ANSWER: A Semantic Approach to Film Direction

Ajay Chakravarthy, Richard Beales, Paul Walland
IT Innovation Centre
University of Southampton
Southampton, UK
{ajc, rmb, pww}@it-innovation.soton.ac.uk

Angelos Yannopoulos
Telecommunications Laboratory
National Technical University of Athens
Athens, Greece
ang@telecom.ntua.gr

Abstract— In this paper we present ANSWER, an innovative approach to film direction. Here we describe a methodology to semantically model the film domain in a way which is coherent with the director's intent during film production. To achieve this, we are developing a system architecture which will provide the director with the necessary tools and services to author a scene description through intuitive gesture based graphical user interfaces, which will in turn populate the underlying model with a rich set of semantic descriptions. These semantic descriptions will be used to render the scene graphically through animated pre-visualizations. A director using the ANSWER methodology will be able to understand and assert certain film making decisions before film production begins.

Semantic Web; RDF; Ontologies; Reasoning; Visualization Engine; Film Production; Post Production.

I. INTRODUCTION

Ontologies are the backbone of the semantic web and are essential in order to enable knowledge sharing and re-use across domains. A well designed domain ontology should capture all the essential elements that conceptualize a domain, and the domain expert (Directors in the case of ANSWER) should be easily able to understand and pose competency questions based on the ontology. Further automation of information processing should be possible by software agents that make use of the ontology [6]. This paper presents the approach we are following to model the ANSWER Ontology. The ANSWER ontology will model the semantics necessary to interpret the process of film production consistently for all users (directors, producers, DOP's (Director of Photography), grips, post-production, lighting, etc). In its initial stages the ontology will cover the domain of film production (and other media applications such as commercials, pop videos, 3D effects, cartoons and computer graphics). However, the film and games industries are converging in the way that their creative content is produced, so we will extend the results to include the needs of the games industry offering a bridge between digital media production and animation for game design.

In Section 2, we will first introduce the system architecture that will enable the creation, evolution, maintenance and population of the ontology. Section 3 describes the development methodology being applied to develop the context and core ontologies. Section 4 will give an insight into the intuitive gesture based graphical user interface. Finally in section 5 we will highlight our plans to apply the ANSWER methodology to potential use cases.

II. ANSWER ARCHITECTURE

The ANSWER architecture (shown in conceptual form in Figure 1) is based on a web service architecture model. This way each component is sufficiently decoupled from each other in order to enable both interdependent and independent functioning of these modules. The graphical user interface (GUI) is provided by the Editor and enables the user (Director) to create a scene description using DirectorNotation, a symbolic language intended to express the content of film, much as notes provide a language for the writing of music [3][4][5], in an intuitive and user friendly manner. While the user is populating the notation score, the editor passes this information over to the Rule Engine. The Rule Engine is a web service provider that performs semantic analysis of the DirectorNotation and converts the score into a semantic model after applying reasoning based on the underlying ANSWER ontology. Here the elliptical user input is transformed into its "fully resolved form", which is still high-level compared to the inputs of a typical graphics engine, but is complete in the sense that all required information is provided explicitly rather than implicitly [7]. The Instant Reality Visualization Engine transforms user notation input, as delivered by the Rule Engine, into a rendered 3D animation for pre-visualisation purposes.

The conceptual semantic representation of the symbols introduced within DirectorNotation will be covered both by the context and core ontologies. (Section 3.1). Hence the GUI which enables the user to represent a scene will generate instances which will be mapped both to the lower level core ontology as well as the broader context ontology. In addition to ontology population, the GUI will also act as a means to receive user feedback from domain experts. This feedback will be an essential input for the evolution of the ontology. Further the metadata generated at this stage will also be passed on to the synchronization module for post production.

Bones[1] is an open post-production framework developed by Thomson Grass Valley, which enables the end-user to retrieve information regarding the markers of both the actual footage and the 3D animation from the knowledge base in order to perform synchronization. Bones will display both

---

[1] http://www.thomsongrassvalley.com/products/film/bones/

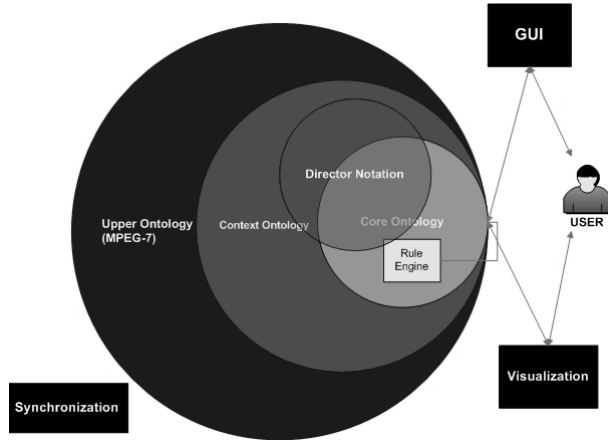streams (3D animation and actual footage) synchronised using corresponding markers.

