I T H E A

International Journal

INFORMATION TECHNOLOGIES & KNOWLEDGE

2007 Volume 1 Number 3

International Journal
INFORMATION TECHNOLOGIES & KNOWLEDGE
Volume 1 / 2007, Number 3

IJ ITA is official publisher of the scientific papers of the members of the Association of Developers and Users of Intellectualized Systems (ADUIS) and the Association for Development of the Information Society (ADIS)

IJ ITK rules for preparing the manuscripts are compulsory.
The **rules for the papers** for IJ ITK as well as the **subscription fees** are given on ***www.foibg.com***.
The camera-ready copy of the paper should be received by e-mail: info@foibg.com.
Responsibility for papers published in IJ ITK belongs to authors.

General Sponsor of IJ ITK is the **Consortium FOI Bulgaria** (www.foibg.com).

International Journal "INFORMATION TECHNOLOGIES & KNOWLEDGE" Vol.1, Number 3, 2007

Printed in Bulgaria

Edited by the **Institute of Information Theories and Applications FOI ITHEA**, Bulgaria, in collaboration with the **V.M.Glushkov Institute of Cybernetics of NAS, Ukraine**, and the **Institute of Mathematics and Informatics** and the **Institute of Information Technologies, BAS, Bulgaria**.

Publisher: **Institute of Information Theories and Applications FOI ITHEA**
Sofia, 1000, P.O.B. 775, Bulgaria. www.foibg.com, e-mail: info@foibg.com

ISSN 1313-0455 (printed) **ISSN 1313-048X (online)** **ISSN 1313-0501 (CD/DVD)**

ADVERGAMES: OVERVIEW

Eugenio Santos, Rafael Gonzalo, Francisco Gisbert

***Abstract**: Advergame is a new marketing concept that has appeared due to the fact that young people are always connected to the Internet, are using mobile services such as SMS and MMS, or are chatting with instant messenger services and they spend too much time just playing in a stand alone way or in a network game. A new revolutionary service is the advergame one; that is a game with advertisment capabilities. Any company can develop an advergame that is, a game with some kind of advertising process of this company. This paper introduces some idea and concepts when developping an advergame..*

***Keywords**: Advergames, Mobile Computing, Games Development.*

***ACM Classification Keywords**: D.1.m Miscellaneous.*

Introduction

Online games are the future of the interactive entertainment industry, seeing the convergence between the traditional media, and entertainment industry, and the gaming industry in an effort to develop new and sustainable business models and revenue streams in an increasingly online world. They move the gaming industry into a more functionally rich online environment from which the majority of the revenue stream will come -- an e-business environment. But moving to this new model presents a number of challenges to the games developers, the players, and the service providers who ultimately will need to support this new environment.

However, it also presents a number of exciting opportunities for new business models, new markets, and new growth. The main problem faced is a solution integration issue. The player wants to pay for online content with their existing channels, but they also want security and privacy. The developers need cross-platform integration and support for multiple services, channels, and providers. The service providers need to build reusable business function that is robust, efficient, and generic -- it should work for all business models, not just the gaming industries.

Online games come in many forms. Perhaps the most recognized are the highly visual, action-oriented pop-ups familiar to NYTimes.com users. They're primarily used by advertisers for branding purposes and are generally delivered via pop-ups and in various other ad formats on third-party sites. The objective is to attract traffic and acquire new customers.

Instant-win promotions and contests requiring some level of consumer participation are increasingly popular. Again, their purpose goes beyond branding into acquisition and building databases of customers and prospects. These games can take many forms, from a roulette-style wheel spun by the user to determine whether she's got a winning game card to an online drag race in which consumers challenge an automated car for a chance to win a related prize.

One marketer specializing in such online instant promotions had noteworthy results with incentive-based online games and contests.

It's hard to say, but most marketers' resistance to the game industry probably has to do with the pimply, geeky gamer stereotype. If we're not in an industry targeted to the teen market, we probably don't believe that games can have any impact on our marketing efforts. That's wrong, and more companies than ever out there are trying to change our collective marketing perceptions -- and convince us to start using games as advertising.

The concept's called advergaming, and regardless of its clunky moniker, it's a concept that has worked well for brands such as Nike (with its Nike Shox email game campaign), Ford (which used a racing game to promote its new Escape), and Pepsi. Companies such as YaYa, WildTangent, and The Groove Alliance have used stunning 3-D technology to create games that rival many commercial desktop and console games. Other companies, such as XI Interactive have brought together some of the finest minds in the gaming industry to create killer sports games. They combine single-player fun with innovative viral techniques that get consumers engaged with brands.

Even now, most of these games can be played over dial-up connections with middle-of-the-road computers. But with higher-speed computers and broadband connections becoming commonplace, these games are destined to become killer marketing vehicles for the future.

Simply put, games engage users for long periods of time, immersing them in an environment where they can develop an affinity for the brand. Rather than merely watching the action (as they do when viewing a sponsored sporting event on TV), advergame consumers actually become part of the action. Also, since the experience can be closely scripted in a near TV-like manner, the action can be interrupted to show TV-like commercials, or the views can be scripted to ensure advertiser messages are seen. It's a great combination of interactivity (for the user) and control (for the advertiser).

Some companies (such as Life Savers) create destination sites (check out Candystand) that host heavily brand-identified games. Others (such as Nike) have created effective viral campaigns in which users can play each other via email, inviting friends to beat their scores. Games have also been incorporated into banners and other rich media ad vehicles.

The market (and audience) for games is huge now and is going to continue to grow in the future as today's kids become tomorrow's sophisticated consumers. It's time to consider games as a viable marketing vehicle.

