Intelligent Media Agents in Interactive Television Systems

Hartmut Wittig
Carsten Griwodz

IBM European Networking Center
Vangerowstraße 18
D-69115 Heidelberg

{wittig,griff}@vnet.ibm.com

Abstract:

Future interactive television applications are the upcoming front-end to interactive radio, television, video rental, home shopping, multimedia communication, and information retrieval. The challenge of an interactive television system is not that the medium is digital, nor the ability of browsing media resources, nor reception of multiple kinds of media. Users do not need more information than today. The challenge is to provide information which corresponds to users' interests and needs. This paper describes the vision of a multimedia iTV system whose appearance is personally adapted for each user. The concept of personal iTV systems is presented by explaining the task of personalization and the technical fundamentals of intelligent assistance.

1 Introduction

We experience the appearance of more and more television and radio channels, and of information bases containing terabytes of data. There has never been a time when more data has been produced than today. People can control this information flood only by using strategies for information filtering. Because people have different interests and needs, everyone has his own information filtering strategy which is based on his personal knowledge and experience. In the next years, digital data compression technologies and global broadband networks will be brought into traditional television systems. For example, in the next years more than 500 channels are expected. Because of this mass of information people cannot use their conventional strategies for information filtering.

The key issue is not to get more content but to receive less and *personalized* information. Today, static information filtering can be easily done by comparing attributes with a given profile. *Agent based computing* is one approach to overcome the inflexibility of preliminary help and assistance. Many examples of the application of agents, their programming, and interaction can be found (e.g., [15], [9], or [10]). This paper applies the *intelligent agent* approach to interactive television systems and introduces the novel approach of parsing and interpreting user interactions in the iTV end systems. The intelligent agent thereby provides easier and adaptive handling of daily tasks such as program selection, home shopping, and information retrieval. We introduce the term of the *intelligent Media Agent (iMA)* as the name for an intelligent agent in iTV systems. The iMA acquires knowledge about the users by collecting knowledge about specifics of users, their behavior, their favorites, and their weaknesses.

2 Intelligent Media Agents tasks

The term *agent* is used for any intelligent software vehicle. To the agent theorists, agents are described as programs which imitate intelligent human behavior by *doing things, making decisions, autonomous acting.* and *learning* (see [2], and [18] for more information). By combining the advantages of artificial intelligence and multimedia systems, the term *intelligent Media Agent* is formed.

2.1 Personal TV guiding and movie generation

The part of the iMA which is responsible for all tasks dealing with movies is called *movie robot.* The movie

0-8186-7105-X/95 $4.00 © 1995 IEEE

robot forms the dominant part of the intelligent media agent. There are two approaches to get an overview of the TV program. A very popular method is *channel zapping*; users switch from channel to channel to get an impression of all programs running. Printed TV guides are commonly used for previews to the TV program. However, assuming there will be a TV system accommodating each home with about 500 channels, both conventional strategies will fail: Zapping takes nearly one and a half hour before every channel has been seen (10 seconds per channel). A TV program guide would comprise about 350 pages per week (one page comprises 10 TV channels).

One of the main tasks of an iMA is the collection of *personalized TV guides* containing only movies or programs which are of interest to the user. The iMA allows the description of user demands and wishes, converts them into personal profiles, and applies these profiles to programs and movies when assembling the personal TV guide. In addition to the simple off-line mode where the users specify their profiles, the agent applies mechanisms for automatic profile generation by watching the operations of the users. The iMA uses this feedback to check its offerings and to optimize the personal TV guide.

Interactive movies are movies which contain *branches* in the story, where users can interactively choose major aspects of the next sequences. For example, it could be up to the users to decide whether the antagonist (in which case the movie is tragic) or the protagonist succeeds (a "happy ending"). There is a trade-off between the granularity of the movie, and the degree of adaptation to the wishes of users. The balance of this parameter could be an additional task of the iMA for interactive movies. The iMA adapts to each user by regulating the number of branches to the specifics of the users. If the agent detects that specific kinds of decisions are not important enough for the user to interrupt the program, they are done on behalf of the user by the iMA.