Figure 1: ANSWER conceptual architecture.

In the next section(s) we will detail two aspects of the architecture: the ANSWER ontology and the Graphical User Interface. Other components of the architecture (Rule Engine, Synchronization, Bones and Game Engine) will not be discussed further since they are outside the scope of this paper.

III. ONTOLOGY DEVELOPMENT METHODOLOGY

The ANSWER project has a core vision involving artistic notation and a core ontology which models it, supporting its automated processing. The project also involves a series of experiments intended to build the most useful tools possible around this core. Examples of such experiments are: various ways to visualize filming-intentions in automatically-generated, non-production-quality animations; ways of creating a "linking", context ontology, that will allow broad re-use of the core ontology in many more applications; application of the artistic notation within a Dialogue Editor designed for game development; and more.

We are using the DILIGENT [8] augmentation model for collaborative ontology development in ANSWER. Unlike traditional approaches to ontology development [1][2], DILIGENT supports ontology development in a multi user and distributed environment. It uses Rhetorical Structure Theory which tries to model the coherence of texts created during user discussions about the structure and content of the ontology. Also DILIGENT allows cost estimation of the ontology based on a formal and statistical methodology.

The ontology development process has been split into two sub tasks: the context ontology covers the broader aspects of film and digital media and will also encompass the broader concepts from the games domain; the core ontology is application specific and will cover the concepts necessary to populate the scene description using the DirectorNotation. In order to maintain consistency over the structure of the ontology we will adhere to standardized ontology engineering development methodologies some of which have

been developed as a part of the WonderWeb[2] project and have been coined under the name CODeP (Conceptual Ontology Design Engineering Patterns).

A. Context ontology

The ANSWER context ontology is being developed in two stages, firstly an initial version was designed after carrying out a thorough study of the film domain [3] and collaborative sessions were carried out with experts in the field to further refine the ontology. In the second stage the initial version of the ontology is hosted on a collaborative Protégé server [10] and iterative ontology development sessions are carried out between project partners. Some of the main aspects captured in the context ontology include FILM_GRAMMAR which refers to some standard conventions followed by directors while shooting a scene. The reasoning applied over these rules will ensure both semantic and syntactic validity of the notation scores. An example of a rule used in film production is the 30 Degree rule: A shot of an object from a particular camera angle X in a time frame T should change by at least 30 degrees angle if the same object is being shown in the next frame. Other prominent elements of the context ontology include Theatrical Presentation Elements. These represent the most commonly used words in the film domain: Shot Angles, Camera Angles, Camera and Lighting Types etc. Figure 2 shows the visualization of part of the Theatrical Presentation Element concept within the context ontology.
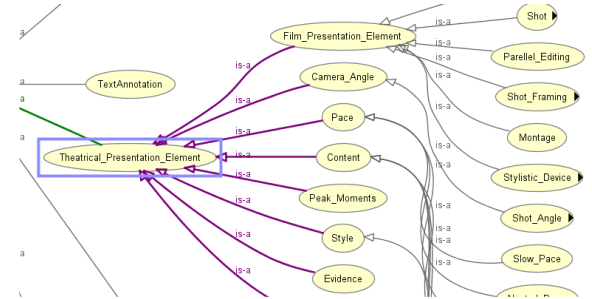

Figure 2: ANSWER context ontology (partial view).

We have re-used the MPEG-7 [9] upper level ontology as the foundational ontology. MPEG7 was deemed to be a very relevant core ontology within ANSWER because the elements that standardize MPEG7 provide support to a broad range of applications, ranging from Multimedia digital libraries, broadcast media selection, multimedia editing, home entertainment devices etc. Further it also covers many of the specific aspects related to content description such as Colour and Texture Descriptors, support for 2D and 3D structures and Motion Descriptors which describe the Camera Motion, Motion Trajectories and Motion Activity. MPEG-7 due to its granularity supports knowledge sharing across wide application domains.

The Core (or Task/System) ontology will include DirectorNotation specific concepts like the Movement Description of actors, and the dependency information of how the notation symbols are placed. This is currently under construction.