Advergames

The strict definition of an advergame is a Web-based computer game that incorporates advertising messages and images. However, we like to think of it as a tool that adds stickiness to your site as well as a little fun. Advergames allow you to market your product or brand subtly. Benefits of an Advergame are:

- Brand image reinforcement.
- Databases created from the advergame can be used for demographics research.
- Targeted markets can be reached by your advertising (when the game link is emailed).
- Visitors may spend more time on your site.
- Increased traffic due to viral marketing.

An advergame is not just for kids anymore - many surfers play advergames. These surfers include but are not limited to:

- 59\% of the boys ages 13 to 17 who go online.
- 62\% of the men ages 18 to 24 who go online.
- The largest group of women game players are between the ages of 45 and 54.

Like other advertising promotions not all advergames are equal. Here are some advices:

- Add a sweepstakes element to your game to provide an incentive for return visits or multiple plays.
- Include a viral component to encourage individuals to e-mail the game or its URL to friends to maximize word-of-mouth marketing.
- Consider a partnership with a company that can market your game to your desired audience. Simply building a great game and placing it on your website won't maximize impact.
- Make the game part of a larger media buy, such as by placing it within an interstitial and using buttons or banners on other sites to raise awareness and drive traffic.
- Take measures to reach your target audience during their leisure time. It's more likely they will play your game, and studies show people are more receptive to marketing messages when they are having fun

An advergame can help generate leads, build long-term brand awareness, and increase site stickiness as well as repeat web site visitors. It is a cost effective tactic that any sized e-business should at least check into for possible inclusion in marketing efforts.

Games in general are undeniably popular. Some of the more successful interactive games used for advertising purposes are parodies of the tried-and-true brand name games such as Jeopardy, Wheel Of Fortune, Who Wants To Be A Millionaire? and Roulette. In the end, what makes for a good interactive game for the player is a compelling, immersive experience.

What makes for a good advergame from the perspective of the advertiser footing the bill is the incorporation of a seamless data capture device. That data is then leveraged to build relationships. Even better are games in which in addition to data being captured, the player is learning about a product or service, its features and benefits, while playing the game.

An advergame can be used as part of an email campaign to a qualified, rented 3rd party opt-in list. The idea is to tease the recipient enough in the email to cause click-through to a special landing page. It is there at the landing page that the forging of an emotional bond with the prospect starts, and that the game is played.

In order to play, or to get to the exciting or more challenging levels of play, information is required of the prospect. At the very least, this includes name and email address. Some interactive games require additional data for each game level. The greater the prospect's desire to continue to play, the greater the likelihood the prospect will willingly provide additional data. It is critical that the prospect not be made to feel he has been brought to the game for the sole purpose of providing personal information. The prospect must feel he wants to be there of his own free will playing that game, or there will be no data at all forthcoming.

Another manner in which leads are generated is through a tell-a-friend viral device. Encouraging such forwarding can result in a viral factor, if you will, of 25\%. This means the amount of leads a game ultimately produces is that much greater than it would have been without the forwarding, or word-of-mouth.

It is the repetitive nature of games that drives up awareness of the sponsor. The same player, the frequency of the brand and/or the brand's message increases, plays each time the game. The greater the frequency, the greater the chances of being remembered.

Research on the specifics of advergaming's ability to build brand awareness does not yet exist. Common sense will tell you that brand awareness does, indeed, increase, but to what degree isn't the point. Awareness should not be sole nor primary objective of an advergaming campaign; it is more of an added benefit to the advertiser. Advergaming's strength is in immediate, and for however long of a shelf life the game may have, lead generation.

Each time the game is played by the same player is also a repeat visit to the site. This is where web site navigation, page design, merchandising techniques, etc. come in to play. A sharp marketer will employ every technique possible to move the player to and through product when the player has completed the gaming session.

The agency with whom an advertiser works to create, implement and manage an advergaming effort will have a mechanism set up to track users (players), qualified leads, viral rate, even viral posting of the game on other venues, if any. The agency will also report on the advertiser's cost per thousand, cost per lead, cost per minute engaged, average engagement time, click-through rate, and conversion rate.

An even more cost effective means of employing advergaming as a marketing tactic is the use of an off-the-shelf game. For this type of game, the agency is re-purposing a game that already exists especially for an advertiser. It is amortizing its fixed production costs over several clients, allowing it to charge a greatly reduced fee for its use. This scenario allows for the simple changing of skins to meet the needs of a particular advertiser, and perhaps a few other adjustments along the lines of incorporating a logo into the game.

The major advantage to a custom interactive game is in the opportunity to tie in product or service features and benefits - even news - to the game play. The obvious disadvantages are extended production time versus re-purposing, and additional cost.

Advergame Architecture

There is a broad range of personal computing devices in the market. Personal computers are the most extended ones, but nowadays there are turning out smaller devices such as handhelds and mobiles phones. These three devices (PCs, handheld and mobile phones) are in a convergence process since they all have a real operating system with graphical capabilities and can be connected to Internet. All of them have an {\it Internet} browser (HTML, WAP and XHTML) and also a Java virtual machine, even SMS services. This is the main reason advergames should be developed using the Java technology or a markup language in order to obtain a portable solution that can be run in any platform (Windows, Linux or Symbian). Maybe simple advergames can make use of WML technology or SMS connecting to a server. Advergames must be implemented in an efficient way mainly due to the limited resources in mobile phones. Connection capabilities of all devices include HTTP and sockets, so client-server and P2P solutions can be used.

A number of different technologies, listed below, are used for games on mobile devices. Some games are programmed to run natively on a phone's chipset, installed on the phone at the factory, and shipped with it. New embedded games cannot be installed by the consumer, and they are becoming less prevalent. Short Message Service is used to deliver short text messages from one phone to another. Users typically pay about 10 cents per message. SMS games are played by sending a message to a phone number that corresponds to the game provider's server, which receives the messages, performs some processing, and returns a message to the player with the results.