Live transmissions allow users to become directors of the live event. They choose camera positions, initiate interviews, or see instant replays. In this mode users should not be compelled with setting up a directing guide. The user either chooses to follow the standard direction offered by TV channels, or delegate the decisions to the iMA with the possibility to interrupt this mode for important changes.

2.2 Adaptation of user interfaces

Similar to traditional television systems, users come from all social strata, have completely different foundations, and might have no particular technical training. Because there are many users which have no idea of handling computers or video recorders, the design of user interfaces for iTV systems must be different from today's inflexible interfaces. Based on common user interface design principles (e.g., in [14]), interfaces must be adjustable to the properties of users. The variety of functions ranges from intelligent help functions (e.g., [7]) to functions which have knowledge on the properties and behavior of users (e.g., [15], [3]). Knowledgeable people are provided with an extended set of control functions, whereas beginners use simple interfaces. The part of the iMA which is responsible for user interface design is called the *user interface robot*. By storing and using rules containing common knowledge of user interface design, the user interface formed by intelligent media agents seems adaptive and personal, and can be easily switched to special-purpose interfaces by changing the rules of the agent.

2.3 Shopping and interactive services support

Interactive home shopping, banking, and a mixture of various on-line services are some of the driving forces for commercial interests in iTV systems. The iMA helps to navigate through virtual malls and builds virtual malls according to the wishes and budget of the user. It helps to seek special offers or to find the best price for an article by comparing catalogs of service providers. The iMA makes a suggestion for the weekly shopping cart of food. The iMA regards the personal shopping habits in the past and the prices of today. Afterwards, users can add or remove articles. Thereby, the agent can help to reduce cost and to adapt the shopping behavior to the particular budget of the household. This subagent is called *shopping robot*. The iMA also supports distance learning services and tutoring by adjusting the speed and degree of difficulty (see also [16], [5]).

2.4 Information retrieval

Interactive TV system end-stations can also serve as the front-end for interactive information retrieval in global networks. Because there is a large number of ways to get the information required and these ways are rarely structured, beginners need much more time than expert users to get the same work done. For the users this means that they need an expensive learning process which gives them experience and knowledge in effective search strategies (meta-knowledge). Approaches to solve the information retrieval problem can be found in the literature (for example, [1]). In most, the core is an intelligent agent which forms the base for making information retrieval and filtering more effective. A model of software agent for information retrieval is given in [19]. Examples are agents which prepare electronic mails or news. These agents are also called *information retrieval robots*.

3 Prerequisites for personalization

Intelligent management of personal information requires that the particular user of the agent is identified and the presentation is dynamically changed according to the identity of the user.

3.1 User identification and authentication

Before producing and applying profiles to one specific person, the user must be identified. The need for cost management and accounting is one of the driving forces for identification and authentication in iTV systems. Strong authentication methods are also required for every kind of financial transaction. Procedures for identification in iTV systems are (1) identification by iTV end-station address, (2) user identification names and passwords, (3) credit/debit card numbers and PINs.

Because agents have knowledge about users, their characteristics, and their behavior, profiles are protected from unauthorized access by using strong authentication methods similar to authentication functions in the banking environment. If user profiles are stored in the distribution nodes of the iTV system they are be protected by using secure encryption methods.

3.2 Dynamic presentation generation

Individual information management requires to substitute the broadcast character of communication by dynamic creation of information which takes into account the personal needs of users. Current architectures of iTV systems do not allow to dynamically generate new presentation elements. Some of these provide support for interactivity across networks, but none of the existing iTV systems explicitly addresses the problem of dynamically changing multimedia presentations, which require that the presentation structure itself is created or modified while the presentation is already running. A detailed characterization of dynamic generation mechanisms within presentation systems, e.g., based on ISO MHEG, is given in [20]. In general, there are client- and server-based concepts for dynamic presentation generation:

(1) The client-based approach requires a self-contained presentation system where data is either collected at the engine or retrieved directly into the engine, and where the presentation structure is modified within the engine in a presentation-defined way. It puts a high strain on the implementation of the presentation subsystems and requires flexible communication protocols, parsing of structural information at presentation run-time, and standardization of dynamic generation procedures. Either the presentation structure itself is required to integrate dynamical nodes, or the presentation engine must communicate with a *supplemental interpreter* which dynamically modifies the presentation structure in the engine. This means that the iMA must be integrated into the engine, or form the supplemental interpreter.