IV. GRAPHICAL USER INTERFACE

While Director Notation is a radically new approach to film planning, our intention is that it should integrate as easily as possible into established production planning workflows. Indeed for an application to have a chance of being adopted in the fast-paced world of the creative industries, we believe this is absolutely essential. Consequently the GUI (Figure 3) presented has to be familiar and intuitive, reflecting, where possible, directors' existing working practices.

Our challenge was to support effective knowledge elicitation and capture and population of the context ontology, without pushing those activities to the fore. Some basic principles were established. The Notation symbol set, the symbol syntax and the ontologies will be loaded dynamically each time the UI is launched so that Notation and ontologies can continually evolve. Working methods associated with paper-based film scripts will be embraced in our UI; we will allow the director to start working with the text of the script, identify outline shots, and then proceed to detailed specification of individual shots with Notation – an additional step that allows him to refine his ideas further.

Knowledge elicitation for the core ontology design is being performed as the project's work on developing the ontology structure (without instances) Core ontology population happens automatically by translating the score into an ontology-based technical representation, when the user uses the implemented system. We are currently experimenting with gesture-based interactions – loose strokes of the mouse or tablet pen, akin to pen annotations on a paper script, that allow the director to summon-up symbol palettes or individual symbols without laboriously clicking through menus. Notation elements are placed onto a vertical timeline; each symbol, or group of symbols, becoming a visual metaphor for an aspect of the director's internal creative vision.

Once placed on the score, the director is able to circle a group of symbols that represent a particular concept and ask the system to suggest a semantically-equivalent 'meta-concept' that can be represented by a single symbol. The aims of this are twofold: to make the score itself easier for the director to read, and to clarify the director's intent to the system so that a better visualisation might result. If the director accepts the suggestion, this information is captured by the system and used to adjust the concept similarity weightings. Alternatively if the director feels none of the suggested meta-concepts are appropriate, he is able to use a 'propose concept' button, which allows him to propose a new named meta-concept and add a rationale for its inclusion in the ontology. This is communicated anonymously to the knowledge engineers so that they can consider its inclusion in the ontology.

Initially this will be done manually – the development team will receive automatically generated emails containing directors' suggestions and assess how each might be integrated into a revised ontology. However we are also interested in automating this process, so that directors can, to some extent at least, directly remould the ontology, without simply causing chaos and confusion to other users. In essence what we are doing is extending the DILIGENT ontology development methodology, with its emphasis on communication between domain experts and knowledge engineers, from the ontology planning phase into day-to-day operation of the system.

To-date we have performed informal evaluation of the system and are currently formalising a concrete approach and are in the process of creating an Expert Group with whom we will conduct formal evaluation/trials.
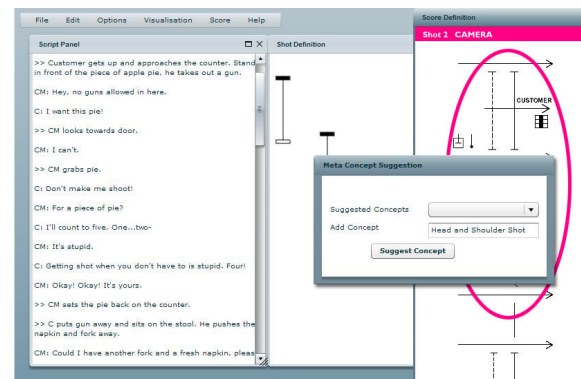

Figure 3: ANSWER GUI

V. POTENTIAL USE CASES

The initial application of the ANSWER methodology is expected to be in the preparation and sharing of a director's vision of a production, however it is anticipated that this tool will become much more widely used in future, and will have benefits throughout the film production chain from pre-production to post-production as well as having application in games design and realization. In order to understand the working practices of those involved in film production, and how those working practices might evolve when the ANSWER tool is available, we are actively involved in engaging directors and other industry professionals in the critical assessment of the tools as they are developed. The principal engagement with the tool for any professional will be through the Graphical User Interface (GUI), and so the design approach to this interface is of paramount importance. We also need to recognise that many different professionals are involved in a typical film production, often exceeding 100 individuals, and each group of professionals, be they directors, cameramen, lighting experts, grip etc. will have their own specialised vocabulary, yet they all have a common understanding of purpose and can generally agree on a director's intention. What the ANSWER approach will add will be a tool which allows directors to explore and develop their artistic intentions without being burdened by technical limitations, and which can be interpreted, via the

associated ontology and applied rule-set, to create an animated visualisation. At the human level it is this visualisation which can be used to communicate with others on set the intentions of the director, which can then assist in achieving the correct interpretation of a scene, and in doing so reduce the number of takes, time on set and hence overall production cost.