Just about every phone shipped since 1999 includes a Wireless Application Protocol (WAP) browser. WAP is, in essence, a static browsing medium, much like a vastly simplified form of the Web, optimized for the small form factors and low bandwidth of mobile phones. WAP games are played by going to the game provider's URL (usually through a link on the carrier's portal), downloading and viewing one or more pages, making a menu selection or entering text, submitting that data to the server, and then viewing more pages. Phones will continue to contain WAP browsers, and developers may find WAP useful to deliver more detailed help or rules to players than can be contained in a game application, since most games are still subject to strict memory limits.

Java 2 Micro Edition (J2ME) is a form of the Java language that is optimized for small devices such as mobile phones and PDAs. Nokia (and most other phone manufacturers) have made a strong commitment to Java phone deployment. Tens of millions of Java-enabled phones are already in consumers' hands. J2ME is limited by comparison to desktop Java, but it vastly improves the ability of mobile phones to support games. It allows far better control over interface than either SMS or WAP, allows sprite animation, and can connect over the air network to a remote server. Because of its capabilities and the widespread and growing deployment of Java-enabled phones, it is a natural for mobile game development today.

It is important to take into account the programming language when developing an advergame. This software must be installed in many different devices with different operating systems and of course it must not be recoded everytime a new device appears in the market. This is the main reason to chose JAVA as the programming language since most of the today devices have a Java Virtual Machine:

- Mobile Phones based on Symbian Operating System have an embedded virtual machine and
- those non based on Symbian also have one, both of them using the J2ME architecture.
- HandHeld devices can install a software tool such as Jeode from Insignia or J9 VM from IBM.
- Personal computers usually have a java runtime-environment that can be obtain from java.sun.com for the Microsoft Windows or Linux platforms.

An developped advergame with the J2ME with the MIDP profile technology can be run in a mobile phone, in a handheld device and in a personal computer. But it the advergame is developped with Standard Java maybe only just a few mobile phones will be able to run it, and of course handhelds devices and personal computers will be able to run it.

The MIDP APIs are logically composed of high-level and low-level APIs. The APIs are designed for applications or services where the handset functions as the client device. The user gains access to applications and services that run on the handset through a network service provider. The high-level APIs are designed for applications where software portability across a range of handsets is desired. This is important if you are writing an application or service that a network service provider plans to deploy to a selected set of handsets. To achieve this portability, the APIs use a high level of abstraction. The trade-off is that the high-level APIs limit the amount of control the developer has over the human interface look and feel. The underlying implementation of the user interface APIs, which is accomplished by the handset manufacturer, is responsable for adapting the human interface to the device hardware and native user interface style.

Among the most interesting capabilities of MIDP are the following ones:

- Persistent storage in the device.
- Graphical libraries.
- Connection via sockets or http.
- Portability amond different devices.

J2ME is not the only interpreted language deployed on phones, but it is an industry standard backed by many manufacturers and therefore offers a large and growing installed base. Some proprietary interpreted languages

have significant regional presence, including Qualcomm's Binary Runtime Environment for Wireless in North American and a standard called GVM supported by some Korean carriers. Games initially developed for the large J2ME installed base can be recoded in these proprietary languages if a sound business case presents itself.

A simple advergame can be also developped with the WML language. Only a WAP browser (Opera, EzWap, WinWap, Personal Internet Explorer, etc.) will be needed in onder to play it. The problem is that a connection to a server must be stablished and it is impossible to play off-line. Also WML has neither device storage capabilities nor graphics libraries in order to obtain good results.

MIDP 2.0 is backwards compatible with MIDP 1.0, hence it provides all functionality defined in the MIDP 1.0 specification. In addition it provides Over-The-Air (OTA) provisioning. This feature was left to OEMs to provide in the MIDP 1.0 specification.

An enhanced user interface has been defined, making applications more interactive and easier to use. Media support has been added through the Audio Building Block, giving developers the ability to add tones, tone sequences and WAV files even if the Mobile Media API optional package is not available. Game developers now have access to a Game API providing a standard foundation for building games. This API takes advantage of native device graphic capabilities. MIDP 2.0 adds support for HTTPS, datagram, sockets, server sockets and serial port communication. Push architecture is introduced in MIDP 2.0. This makes it possible to activate a MIDlet when the device receives information from a server. Hence, developers may develop event driven applications utilizing carrier networks. End-to-end security is provided through the HTTPS standard. The ability to set up secure connections is a leap forward for MIDP programming. A wide range of application models require encryption of data and may now utilize the security model of MIDP 2.0 based on open standards.

There are may connection methods when accessing a server or client depending on the device or the programming language when running an advergame. It is clear that some information needs to be sent among players or among clients-servers in order to accomplish the advergame, keep information tracking of the users, and the flow advertisment control.

Short messages or multimedia ones are a useful tool when sending information, only telephony services are neccesary. Mobile telephones, PDA's with expansion cards or personal computers can be used to send messages. One of the main disanvantages is that a non real-time game can be developed, but it is a good method to keep track information about the users. When a message reaches the game server some infomation (telephone number, provider, user status, etc.) can be updated in a database in order to send push messages. Typical advergames that use SMS technology are:

- Test questions, when there are enough messages in the server a prize can be randomly delivered to any mobile phone number.
- Last minute offers, when there are some tickets (cinema, football, theater, etc.) that are not sold they can be delivered to mobile phones with a previous registration or asking the server with a SMS.

Only a SMS server and a agreement with an SMS provider is needed to implement this service. There are many companies that are actually offering such service. Even a Java implementation could be coded but it is needed an special agreement with the SMS provider in order to deliver all the messages to a given IP. SMS is not a particularly good technology for games, because it is dependent on text entry by the user, and thus is, in essence, a command-line environment. It is also expensive for a game of any depth, since a mere 10 exchanges with the server will cost a user 1 dollar or more. Although the deployment of Multimedia Message Service (MMS) technology makes message-based games more appealing, this is still not a great gameplay environment.