(2) The server-based approach is that information like user input or user profiling data is transferred back to the information providing side, the server side of the distributed presentation, where the information can be stored, processed, and augmented with user-specific data, and the structural information can be built. It is very flexible, but it requires that a presentation itself be distributed, and that even the structural information be distributed in itself. It does not restrain the iMA from being completely or partially located on the end-station. The structural data which is passed to the engine is already personalized and needs no further dynamic modifications.

The distribution is easily conceived in a page-oriented model (e.g., mark-up languages such as HTML) since no overall presentation structure is given and time-constraints are not even defined. Models without this restriction (e.g., ScriptX and MHEG) allow more flexible mechanisms of dynamic generation. This dynamical behavior is presentation-dependent and completely hidden from the media players. Because multimedia presentations are retrieved as if from a normal presentation server, we call this server a *virtual object store (VOS)*. Figure 1 provides a general sketch of the VOS.

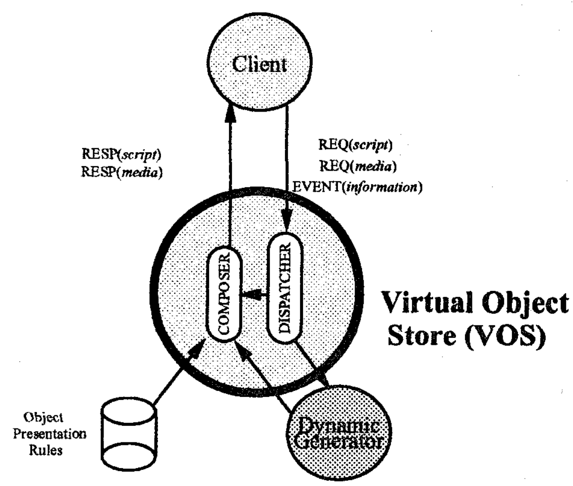

Figure 1: Dynamic presentation generation

This VOS approach does not require that the VOS be physically present on the server side of an interactive TV system. Depending on the equipment of the iTV end-station the VOS may fully or partially run on the end-system. For the engine, the VOS is a server entity from which presentation objects are retrieved. Structurally, the VOS consists of a *dispatcher* and a *composer*. The dispatcher reads each incoming message to test whether it

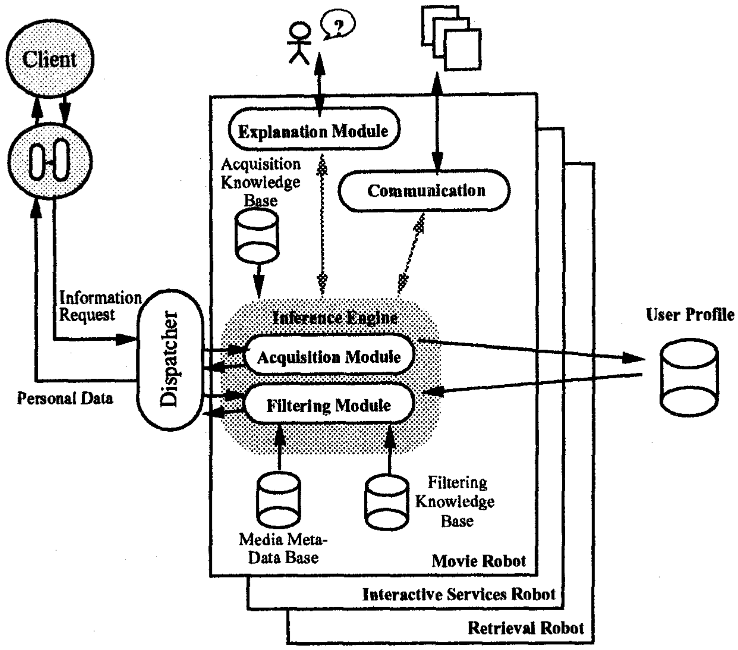

Figure 2: Intelligent media agent

is a script, media request, or information which must be forwarded to the iMA. In the first case, the request is forwarded directly to the composer which is able to deliver the presentation data. In the second case, the dispatcher delivers the information to a *dynamic generator* which stores data for later evaluation or for combination with other information gathered. The composer fulfils requests for media or structure objects from the client. These objects can either be located in a persistent storage device like a hard disk, in which case the retrieval process is a simple download, or they can be dynamically created. For dynamic creation, the dynamic generator provides the dynamic data, but the framework (the rules for object creation) of the presentation can still be stored on a persistent storage device. The mapping to the particular presentation standard is than handled by the composer, and the resulting presentation object is delivered to the client.

