no segmentation occurs. However, when the predictions are wrong by some margin of significance, the current event is considered to end and a new event begin in the process of selecting or generating a new event model. These explanations of event segmentation have been supported by evidence from studies of segmentation of event boundaries in written and video narratives [35]. Narratives are

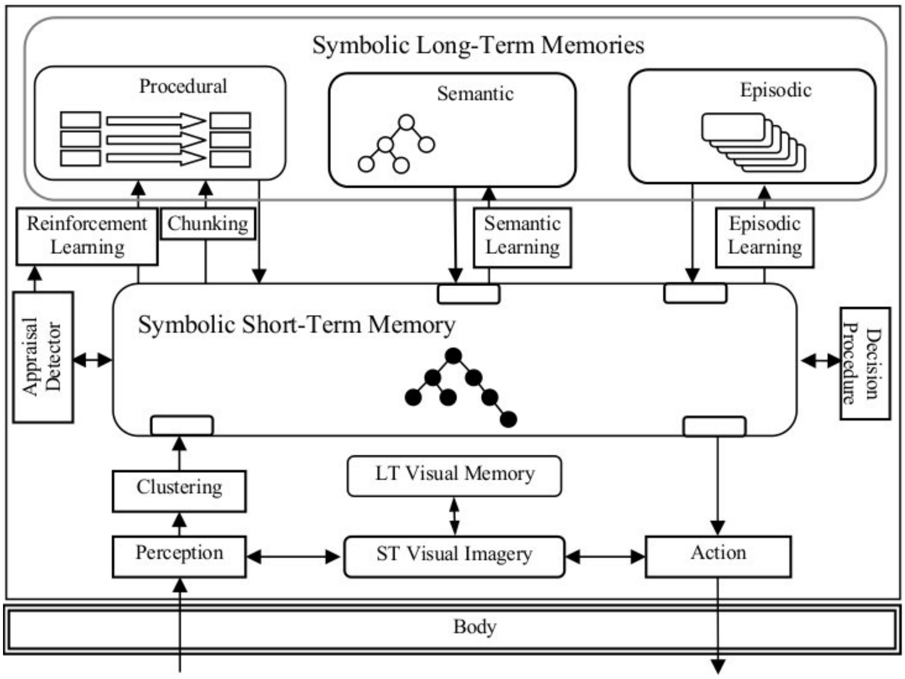

Figure 2 The Soar cognitive architecture [14].

constructed as segmentation occurs at broader granularities over episodic memory, to the point of eventually contributing to production of the life-long autobiographical memories that "make up our own personal narrative of who we are and what we have experienced" [27, ch. 8].

3 An Approach with the Soar Cognitive Architecture

Although it has been explored in a neural network framework [28], EST has yet to be applied in a symbolic architecture. Soar [15] (see Figure 2) is a general cognitive architecture with development overseen by John Laird and is one of the most popular cognitive architectures in current use, with deployments ranging from robotic intelligence to complex battlefield simulation to military training of human soldiers. In addition to an AI system, Soar represents a theory of general human cognition [22]. Soar is a rule-based system in which perception is represented as a graph structure in either working memory or long-term memory. Soar is also agent-based, meaning that instances of Soar run as individual agents independent of, but often interacting with, each other. A given application can call upon large numbers of Soar agents, each running as its own process with its own long-term memory and working memory systems. Soar agents make decisions based on the matching of rules, which depend on the agent's perception of the current state of the world and of its personal state. As a symbolic architecture Soar is well-suited to capturing top-down information such as explicit scripts or subjects of high-level complexity like narrative, whereas it can be difficult to obtain narrative training sets that are both suitably representative and sufficiently sizable for the needs of connectionist models.