As the ANSWER project progresses, we expect that the tool will be extended through the production workflow such that, by synchronising the Director Notation description of the scenes with what is actually shot on set, the post-production process can be improved, since the editing and effects artists will be able to be guided by the intentions of the director, and hence reduce the time during which the director and post-production artists have to work together to reach an acceptable result.

Games production is becoming increasingly similar to film production, in that each encounter by a player with an active game character is played out in a similar way to a shot in a film, with the enacted shot being selected on the basis of context from a large number of possible shots. Hence, these 'shots' need to be created containing appropriate character behaviour and dialogue instances to cover each possible game situation. The description of these shots using the ANSWER approach will ensure that sequences will flow in a natural way, and will speed up the production process, since shots having minor differences can be created very rapidly by modification of an initial template.

Through discussion with production professionals at conferences (such as the IBC[4] event in Amsterdam) and expert group workshops we are gaining valuable feedback on the design and structure of the GUI, and discovering the different ways in which directors work. The principal benefit of our approach is to enable the user to develop and modify their intentions using a range of tools, from textual annotation to shot descriptions and the graphical director notation, and then to relate them immediately to a visual representation and receive warning of inconsistencies which can be corrected before any resource is committed to filming on set. During the course of the project we will expand the expert user base and include other film production professionals to allow us to develop the GUI and the underlying ontology in a way which will most accurately reflect the needs of the users. As they gain experience with using the GUI we anticipate that their requirements of the interface will evolve, and we believe that the design approach we are taking will enable us to work with them to develop an effective production tool.

VI.   CONCLUSIONS AND FUTURE WORK

In this paper we have presented two aspects of the ANSWER architecture: the Ontology and the Graphical User Interface, which will enable the director to describe a scene using a DirectorNotation score and populate the underlying semantic model. Using this approach we have demonstrated how knowledge about the director's intent can be captured such that the metadata can be shared across a variety of application domains. Also it can be an effective way to save costs incurred during the film production process, by enabling the director to get a deeper understanding of the film making decisions before the shooting begins. We are currently establishing an international expert group and plan to carry out a detailed user evaluation of both the ontology and the user interface early next year.

REFERENCES

[1] A.Gangemi et al. Ontology integration: "Experiences with medital terminologies. In the Formal Ontology in Information Systems", e.d., Nicola Guarino, pp. 163-178, Amsterdam, (1998). IOS Press

[2] A.Gomez-Perez et al. Ontological Engineering, Advanced Information and Knowledge Processing, Springer 2003.

[3] A.Yannopoulos, K.Savrami, T.Varvarigou. "DirectorNotation as a Tool for AmI & Intelligent Content: an Introduction by Example." accepted for publication in the International Journal of Cognitive Informatics and Natural Intelligence (IJCiNi).

[4] A.Yannopoulos, K.Savrami, T.Varvarigou. "An introduction to DirectorNotation." accepted with revisions, Multimedia Systems, Springer.

[5] A.Yannopoulos, K.Savrami, T.Varvarigou. "Artistic Notation Systems Integrated in Software Engineering of Knowledge Technologies: emphasising film, animation and games." Accepted with revisions, ACM JOCCH (Journal of Computers and Cultural Heritage).

[6] B. Smith. Ontology. Preprint version of chapter "Ontology", in L. Floridi (ed.), Blackwell Guide to the Philosophy of Computing and Information, Oxford: Blackwell, 2003, 155–166.

[7] Deliverable D3: "System, Components and Technology Specifications." D3 Project Deliverable: ANSWER Project (2008).

[8] H.S. Pinto et al. "DILIGENT: Towards a fine grained methodology for Distributed, Loosely Controlled and Evolving Engineering of Ontologies."

[9] M.J. Martinez. "MPEG-7 Overview: International organization for standardisation (ISO//IEC/JTCI/SC29/WG11. Coding of moving pictures and audio". Palma be Mallorca, October 2004.

[10] T. Tudorache, N. F. Noy, and M. A. Musen. "Supporting collaborative ontology development in Prot´eg´e." In 7th Intl. Semantic Web Conference, ISWC 2008, Karlsruhe, Germany, 2008.

---

[4] http://www.ibc.org/cgi-bin/displaypage.cgi?pageref=100