4 Architecture

The iMA consists of the movie robot, user interface robot, services robot, and the retrieval robot. These robots have autonomous inference mechanisms, different techniques for knowledge acquisition and filtering, and work on different media meta-data. However, these robots share their knowledge about users.

4.1 Architecture

The architecture of the iMA is shown in Figure 2. Information and requests are forwarded to the dispatcher of the iMA. The dispatcher is the only module which must be shared among all robots. The dispatcher analyses incoming information (e.g., interaction, events) and gives it to the registered robots.The dispatcher also supports inter-robot communication.

Robots use the information provided by the dispatcher for knowledge acquisition, as trigger for filtering operations, or for both. For example, if an interaction event is forwarded to the acquisition module, the user profiles are re-computed by using the acquisition strategies which are stored in the acquisition knowledge base. The knowledge and data bases of the iMA contain knowledge of users and of acquisition and filtering strategies, and information about media. In the case of user profiles local sharing of knowledge is very helpful, because consistency can easily be achieved. Thereby, robots profit from knowledge acquired by other robots. The next sections will describe the details by taking the movie robot as example.

4.2 Media and user profiles

Media profiles are stored in the media meta-data base and contain a description of media, their genre, statistics, and evaluation information. User profiles contain information about wishes, requirements, and personal preferences. In today's TV guides there are some fragments of classification and evaluation: daily tips are given, some guides have a ranking to show what is very remarkable, and there is also a sort of evaluation of some categories (e.g., suspense, humor, or sexuality). A classification scheme for the TV allows the movie robot of the iMA to characterize TV channels as well as the program units.

Attributes of medium

The media database contains statistical information and evaluation data. As a prerequisite for further extensions the intelligent media agent keeps structural attribute information in a meta-database which is used to generate the media database and media attributes. The *statistical information* is very similar to information which can be found in a lexicon. Statistical information is fixed and objective. An example is given:

Title:	STRING	Red Hot
Director:	STRING	Paul Haggis
Country:	STRING	Canada
Year:	INT	1993
Length:	INT	5713 sek
Actors:	STRING	Balthazar Getty, Armi
Speech:	STRING	English

There is also the vision to integrate metrics about scenes in the movie. It is expected that new methods of image recognition and video parsing will have deeper understanding about the content and its semantics. First steps towards the automation can be found in [6], [21]. By using content-based retrieval methods additional statistics can be provided, e.g., the number of kisses, the number of cuts, or the number of people killed.

The *classification* of movies in *genres* is another set of movie attributes. The division of programs into categories is subjective. Simple classification schemes for movies can be found in almost every TV guide. The class "movie" is divided into subclasses: action movies, thrillers, comedies, romantic movies, and many others. User desires and demands can approximately be described by specifying a set of genres. There is a number of top-level classes (e.g., humorous programs) which can contain several subcategories (e.g., comedy, satire). This scheme is used by the movie robot to arrange the program according to its content. By giving every node of the classification tree a certain value (e.g., from 0 to 100) to represent the degree of affiliation of the movie to the categories, nuances of affiliations can be expressed. A movie can be a member of many categories to express that this movie has various facets and cannot be assigned to only one category. The classification scheme of the iMA is given in [20]. The example movie "Red hot" is classified as follows within an interval [0;100]:

Love = 65
Music => Rock = 30
Education = 5

It contains elements of several genres: Because the movie is a typical love-story the most dominant attribute is "love". The movie contains aspects of a music movie (sub-category rock). All other nodes in the classification tree are of no interest to describe the genre of this movie.