Soar's episodic memory modules (epmem) depicted in the top right corner of Figure 2 were added relatively recently and are our central focus. Soar's epmem works by storing snapshots of the working memory state (i.e. the Soar agent's awareness) at each time step, attaching to each snapshot a unique index representing the time of the memory. Once Soar has recalled an episodic memory it is possible to increment forward or backward through the neighboring episodes. Retrieval of episodic memory occurs as queries are issued searching for matching or partially matching features in the graph-structure knowledge representation. Results are given a match score based on how much of the query-graph matches the graphs in an episode, and the best match is returned.

The aim of this project is to outline the addition of rudimentary narrative intelligence within the Soar theory of cognition; we propose to start with narrative intelligence on the most basic of levels, not aspiring beyond child-level narrative intelligence at this point. With this starting point groundwork is laid for future work refining the model.

The implementation proposed proceeds as follows: Soar provides sensory input which is represented in working memory and stored over time as episodes in epmem. These provide the information stream required by EST to make the predictions that result in discrete events. These events are the building blocks of narratives.

3.1 Predictions

At the heart of EST is the making of predictions, which may receive input from a variety of sources including scripts and schema, behavioral character models, genre expectations, and other inputs from semantic memory. As has been previously mentioned the resources available for these processes develops with the experience of the agent. As this exploration considers naive agents with a minimum of prior knowledge it is desirable to have universal heuristics that can form the basis for prediction across domains. Making the simplification that a world consists of agentive and non-agentive components we consider two heuristics. Both of these stand to be superseded as knowledge is gained by the agent.

The heuristic of inertia pertains to non-agentive components of the world, such as spatial configurations. The agent may predict that its environment will continue to exhibit the same features that it now exhibits.

The heuristic of auto-simulation applies to agentive components of the world and takes one of the simplest approaches to a theory of mind by assuming that a perceived agent will act in the same way as the perceiver.

Simplistic as they are, these heuristics provide a ground case to create predictions in any situation, the violation of which delineates the events necessary to form narratives. The result is a stream of events that is, in the worst case of a rapidly and inscrutably changing environment, identical to epmem. With any stability of environment or shared rationality of the agents the product will be an abstraction over the episodes.

3.2 Linking events into narratives

Many definitions of narrative allow for single-event narratives, as when a toddler recalls repeatedly that today “I fell down.” Such interpretation draws no distinction between *event* and *narrative*, a point of ambiguity further promulgated by Zacks’ explanations of EST. The distinction here proposed is not one of structure but of function. EST provides events as a natural kind by which we perceive the world, just as we discern discrete objects. According to EST this perception can occur reflexively. Narrative – particularly personal narrative – is, on the contrary, deliberate and negotiated, the product of an ongoing decision-making process [23] that grows more sophisticated as the narrator matures [4].

Because the aim of this paper is to suggest a means for narrative intelligence that can serve as a (child-like) basis for future work, it is sufficient to allow for single-event narratives while admitting that among the most prominent future work will be the reasoning processes by which more sophisticated narratives can be created from the events produced by EST. These narratives will develop alongside the addition of semantic-memory narrative structures that will influence the top-down processing of EST.

3.3 Considering a Domain: Eaters

While Soar applications are fully capable of recording the richness of real-world perception (e.g. in robotic applications), generating the events with EST which are requisite for narrative generation requires that the system be capable of making useful predictions, which in turn requires rules capturing the complexity of the domain. Games make useful simplified domains. Currently, Soar comes with several game domains that can make testing-grounds for introductory exploration of this approach; we take as an example the Eaters domain [21].

The Eaters game is a two-dimensional Pacman-like game in which one or more colorful "eaters" navigate within a randomly generated maze with the goal of achieving the high score by consuming food pellets of lesser or greater point-values. The eaters are capable of two types of action: moving one space at a time in any of the four cardinal directions, which type of movement has no cost, or jumping up to two squares away, which costs the equivalent of a lesser food pellet. By jumping, an Eater can pass over an obstacle but never consumes food over which it has jumped. When eaters collide, they are each randomly transported elsewhere in the world and their scores are averaged with each other. Each Eater agent has a limited range of vision and discovers the world as it moves. This feature of partial-observability is desirable for mechanisms that rely upon prediction, as does an EST-based approach to narrative intelligence.