A more complex service than SMS can be implemented if devices supports HTTP connections. All new devices are capable of such services. This is an approach similar to WWW services in personal computers. A server is necessary in order to control all the information since the IP address is usually dinamic. Java Servlets technology is an usefull tool to implement such services.

This service can use WML and WMLScript to connect to a server or use J2ME to stablish HTTP connections. In any case schema is similar to normal web pages. A real-time advergame can be achieved with this method since it does not depends on the message delivering just as in the previous section. The information is sent in real-time. This is a flexible method since only a browser is needed (or a JVM). The main advantage is the graphical user interface. But no P2P communication can be carried out. Either version of WAP offers a friendlier interface than

SMS, and is generally less expensive for consumers who pay for airtime only, rather than by the message. But it is a static browsing medium; little or no processing can be done on the phone itself, and all gameplay must be over the network, with all processing performed by a remote server.

Socket use gives J2ME developers the flexibility to develop all kinds of network applications for wireless devices. However, not every wireless manufacturer supports socket communication in MIDP devices, which means that wireless applications developed using sockets could be limited to certain wireless devices and are less likely to be portable across different types of wireless networks. To use a socket, the sender and receiver that are communicating must first establish a connection between their sockets. One will be listening for a request for a connection, and the other will be asking for a connection. Once two sockets have been connected, they may be used for transmitting data in either direction. All today devices are using an embedded JVM, tipically supporting J2ME version 2.0 (with sockets). Main advantages of this implementation are:

- Sockets management.
- 2D and 3D graphical {\it APIs}.
- Persinstent storage on the client.

The use of sockets is usefull when dealing P2P services. The only problem is that the IP address is dinamically assigned to the client so a server is needed. This P2P service needs to send some information to an advergame server to keep track information of the players. This solution is the best one since with sockets all previous schema can be implemented.

Conclusions

Advergame is a new marketing concept that brings users a way to interact with others and also to participate in quizs. User information can be update in a database in order to send push messages or do mailing while the client in playing some game. Some advertisements can appear in the game or even play with advertisements. The user can win prizes to keep the his attention.

This paper has presented some tecnologies that can be used to develop an advergame. Java services are the best solution since it is a portable solution and all today devices have an embedded virtual machine.

Bibliography

[1] Blockdot (2001). Advergaming 101. Available online: http://www.blockdot.com/advergaming/stats.cfm.

[2] Chen, J., Ringel, M. (2001). Can advergaming be the future of interactive advertising? Fast Forward. Available online: http://www.kpe.com.

[3] March, T. (2001, Spring). How to bag the elusive human attention span. Digitrends. Available online:

http://www.digitrends.net/marketing/13639_16525.html.

[4] Pintak, L. (2001, May 23). It's not only a game: Advergaming set to become a billion dollar industry. Available online: http://www.turboads.com/richmedia news/2001rmn/rmn20010523.shtml.

[5] Rodgers, A. L. (2002, January). Game theory. Available online: http://www.fastcompany.com/build/build feature/yaya.html.

[6] YaYa (2002a). Why games? Available online: http://www.yaya.com/why/index why.html.

[7] YaYa (2002b). YaYa creates viral Internet games that build brands and drive revenue. YaYa online press kit. Available online: http://reports.yaya.com/presskit.pdf.

Authors' Information

Eugenio Santos Menéndez. Dpto. Organización y Estructura de la Información. Escuela Universitaria de Informática de la Universidad Politécnica de Madrid, Ctra. Valencia, km. 7, 28031 Madrid (Spain); e-mail: esantos@eui.upm.es

Rafael Gonzalo Molina – Dpto. Inteligencia Artificial. Facultad de Informática de la Universidad Politécnica de Madrid; Boadilla del Monte, Madrid (Spain); e-mail: rgonzalo@fi.upm.es

Francisco Gisbert - Dpto. Lenguajes, Sistemas e Ingeniería del Software. Facultad de Informática de la Universidad Politécnica de Madrid; Boadilla del Monte, Madrid (Spain); e-mail: fgisbert@fi.upm.es

COGNITIVE BIAS IN KNOWLEDGE ENGINEERING COURSE

Tatiana Gavrilova, Seppo Puuronen

***Abstract**: The paper presents experience in teaching of knowledge and ontological engineering. The teaching framework is targeted on the development of cognitive skills that will allow facilitating the process of knowledge elicitation, structuring and ontology development for scaffolding students' research. The structuring procedure is the kernel of ontological engineering. The 5-steps ontology designing process is described. Special stress is put on "beautification" principles of ontology creating. The academic curriculum includes interactive game-format training of lateral thinking, interpersonal cognitive intellect and visual mind mapping techniques.*

***Keywords**: ontologies, knowledge engineering, analytical skills.*

***ACM Classification Keywords**: I.2 Artificial intelligence*

Introduction

Young researchers of computer science and information technologies use a lot of informal rules-of-thumb advice that may help but not a systematic guidelines. Major recommendations deal with library work, citation and other important but not essential components of research activity. Web resources now help to streamline the process of finding, selecting, and entering information from the Web, corporate databases, and reference materials into a paper. We are speaking about not syntax but semantic of study. Our course is aimed at information engineering – structuring and shaping the research framework. This issue has hardly been explored yet.

Students are engaged in various semantically-based activities. Indeed, any conceptual data modeling is a form of informal Semantic Modeling [26]. They do all the "bottleneck" activity including data and knowledge elicitation, structuring, formalizing. But do they do it professionally? They need knowledge engineers' analytical skills.