The third type of attributes is the *qualitative description* of movies. Qualitative information is required to distinguish between good and bad movies. While the classification attribute only describe the genre of a movie, in the first prototype the attributes "suspense", "humor", "action", and "sexuality" are used to qualitatively describe a program. These attributes are commonly used in printed TV guides. Qualitative information is very subjective because it reflects the personal opinions of the human who specified these attributes. The example movie "Red hot" has the following quality within the interval [0;100]:

Suspense:	INT	65
Humor:	INT	34
Action:	INT	14
Sexuality:	INT	85

High attribute values indicate that the movie fulfils this quality criterion. Sexuality and suspense of "Red Hot" are very good. The high value of sexuality does not mean that "Red Hot" is a sexual explicit movie. This is expressed by a high value of attribute sex in the above classification scheme.

Attributes of user profiles

User profiles consist of a fixed and a variable part. The fixed part is used to store statistical information. Examples are the user name, the date of birth, address, credit card number, or profession. When no other information about the user wishes is available (e.g., if the user want that the robot automatically collects the profile) statistical information is used for automatically setting up an initial user profile. Depending to the age and profession of the user the movie robot derives standard profiles. The variable part of the user profile contains the wishes and demands of users for the TV program. The knowledge acquisition module automatically changes the variable part in order to adapt the user profile to the user's habits of watching. The filtering module uses the profile to select the program which fits best the requirements. One part of the variable user profiles contains information, favorite genre and actors, or the quality of program a user expects. This part is similar to the attributes in the movie data base. Additionally, in the user profiles it is explicitly distinguished between wishes, demands, and "does not matter".

A set of subprofiles can be specified to reflect the variety of user demands. Each subprofile contains attributes which describe the interests of the user according to the specific subcategory. The application of methods from the fuzzy set theory allows us to specify ranges of values which represent uncertainty and vague wishes. The general model of attributes is given in Figure 3. This figure shows the relation between an attribute m and its value in the user profile $U(m)$. The value of m_{base} defines an important value of this attribute in the subprofile because it is the middle of those points where the fuzzy function $U(m)$ is reaches its maximum. If more than one maximum need to be specified this has to be done by defining another subprofile. An interval for the validity of the maximum is defined by m_2 and m_3. The values of m_1 and m_4 define the points of the function where $U(m)$ reaches its minimum. With a set of these fuzzy attributes user subprofiles and user profiles are formed.

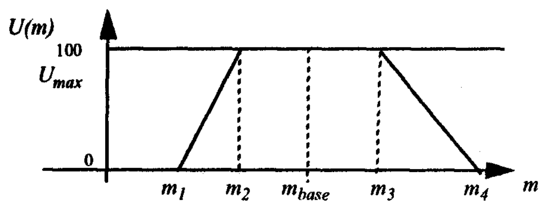

Figure 3: Fuzzy specification of attribute *m*

The knowledge about the behavior of the user is represented in *behavior variables*. These variables represent the accumulated knowledge about the behavior of users. The characterization comprises the attitude and acceptance of the user to iTV, information about the zapping behavior, the favorite times for watching and shopping over the day, week, or month, and the acceptance of specific kinds of advertisement.

4.3 Knowledge acquisition

The knowledge acquisition module cares about the interpretation of incoming events, and uses experiences to compute and change user profiles. Three modes of operations can be distinguished: (1) manual user profile input, (2) automatic user profile generation and adaptation, and (3) automatic adaptation of manually generated user profiles. In the first mode, the user exactly specifies what kind of program he wants to see. The manual specification requires a user profile front end. User profiles can thereby exactly be installed. The main disadvantages of manual specification are that users might be confused with such a complex task, and if the wishes changes the user profiles must manually be adapted. In the automatic mode the movie robot watches the user interactions for a certain time span, and records programs and their duration. From this list of selected programs and the evaluation of the programs in the media data base, the agent recognizes favorite programs, genres, and the behavior of the user. The third mode is a mixture of the above.