3.3.1 Heuristic Prediction in Eaters

Even within so simple a domain as Eaters prediction is still possible and interesting. Because of the partially-observed nature of the domain a natural opportunity for prediction is in world-state itself; for this the heuristic of inertia applies. It happens in Eaters that in randomly generated maps pellets of the same type continue in vertical rows, and that walls may turn but never stagger (do not proceed diagonally or in stair-case formations). The heuristic of inertia means that if the agent has a normal food pellet in front of it as it moves forward, it will predict there to be another food pellet in front after it moves; if not, an event is produced segmenting experience from the previous "normal pellet above" sequence of events. Later reasoning could use this event as a cue to infer that another agent has traversed this path. Likewise, once another Eater has been sighted by an aggressive agent, the heuristic of auto-simulation may come in to play to expect the other Eater to approach. If this doesn't occur, the event might be used in future reflection for the altering of expectations about the unseen portions of the map, or about the schema ("aggressive") of the other agent.

3.3.2 Top-down Narrative Structures in Eaters

A variety of narrative structures could readily be encoded into semantic memory to influence understanding in Eaters. Some such influences could directly influence the production rules applied in Soar by altering the event model being applied. Different event models could include a model for exploration which might apply the afore-mentioned heuristics; prediction error could cue changing to hunting models in which expectations are drawn from heuristics that anticipate perceptual changes that indicate passage of another Eater (e.g. following a trail and expecting pellets to be absent as the trail continues).

3.3.3 Eaters' Narratives

The store of events produced by EST includes segments indicating such things as when a trail of pellets concluded at a wall, or when another eater became visible. In addition to the consideration of these individual events as comprising narratives in their own right, sequences of these events become candidates to be narratives that should be regarded as on a higher hierarchical level than are individual events. Once again the role of top-down structures is important to this production of more complex narratives: as purported by Zacks [35], the changing of event models represents, itself, a key event (e.g. when the agent switches from an exploration model to a hunting model). While the brief model that has been laid out is capable of providing a simple set of event-narratives, these narratives stand to become increasingly interesting and useful as mechanisms for learning semantic structures are introduced.

One of the key features of perception, and hence EST, is the hierarchical nature of perception. Simplified domains like Eaters offer data at a relatively shallow level of abstraction; one way of achieving hierarchical levels of events – and hence higher-level narratives – is by reflection upon episodic memory, by which process broader narrative structures can be applied and recognized. Continuing the Eaters example, reviewing epmem (which contains copies of each state of working memory) can make a place for the application of meta-heuristics, like expecting the heuristic of inertia to apply (say) 70% of the time. This mechanism of heuristics over epmem sequences (rather than singular working memory state) is both naturally precedented by the concept of narrative intelligence, which implies extended temporal breadth, and significant for establishing the recursive nature of narrative.

4 Discussion and Conclusions

The approach to narrative intelligence proposed in this thesis is a preliminary one; it is child-level at best, and awaits further contributions to realize crucial narrative-learning methods that will provide narrative structures, schema, and semantic memory components that are crucial to the next stages of narrative cognition. Such structures proposed by researchers like Propp form the basis of modern narratology and continue to be explored [25, 6, 5]. This model does, however, provide a base-level account for the development of personal narratives from experience. The contribution of this work is to take steps toward a theory of cognitive narrative that bridges the gap between perception and narrative cognition and is, therefore, a comprehensive starting-point for agentive systems. However child-like (even toddler-like) these minimal narratives may be at the start, the function that can provide them will meet needs of both quality and quantity. A system that is able to continually produce narratives from its experiences has the potential to offer the sort of statistical data valuable for categorization and norm detection, both considered some of the fundamental purposes of cognitive narrative in humans [8]. It also offers a promising starting-place for automated generation of scripts within a domain, which could be a useful complement to crowd-sourced script generation that can be costly and unpredictable [18]. Together, these capabilities may serve in support of advanced cognition like goal-based reasoning [30], whereby consideration of narrative schema could provide resources for adaptation or change of goals in dynamic scenarios.