Knowledge Engineering traditionally emphasized and rapidly developed a range of techniques and tools including knowledge acquisition, conceptual structuring and representation models [1], [24]. But for practitioners it is still a rather new, eclectic domain that draws upon areas like cognitive science. Accordingly, knowledge engineering has been, and still is, in danger from fragmentation, incoherence and superficiality. Still few universities deliver courses in practical knowledge engineering.

This paper describes recent experience in teaching such course. The total number of students that were taught is more than 160 – in Finland, Russian and USA.

Teaching Framework

Knowledge Engineering course (KEC) is based on university courses in intelligent-systems development, cognitive sciences, user modeling, and human-computer interaction delivered by authors in 1988-2005. KEC proposes information structuring multi-disciplinary methodology, including the principles, practices, issues, methods, techniques involved with the knowledge elicitation, structuring and formalizing. Emphasis is put not on technologies and tools, but in training of analytical skills. KEC also includes Ontological Engineering that is further development of knowledge engineering towards ontology design and creating.

Students are introduced to major issues in the field and to the role of the knowledge analyst in strategic information system development. Attention is given both to developing inter-personal information communication skills and analytical cognitive creative abilities.

The class feature short lectures, discussions, tests, quizzes and exercises. Lectures are important but the emphasis is put on learning through discussions, simulation, special games, training and case studies. A good deal of the course focuses on auto-reflection and auto-formalizing of knowledge, training of analytical and communicative abilities, discovery, creativity, cognitive styles features, and gaining new insights.

KEC consists of 4 inter-connected modules:

- Getting Started in KE (12 hours),
- Practical KE in depth (12 hours),

- Ontological Engineering (12 hours),
- Business Processes Modelling (12 hours).

Different combination of sub-topics is possible. Fig.1 illustrates the structure.

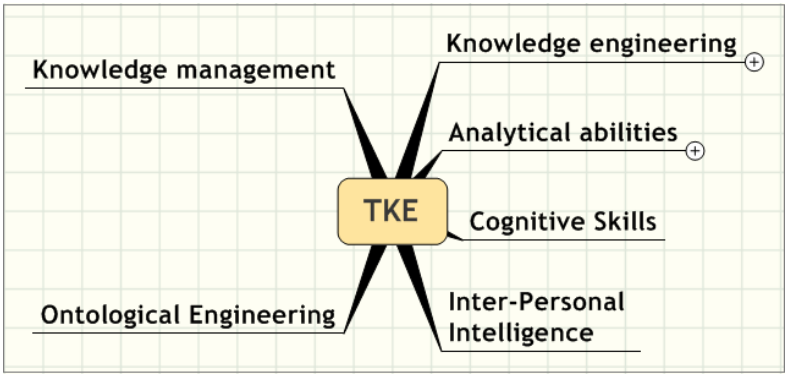


Fig.1. Outline of training on knowledge engineering

The main difference of KEC to existing curricula is cognitive (not technological) bias. Fig.2. shows the main issues covered by tests and practical exercises. Students of IT-departments often under-value the significance of psychological background of categorization, laddering and lateral thinking. But during learning process some of them feel "insight" and become very enthusiastic.

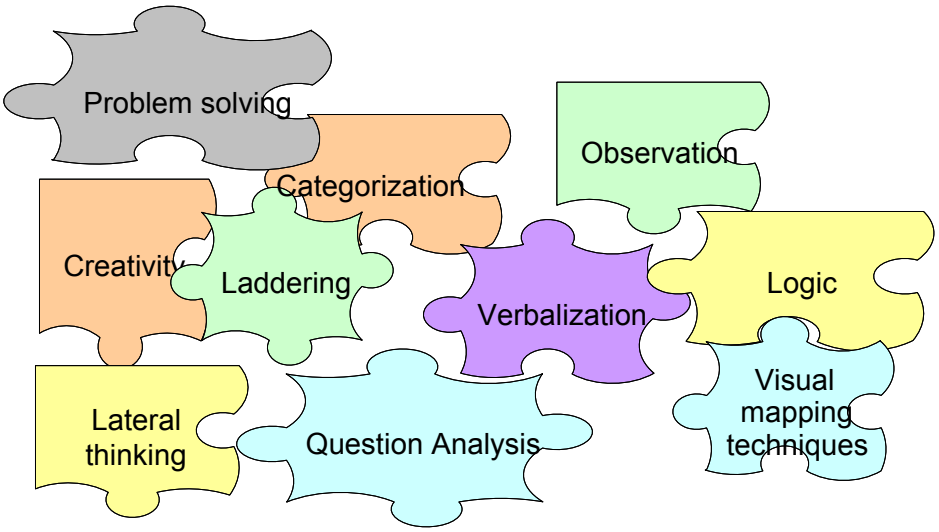


Fig.2. The main topics of practical exercises

The practical methods of knowledge elicitation gain the main interest during teaching as they really work. The training is organized that students study main techniques in pairs "expert – knowledge analysts" with a shift of roles when needed. Some psychological assessment techniques help students to realize their strong and week points in inter-personal communication and intelligence.

But only knowledge structuring exercises show the importance of obtained analytical skills for the students. Even simple tests from their own research domains are rather difficult at the first classes.

The study is aimed on semantics not syntax of knowledge engineering. We suppose that systems and languages may be self-studied while general scope and knowledge-stressed approach should be trained thoroughly. Students often under-estimate the role of cognitive styles, verbal skills and logics in information processing. It is supposed to be common sense while it needs to be taught.