The acquisition knowledge is represented in a set of acquisition rules. Each rule consists of a precondition and a postcondition. If the precondition is true the rule is suited to interpret the incoming event, the rule is put on the *agenda*. The preconditions of this rule will be applied. Examples for knowledge acquisition rules are given in Table 1. The first rule describes how channel zapping is detected. If there is a high rate of channel selections, the precondition of the first rule is true. In the postcondition, the internal zapping variable of the user is increased by a function containing a parameter for the time between two switches. The second rule is used to add new subprofiles to the user profile of a "zapper". If the movie robot sees that the zapper had seen a program for a longer time than the zapping time, this is interpreted as an interest for this program. The profile of this program is added to the user profile, or is merged with existing profiles. The complete knowledge acquisition base contains hundreds of rules.

Because it is too expensive to keep the profiles of all programs ever seen in a history list, a clustering algorithm described in [13] is used to compare and merge user profiles.

4.4 Filtering module

The filtering module compares a user profile with the TV program and tries to find similarities between user wishes and the offered program. The comparison is done by sequentially comparing each subprofile with each movie profile. The result of a comparison is not to find out whether a movie fulfils the user profile, or not. The result is a parameter which expresses the degree of correspondence between the movie and the profile (for details see [20]). A hotlist of movies can thereby be formed which contains the TOP 20 movies of a user. In the description of knowledge acquisition it has been pointed out that a user profile can contain wishes, demands, and "do not care" attributes. This is taken into account by the comparison functions by (1) comparing wishes as described above, (2) refusing movies which do not fulfil demands, and (3) not regarding the "do not care" attributes. Similar to the acquisition knowledge, the accumulated knowledge on filtering strategies is kept in rules of the filtering knowledge base.

Table 1: Examples for knowledge acquisition rules

Precondition	Postcondition
(event = channel_switch) (time_since_last_switch < zap_threshold)	(zap_factor_inc.time_since_last_switch)
(event = channel_switch) (time_since_last_switch > zap_threshold) (zap_factor.low) (old_channel. not_in_profile)	(merge_profile.old_channel, time_since_last_switch)

4.5 Explanation and communication modules

The explanation module helps users to understand decisions of the iMA. The intelligent media agent will explain its ranking functions, and describe what are the influences to its decisions, and why the robot not decided to choose this specific movie. Users can also complain about the quality of decisions. Thereby the agent can register that his strategies are partly wrong, and try to change and optimize its parameters. This dialog between users and the explanation module cannot only be used to answer questions of users. Depending upon the kind of questions and their quality, additional assumptions about the user can be made.

The communication module is used for further exchange of acquisition and filtering strategies among intelligent media agents. This can be done by using standard knowledge interchange formats (e.g., Knowledge Interchange Format KIF [11], Knowledge Manipulation and Query Language (KQML) [8]). An alternative approach is the Common Object Request Broker Architecture (CORBA) by the Object Management Group which allows the sharing of distributed objects as representations of knowledge.

5 Conclusions

Main parts of the iMA and the movie robot are implemented. An administration tool allows adding, deleting, and changing of user and movie profiles. Matching functions evaluate the similarity of user profiles and media profiles. Clustering methods derive new and adapt existing user profiles from the set of programs viewed. The knowledge databases are implemented as relations of the POSTGRES database system. The content of the german TV program is saved into the movie meta data base. To simulate the five hundred TV channel which are forecasted, old programs are stored in the database as a new channel.

The first iMA prototype uses the interactive television system GLASS (Globally Accessible Services [4]) and generates MHEG presentation output. The prototype user front-end is shown in Figure 4. A complete implementation involves an expert system shell (CLIPS) and is scheduled for the middle of 1995. The password authentication method is implemented and the dynamic presentation module (VOS) is designed. In anticipation of the hundreds of TV channels, it is planned to integrate foreign TV guides (e.g., the American "What's On Tonite!" internet TV guide).