A major question highlighted by the Eaters example with primary relevance to a system's episodic memory has to do with the timing of experiential reflection and personal narrative generation. Although the Eaters example suggests narratives being produced concurrently with perception, much more truthful to work like Ochs' and Capps'[23] is narrative generation that occurs as reflection upon the contents of memory. Indeed, multiple revisits to whatever primitive narratives are produced around perception time will be essential to acquiring higher narrative forms.

Regardless of the episodic memory implementation, a system that produces experiential narratives will also capture qualities of coherence that are desirable in a narrative system. Insofar as narrative is defined as being concerned with having a "continuant subject," [17] experiential narratives minimally satisfy that by providing the experiencer as subject. This fact is not insignificant for applications in Human-Computer Interactions, Expressive AI, or Affective Computing, where "self" for continuity of subject may provide resources for desirable development of personality and style within an agent [12] and ultimately for the development of life story [27].

An event/prediction-based model of cognitive narrative also extends an invitation to insights from the dramatic arts, whose perspective of narrative as affective is highly relevant to the predictions of EST in response to suspense [24], some of which have already applied Soar [19, 11].

A concluding line of work worth mentioning would be observer-systems which would consider primarily other agents as the subject of their predictions and narratives. Such systems would enhance the quality of the narratives generated by developing narratives based on human or expert-system performance and would be important steps toward tasks such as automated sports commentary [1], summarization [26, 16], and theory of mind [7]. One of the severe challenges facing the development of effective observer systems is having an approach to narrative intelligence that can be generalized across domains. The development of general story-generation algorithms suitable for general cognitive architectures is one strategy for approaching such useful systems; hopefully the approach discussed here is a step in that direction.

Eventually narrative intelligence will be an instrument for general intelligence, at which time we could expect that agents with greater narrative intelligence would have a competitive advantage in games like Eaters. As an introductory exploration, the chief product of the approach proposed are the narratives themselves, preliminary to more advanced functions of intelligence.

References

1 Nicholas D Allen, John R Templon, Patrick Summerhays McNally, Larry Birnbaum, and Kristian J Hammond. Statsmonkey: A data-driven sports narrative writer. In *AAAI Fall Symposium: Computational Models of Narrative*, 2010.

2 Alan Baddeley. The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, 4(11):417–423, 2000.

3 David Blair and Tom Meyer. Tools for an interactive virtual cinema. In *Creating Personalities for Synthetic Actors*, pages 83–91. Springer, 1997.

4 Gilbert J. Botvin and Brian Sutton-Smith. The development of structural complexity in children's fantasy narratives. *Developmental Psychology*, 13(4):377–388, 1977.

5 Mark Alan Finlayson. Learning narrative morphologies from annotated folktales. In *Proceedings of the 1st International Workshop on Automated Motif Discovery in Cultural Heritage and Scientific Communication Texts*, 2010.

6 Pablo Gervás. Propp's Morphology of the Folk Tale as a Grammar for Generation. In Mark A. Finlayson, Bernhard Fisseni, Benedikt Löwe, and Jan Christoph Meister, editors, *2013 Workshop on Computational Models of Narrative*, volume 32 of *OpenAccess Series in*

Informatics (OASIcs), pages 106–122, Dagstuhl, Germany, 2013. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

7 Andrew Gordon. The theory of mind in strategy representations. In *24th Annual Meeting of the Cognitive Science Society*, 2002.

8 David Herman. How stories make us smarter. narrative theory and cognitive semiotics. *Recherches en communication*, 19(19):133–154, 2003.

9 David Herman. Storytelling and the sciences of mind: Cognitive narratology, discursive psychology, and narratives in face-to-face interaction. *Narrative*, 15(3):306–334, 2007.