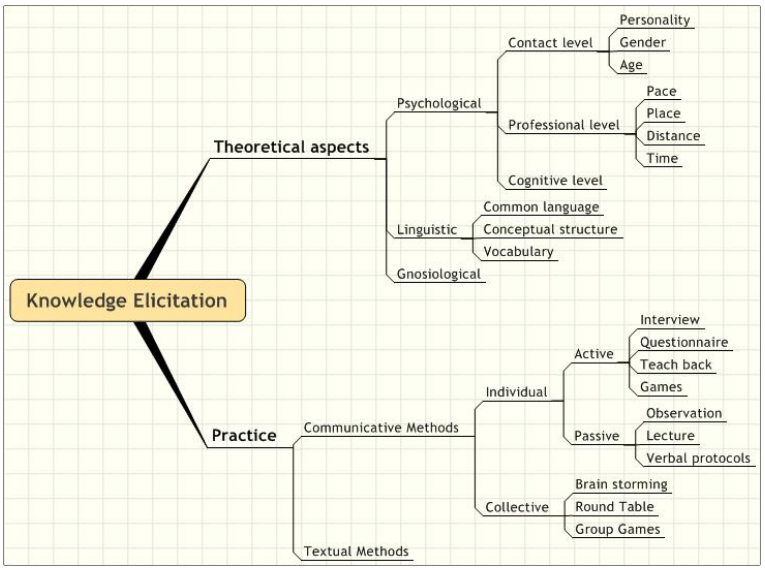


Fig.3. Theory and practice of knowledge elicitation

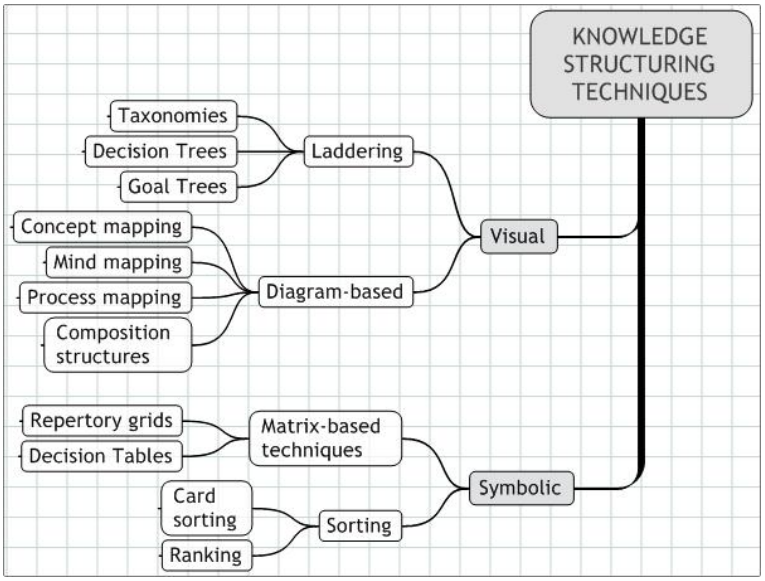


Fig.4. Knowledge structuring techniques

Stress on Ontological Engineering

Ontologies are fashionable now. But our experience in training show that nobody can deal with ontologies without knowledge engineering practice. How to teach ontology design? The theory differs from practical needs…

There are numerous well-known definitions of this milestone term (Gruber, 1993; Guarino and Giaretta, 1998; Jasper and Uschold, 1999; Mizogushi and Bourdeau, 2000; Neches, 1991) but they may be generalized as "Ontology is a hierarchically structured set of terms for describing an arbitrary domain" [9]. In other words "ontologies are nothing but making knowledge explicit" [12].

Since 2000 a major interest of researchers focuses on building customized tools that aid in the process of knowledge capture and structuring. This new generation of tools – such as Protégé, OntoEdit, and OilEd - is concerned with visual knowledge mapping that facilitates knowledge sharing and reuse [18], [19], [22]. The problem of iconic representation has been partially solved by developing knowledge repositories and ontology servers where reusable static domain knowledge is stored. Ontolingua, and Ontobroker are examples of such projects [20], [21].

Ontology creating also faces the knowledge acquisition bottleneck problem. The ontology developer encounters the additional problem of not having any sufficiently tested and generalized methodologies, which would recommend what activities to perform and at what stage of the ontology development process. The lack of structured guidelines and methods hinders the development of shared and consensual ontologies within and between the specialists. Moreover, it makes the extension of a given ontology by others, its reuse in other ontologies, and final applications difficult [11].

Until now, only few effective domain-independent methodological approaches have been reported for building ontologies [4]; [25], [16]. What they have in common is that they start from the identification of the purpose of the ontology and the needs for the domain knowledge acquisition. However, having acquired a significant amount of knowledge, major researchers propose a formal language expressing the idea as a set of intermediate representations and then generating the ontology using translators. These representations bridge the gap between how people see a domain and the languages in which ontologies are formalized. The conceptual models are implicit in the implementation codes. A re-engineering process is usually required to make the conceptual models explicit. Ontological commitments and design criteria are implicit in the ontology code.

Figure 3 presents our vision of the mainstream state-of-the-art categorization in ontological engineering [12], [13], [28] and may help students and analyst to figure out what type of ontology he/she really needs. We use Mindmanager™ and Cmap™ as they proved to be rather powerful visual tools.