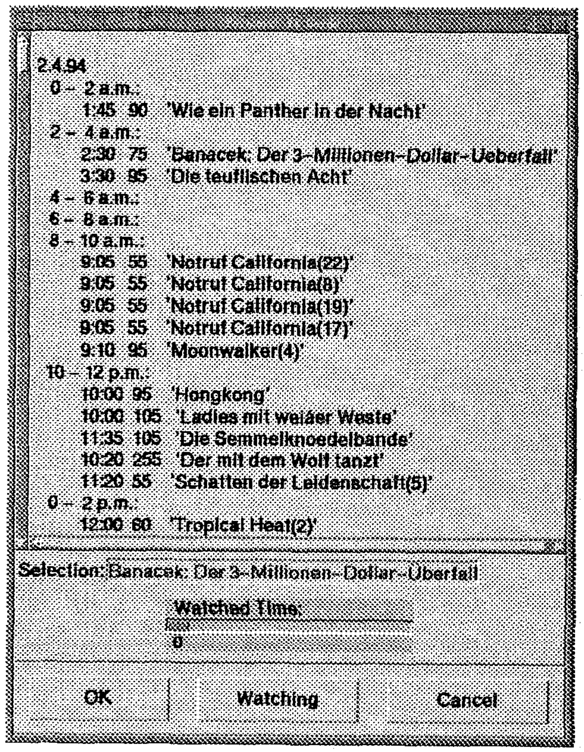

Figure 4: Snapshot of the Tcl/Tk frontend program

References

[1] Alberico, M. Micco: *Expert Systems for Reference and Information Retrieval. Meckler.* London, 1990.

[2] G.A. Boy: *Intelligent Assistant Systems.* Academic Press, San Diego, 1991.

[3] D. Benyon, D. Murray: *Experience with adaptive interfaces.* Computer Journal 31(5), pp. 465-473, October 1988.

[4] DEC, GMD Fokus, Grundig Multimedia Solutions, IBM ENC, Technical University of Berlin PRZ: *BERKOM Globally Accessible Services: System Specification 1.0.* BERKOM, Berlin, May 1994.

[5] W.J. Clancey: *Knowledge Based Tutoring: The GUIDON program.* MIT Press, Cambridge, MA, 1987.

[6] T.-S. Chua, S.-K. Lim, H.-K. Pung: *Content-based retrieval of Segmented Images.* Proceedings of the 2nd ACM Multimedia, October 1994.

[7] J. Erlandsen, J.Holm: *Intelligent Help Systems.* Information and Software Technology, 29(3), pp. 115-121, 1987.

[8] T. Finin et al.: *Specification of the KQML Agent-Communication Language.* Technical Report EIT TR 92-04, Palo Alto, CA, 1992.

[9] I. Greif: *Desktop Agents in Group-Enables Products.* CACM, vol. 37, no. 7, July 1994.

[10] M.R. Genesereth, S.P. Ketchpel: *Software Agents.* CACM, vol. 37, no. 7, July 1994.

[11] M.R. Genesereth, R.E. Fikes: *Knowledge Interchange Format Version 3 Reference Manual.* Logic-92-1, Stanford University Logic Group, 1992.

[12] ISO/IEC: *Information Technology - Multimedia and Hypermedia information coding Expert Group,* Committee Draft, 1994.

[13] R.A.Jarvis, E.A. Patrick: *Clustering Using a Similarity Measure Based on Shared Near Neighbors.* IEEE Transactions on Computers, vol. 22, no. 11, November 1973.

[14] B. Laurel: *The Art Of Human-Computer Interface Design.* Addison-Wesley, Reading NY, 1991.

[15] A. Markus: *Future Directions in Advanced Interface Design.* Communication with Virtual Worlds, Springer, 1993.

[16] P. Maes, R. Kozierok: *Learning Interface Agents.* Proceedings of the 11th National Conference on Artificial Intelligence, July 1993.

[17] D. Riecken: *An Architecture of Integrated Agents.* Communications of the ACM, vol. 37, no. 7, July 1994.

[18] N. Shardlow: *Action and agency in cognitive science.* Master's thesis. Dept. of Psychology, University of Manchester, Uk, 1990.

[19] E.M. Voorhees: *Software Agents for Information Retrieval.* AAAI Spring Symposium on Software Agents, AAAI Press, pp. 126-129, Menlo Park CA, March 1994.

[20] H. Wittig, C. Griwodz: *Intelligent Media Agents for Interactive Television Systems.* IBM ENC Technical Report 43.9501, Heidelberg, January 1995.

[21] Zhang, Y. Gong, S.W. Smoliar, S.Y. Tan: *Automatic Parsing of News Video.* Proceedings of 1st IEEE ICMCS, Boston, May 1994.