Evaluation of Intelligent Camera Control Systems Based on Cognitive Models of Comprehension

Arnav Jhala
Center for Computer Games Research
IT University of Copenhagen
Rued Langgaards Vej 7, Copenhagen, Denmark
arjh@itu.dk

R. Michael Young
Department of Computer Science
North Carolina State University
Raleigh, NC, USA - 27606
young@csc.ncsu.edu

ABSTRACT

We propose a novel evaluation methodology for intelligent camera control systems based on established techniques of measuring story comprehension from cognitive psychology. The proposed methodology can be used specifically for evaluating the effectiveness of the camera system in communicating the story. We introduce a psychological model of question answering called QUEST and present a preliminary evaluation design of videos automatically generated by Darshak, an intelligent cinematic camera planning system. Initial results from our analysis are encouraging and motivate further work in evaluation of intelligent camera control systems as well as cognitive models of story comprehension through the visual medium.

Keywords

Intelligent Camera Control, Computational Models of Narrative, Discourse Comprehension, Visual Discourse

Categories and Subject Descriptors

H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems—*Evaluation/Methodology*

1. INTRODUCTION

Evaluation of intelligent camera control systems is a challenging problem [3]. There are several dimensions across which camera systems can be evaluated. Most current camera systems are evaluated based on their performance in terms of speed of calculating camera positions. While it is difficult to evaluate stylistic capabilities of such systems, it is possible to evaluate their efficiency in communicating the underlying narrative content. In this short paper, we introduce an experimental design to compare the effectiveness of different visualization strategies in communicating a story. Our approach is based on established cognitive models of story understanding [4] that have been successfully used to evaluate plan-based computational models of narrative [2].

Foundations of Digital Games April 26-30, 2009 Orlando, Florida, USA

Christian and Young [2] have shown that partial order planning representation of stories can serve as plausible cognitive structure for representing a user's comprehension of those stories. In their approach they mapped the story represented as a plan data structure to the knowledge structure of an existing cognitive framework called QUEST [4]. They evaluated the efficacy of their model by using the same techniques to evaluate the computational model that were used to evaluate the cognitive model. In this paper, we further extend the evaluation of plan-based representations of the story by investigating the effects of different visualization strategies on comprehension for stories represented by the same underlying knowledge structure. To evaluate the effects of different visualization strategies we prepared three visualizations of the same story, one with a fixed camera position within the setting, one with an over-the-shoulder camera following the protagonist, and one with camera plan automatically generated by Darshak, a discourse planning algorithm [5] that generates plans for camera shots from a given story. Our purpose for running these experiments is two-fold: First, we want to investigate whether visualization strategies do indeed affect comprehension. Second, we seek to evaluate the quality of visualization generated by Darshak using a representation of camera shots as communicative actions. That is, whether visualizations generated by Darshak are coherent, is determined by measuring the perceived attributes of the story by viewers. Empirical evaluation of such a subjective metric is challenging due to the following reasons: a) Viewers rarely share a common definition of coherence. They cannot be asked directly to give judgement on coherence. b) Viewers differ in their perceived coherence and the values reported by them cannot be directly mapped to a unifrom scale. c) Coherence is a property of the fabula plan itself. The audience perception of the story, rendered in a communicative medium and presented, is altered by the transformation of the fabula plan into the communicative elements. Any evaluation of the communicative elements must take into account this inherent coherence in the fabula itself. d) Viewers' likes and dislikes about the story are also affected by the quality of character dialog and animations. It is difficult for subjective surveys to control for these effects.

To address these challenges, we empirically evaluate discourse generated by Darshak through indirect evaluation of cognitive representations formed by the audience members during the presentation of stories it creates. While the cognitive representations of the story are not directly accessible to the audience, their structure can be inferred by measuring audience members' performance on various cognitive tasks related to the story such as recall [1] or question-answering [4]. Measurements of cognitive tasks have been found to be successful in measuring perceived coherence in fabula plans generated by planning algorithms [2].

2. BACKGROUND

In the QUEST model [4] stories are represented as conceptual graph structures which contain concept nodes and connective arcs. These graphs are called QUEST Knowledge Structures (QKSs). They describe the reader's conception of narrative events and their relationships. Nodes and arcs in a QKS structure are based on their purpose in the narrative. For instance, if nodes A and B are two events in a story such that A causes or enables B then the events A and B are represented by nodes in the QKS graph and they will be connected by a Consequence type of arc.