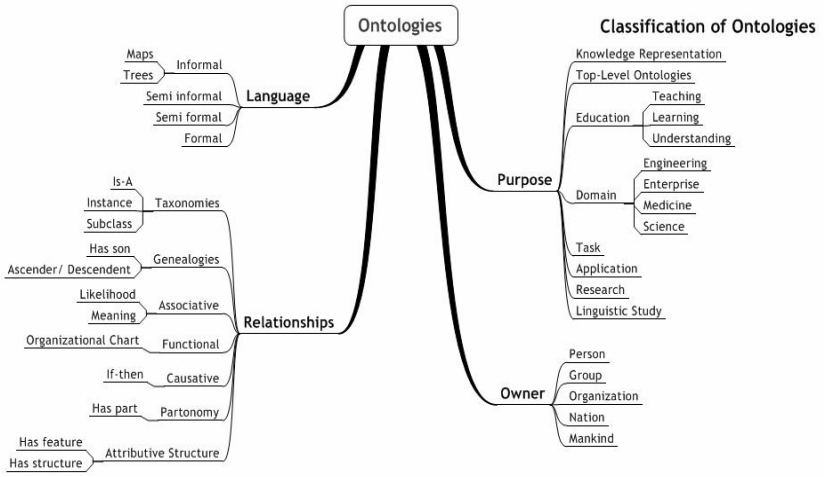


Fig.5. Ontology mind map

We try to simplify a bunch of different approaches, terms and notations for practical use (Fig. 5) and even dare to propose a 5-step recipe which help practical ontology design.

Exercises during training help us to evaluate and update it unless it starts to work.

Ontology Design Recipe

The existing methodologies describing ontology life cycle [28], [16], [9] deal with general phases and sometimes don't discover the design process in details. Five simple practical steps were proposed.

Step 1. **Glossary development:** The first step should be devoted to gathering all the information relevant to the described domain. The main goal of this step is selecting and verbalizing all the essential objects and concepts in the domain.

Step 2. **Laddering:** Having all the essential objects and concepts of the domain in hand, the next step is to define the main levels of abstraction. It is also important to elucidate the type of ontology according to Figure 1 classification, such as taxonomy, partonomy, and genealogy. This is being done at this step since it affects the next stages of the design. Consequently, the high level hierarchies among the concepts should be revealed and the hierarchy should be represented visually on the defined levels.

Step3. **Disintegration:** The main goal of this step is breaking high level concepts, built in the previous step, into a set of detailed ones where it is needed. This could be done via a top-down strategy trying to break the high level concept from the root of previously built hierarchy.

Step4. **Categorization:** At this stage, detailed concepts are revealed in a structured hierarchy and the main goal at this stage is generalization via bottom-up structuring strategy. This could be done by associating similar concepts to create meta-concepts from leaves of the aforementioned hierarchy.

Step 5. **Refinement**: The final step is devoted to updating the visual structure by excluding the excessiveness, synonymy, and contradictions. As mentioned before, the main goal of the final step is try to create a beautiful ontology. We believe what makes ontology beautiful is harmony.

Using these tips the students develop several huge practical ontologies to conduct systemically more structured research. This approach is based on developing of a set of ontologies:

- Problem-ontology definition (ontology N1 describing main concepts)
- Ontology of reviewed approaches (ontology N2 describing the history of the problem)
- Experiment framework design (ontology N3 presenting experimental conception)
- Data structure ontology (ontology N4 presenting input and output data)t
- Mathematical modelling and main results ontology design (ontology N5 describing results)

Not all five ontologies are obligatory, but even an attempt to create them is a first step to perform systemic scientific study.

Conclusion

Students and teachers both are knowledge workers. So students enter “the world of ontologies” with interest and begin to use it in their practical research work.

Our experience in training of knowledge analysts and teaching this course in the period of 1999-2006 confirm the unique role of knowledge structuring for developing ontologies quickly, efficiently and effectively. We follow David Jonassen’s idea of “using concept maps as a mind tool” [14]. The use of visual paradigm enables students to process and understand greater volume of information.

The course is double-ontological as the development of educational knowledge structures in the form of ontologies provides training and learning support. Teaching ontologies scaffold and improve students’ understanding of the courseware and later help to realize substantive and syntactic company knowledge. As such, they can play a part in the overall pattern of learning, facilitating for example analysis, comparison, generalization and transferability of understanding to analogous problems.

Ontological framework scaffolds the student’s research activity. But ontological engineering is rather easy for «old» sciences with good structure. Researchers in new, multi-disciplinary and ill-structured disciplines as HCI, cognitive sconces, management, etc. will face a bunch of difficulties in design and development phases. Ontologies also are rather subjective.

Our paper presents one of the first attempts to show the visionary role of knowledge engineering in helping student research. Ontologies are good for better self-understanding of research and then for knowledge sharing.

We also have experienced to teach the modification of this course to the practitioners. It was in-service training of analysts for IT-departments of some companies. The training was a success, as knowledge engineering is a unique set of methods that help everywhere. It is a real ‘silver bullet’.

Acknowledgements

This work is partly supported by the grant of Russian Foundation for Basic Research No.07-01-00053 and grant of Academy of Finland.

Bibliography

1. Adeli, H. (1994) Knowledge Engineering. McGraw-Hill, New-York
2. Boose, J.H. (1990) Knowledge Acquisition Tools, Methods and Mediating Representations. In Knowledge Acquisition for Knowledge-Based Systems (Motoda, H. et al., Eds), IOS Press, Ohinsha Ltd., Tokyo, pp.123-168.
3. Eisenstadt, M., Domingue, J., Rajan, T. & Motta, E. (1990) Visual Knowledge Engineering. In IEEE Transactions on Software Engineering, Vol.16, No.10, pp.1164-1177.
4. Fensel, D. (2001) Ontologies: A Silver Bullet foe Knowledge Management and Electronic Commerce. Springer.