10 David Herman. *Storytelling and the Sciences of Mind.* MIT Press, 2013.

11 Mikhail Jacob, Alexander Zook, and Brian Magerko. Viewpoints ai: Procedurally representing and reasoning about gestures. In *Proceedings of the 6th Digital Games Research Association Conference, Atlanta, Georgia*, 2013.

12 Nicola King. Memory, narrative, identity. *Remembering the Self Edinburgh University Press, Edinburgh*, 2000.

13 Christopher A. Kurby and Jeffrey M. Zacks. Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, 12(2):72–79, 2008.

14 John E. Laird. Extending the soar cognitive architecture. *Frontiers in Artificial Intelligence and Applications*, 171:224, 2008.

15 John E. Laird. *The Soar Cognitive Architecture.* MIT Press, Cambridge, MA, USA, 2012.

16 Wendy G Lehnert. Plot units: A narrative summarization strategy. *Strategies for natural language processing*, pages 375–412, 1982.

17 Boyang Li. *Learning Knowledge To Support Domain-Independent Narrative Intelligence.* PhD thesis, Georgia Institute of Technology, 2015.

18 Boyang Li, Stephen Lee-Urban, Darren Scott Appling, and Mark O Riedl. Crowdsourcing narrative intelligence. *Advances in Cognitive Systems*, 2:25–42, 2012.

19 Brian Magerko and John E. Laird. Mediating the tension between plot and interaction. *Ann Arbor*, 1001:48109–2110, 2005.

20 Michael Mateas and Phoebe Sengers. Narrative intelligence. In *Proceedings AAAI Fall Symposium on Narrative Intelligence*, pages 1–10, 1999.

21 Shelley Nason and John E Laird. Soar-rl: Integrating reinforcement learning with soar. *Cognitive Systems Research*, 6(1):51–59, 2005.

22 Allen Newell. *Unified Theories of Cognition (William James Lectures).* Harvard University Press, 1994.

23 Elinor Ochs and Lisa Capps. *Living Narrative: Creating Lives in Everyday Storytelling.* Harvard University Press, 2009.

24 Brian O'Neill. *A computational model of suspense for the augmentation of intelligent story generation.* PhD thesis, Georgia Institute of Technology, 2013.

25 Vladimir Propp. *Morphology of the Folktale*, volume 9. American Folklore Society, 1958.

26 Dragomir R Radev, Eduard Hovy, and Kathleen McKeown. Introduction to the special issue on summarization. *Computational linguistics*, 28(4):399–408, 2002.

27 G.A. Radvansky and J.M. Zacks. *Event Cognition.* Oxford University Press, 2014.

28 Jeremy R Reynolds, Jeffrey M Zacks, and Todd S Braver. A computational model of event segmentation from perceptual prediction. *Cognitive Science*, 31(4):613–643, 2007.

29 Mark O Riedl and R Michael Young. Narrative planning: balancing plot and character. *Journal of Artificial Intelligence Research*, 39(1):217–268, 2010.

30 M. Roberts, S. Vattam, D.W. Aha, M. Wilson, T. Apker, and B. Auslander. Iterative goal refinement for robotics. In A. Finzi and A. Orlandini, editors, *Planning and Robotics: Papers from the ICAPS Workshop*, page to appear. Portsmouth, NH: AAAI Press, 2014.

31 Daniel L. Schacter, Donna Rose Addis, and Randy L. Buckner. Remembering the past to imagine the future: the prospective brain. *Nat Rev Neurosci*, 8(9):657–661, 09 2007.

32 Roger C Schank and Robert Abelson. *Scripts, goals, plans, and understanding.* Hillsdale, NJ: Erlbaum, 1977.

33 E. Tulving. *Organization of memory.* Academic Press, New York, NY, 1972.

34 Gillian Wigglesworth. Children's individual approaches to the organization of narrative. *Journal of Child Language*, 24(02):279–309, 1997.

35 Jeffrey M Zacks, Nicole K Speer, Khena M Swallow, Todd S Braver, and Jeremy R Reynolds. Event perception: a mind-brain perspective. *Psychological bulletin*, 133(2):273, 2007.