Techniques used by Graesser et. al. to validate the QUEST model are based on goodness-of-answer (GOA) ratings for question-answer pairs about the story shown to viewers. GOA ratings obtained from the viewers are compared to ratings predicted by the QUEST model. The purpose of this model is to show that viewers build cognitive representations of the story they see that capture certain relationships between events in a story. These cognitive representations can then be queried through the question-answer pairs about the story. QUEST supports questions of types why, how, when, enablement, and consequence. An example story and its corresponding QKS structure is shown by Graesser, Lang, and Robers ([4], Figure 1). Each event and goal is represented as a node in the QKS structure. The links in a QKS structure represent the different types of relationships between events and character goals within a story. Consequence(C), Reason(R), Initiate(I), Outcome(O), and Implies(Im) are the types of relationship arcs between event and goal nodes in a QKS structure.

3. EXPERIMENTAL EVALUATION BASED ON QUEST

We present the design of a pilot experiment to evaluate story comprehension across three different visualization strategies. In this experiment, our method was modeled after experiments carried out by Christian and Young [2]. Christian and Young's experiments were carried out to evaluate the effectiveness of computational models of stories. They were themselves based on the design used for evaluating QUEST [4].

3.1 Design

Our experiments were carried out on 30 participants, primarily undergraduate and graduate students from Computer Science. We used 2 stories (S1 and S2) and three visualization strategies for each story (V1-fixed camera, V2-over-the-shoulder camera angle, and V3-Darshak driven camera) yielding 6 treatments. Treatments were identified by labels with story label as prefix followed by the label of the visualization. For instance, S2V1 treatment would refer to a visualization of the second story(S2) with fixed camera angle strategy (V1) Participants were randomly assigned to one of 6 groups (G1 to G6). Each participant was asked to rate question-answer pairs of three forms of how, why and what enabled for videos representing two unique story and visualization pairs.

Table 1: 2x3 Youden squares design for the experiment.

Viz	Master Shot	Over The Shoulder	Darshak
S1	G1,G4	G2,G5	G3,G6
S2	G5,G3	G6,G1	G4,G2

3.2 Procedure

Each participant went go through three stages during the experiment. The entire experiment was carried out in a single session for each participant. Total time for a single participant was between 30 and 45 minutes. Initially, each participant was briefed on the experimental procedure and was asked to sign the consent form. They were then asked to read the instructions for participating in the study. After briefing, users watched a video of one story with a particular visualization according to the group assignment (Table 1). For each video, users provided GOA ratings for 10 question-answer pairs. Participants were asked to rate the pairs along a four point scale (good, somewhat good, somewhat bad, bad). This procedure is consistent with earlier experiments. Next, they watched a second video with a different story and visualization followed by a questionnaire about the second story. The videos were shown in different orders to common groups in order to account for discrepancies arising from the order in which participants were shown the two videos.

4. CONCLUSION

We presented a novel evaluation technique to measure effectiveness of intelligent camera control systems to communicate a story. The evaluation is based on an established cognitive model of story comprehension. We also presented an experimental design to compare several different visualization strategies. Our initial results are encouraging. We found significant differences between GOA ratings obtained from participants viewing different visualizations of the same stories in support of our hypothesis that different visualization strategies did affect comprehension. We also found significant correlation between GOA ratings predicted by the QUEST predictors and two of the three visualization strategies.

5. REFERENCES

[1] P. V. Broek. The effects of causal structure on the comprehension of narratives: Implications for education. *Reading Psychology*, 10(1):19–44, 1989.

[2] D. B. Christian and R. M. Young. Comparing cognitive and computational models of narrative structure. In *AAAI*, pages 385–390, 2004.

[3] M. Christie, R. Machap, J.-M. Normand, P. Olivier, and J. Pickering. Virtual camera planning: A survey. In *Smart Graphics*, pages 40–52, 2005.

[4] A. C. Graesser, K. L. Lang, and R. M. Roberts. Question answering in the context of stories. *Journal of Experimental Psychology:General*, 120(3), 1991.

[5] A. Jhala and R. M. Young. Representational requirements for a plan based approach to virtual cinematography. In *AIIDE*, 2006.