5. Gavrilova, T., Voinov, A. (1998) Work in Progress: Visual Specification of Knowledge Bases // Lecture Notes in Artificial Intelligence 1416 “Tasks and Methods in Applied Artificial Intelligence”, A.P.del Pobil, J.Mira, M.Ali (Eds), Springer, pp. 717-726.
6. Gavrilova, T.A., Voinov, A., Vasilyeva E. (1999) Visual Knowledge Engineering as a Cognitive Tool / Proc. of Int. Conf. on Artificial and Natural Networks IWANN’99, Spain, Benicassim. - pp.123-128.
7. Gavrilova, T. (2003) Teaching via Using Ontological Engineering // Proceedings of XI Int. Conf. “Powerful ICT for Teaching and Learning” PEG-2003, St.Petersburg, p. 23-26.
8. Gavrilova, T., Kurochkin M., Veremiev V. (2004) Teaching Strategies and Ontologies for E-learning // Int. Journal “Information Theories and Applications”, vol.11, N1, pp.35-42.
9. Gómez-Pérez, A., Fernández-López, M., Corcho, O. (2004) Ontological Engineering with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web, Springer.
10. Gruber, T. (1993) A translation approach to portable ontology specifications. Knowledge Acquisition , Vol. 5, pp. 199-220.
11. Guarino, N. & Giaretta, P. (1998) Ontologies and Knowledge Bases: Towards a Terminological Clarification. In Towards Very Large Knowledge Bases: Knowledge Building & Knowledge Sharing, IOS Press, pp.25- 32.
12. Guarino, N., Welty, C. (2000) A Formal Ontology of Properties. In R. Dieng and O. Corby (eds.), Knowledge Engineering and Knowledge Management: Methods, Models and Tools. 12th International Conference, EKAW2000. Springer Verlag, pp. 97-112.
13. Jasper, R. and Uschold, M (1999). A Framework for Understanding and Classifying Ontology Applications. In 12th Workshop on Knowledge Acquisition Modeling and Management KAW'99.
14. Jonassen, D.H. (1998) Designing constructivist learning environments. In Instructional design models and strategies (Reigeluth, C.M. (Ed), 2nd ed., Lawrence Eribaum, Mahwah, NJ.
15. Miller, G. (1956) The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Proccessing Information. The Psychological Review, vol. 63, pp. 81-97.
16. Mizogushi, R. and Bourdeau J. (2000), Using Ontological Engineering to Overcome Common AI-ED Problems. International Journal of Artificial Intelligence in Education, v. 11, pp.1-12.
17. Neches, et al , (1991) Enabling Technology for Knowledge Sharing. AI Magazin, Winter, pp.36- 56.
18. OilEd, Bechhofer, S. and Ng G. Accessed from http://oiled.man.ac.uk/ at December 7, 2004.
19. OntoEdit, AIFB, University Karlsruhe Accessed from http://www.ontoknowledge.org/tools/ at December 07, 2004
20. OntoBroker, Accessed from http://ontobroker.aifb.uni-karlsruhe.de/ at December 7, 2004.
21. Ontolingua, Stanford University. Accessed from http:// www.ksl.stanford.edu/ software/ontolingua/ at December 7, 2004.
22. Protégé, Stanford Medical Informatics. Accessed from http://protege.stanford.edu/ at December 07, 2004.
23. Sowa, J. F. (1984) Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley, Reading, Massachusetts.
24. Scott, A., Clayton, J.E. & Gibson E.L. (1994) A Practical Guide to Knowledge Acquisition, Addison-Wesley.
25. Swartout, B., Patil, R., Knight, K. & Russ, T. (1997) Toward Distributed Use of Large-Scale Ontologies. In Ontological Engineering, AAAI- 97 Spring Symposium Series, pp.138- 148.
26. The CIO’s Guide to Semantics, © 2004 Semantic Arts, Inc. www.semantic-conference.com
27. Tu, S., Eriksson, H., Gennari, J., Shahar, Y. & Musen M. (1995) Ontology-Based Configuration of Problem-Solving Methods and Generation of Knowledge-Acquisition Tools. In “Artificial Intelligence in Medicine”, N7, pp.257-289.
28. Uschold, M., Gruninger M (1996). "Ontologies: Principles Methods and Applications", Knowledge Engineering Review,vol1, N1.
29. Wielinga, B., Schreiber, G. & Breuker J. (1992) A Modelling Approach to Knowledge Engineering. In Knowledge Acquisition, 4 (1), Special Issue, pp.23-39.
30. Werthheimer, M. (1959) Productive Thinking, HarperCollins.

Authors' Information

Tatiana Gavrilova – Professor Saint-Petersburg State Polytechnic University, Politechnicheskaya 29, 195251, St. Petersburg, Russia, gavr_csa@rambler.ru

Seppo Puuronen – Professor of University of Jyväskylä, P.O. Box 35, FIN-40014, Jyväskylänyliopisto, Finland, sepi@cs.jyu.fi

AUTHORS SUPPORT IN THE TM4L ENVIRONMENT

Darina Dicheva, Christo Dichev

***Abstract**: The TM4L environment enables the development and use of ontology-aware courseware based on the Semantic Web technology Topic Maps. In this paper we discuss its features in the light of authoring support, giving illustrative examples to highlight its use.*

***Keywords**: Ontologies, digital libraries, courseware, Topic Maps.*

***ACM Classification Keywords**: H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia – navigation.*

Introduction

Current Web-based educational practices indicate that courseware authors' ability to gather and generate information exceeds their ability to organize, manage, and effectively use it. Ontologies are a key technology emerging to facilitate Web information processing by supporting semantic structuring, annotation, indexing, and search. Ontologies allow organization of learning material around components of semantically annotated topics. This enables ontology-based educational systems to do efficient semantic querying combined with structured intuitive navigation and access to the learning resources. However ontology development process is costly, which should be compensated by reusing the work. We have developed an authoring tool, the TM4L Editor [1], which enables the development of ontology-aware courseware based on the emerging technology Topic Maps [2]. To our knowledge TM4L is the only general educational Topic Maps Editor presently available. It is free software available at http://compsci.wssu.edu/iis/nsdl/ download.html (for the period May 2005 - February 2006 it had 3,259 downloads.)

In the Topic Map (TM) paradigm an ontology is an accurate description of the essential entities and relations which are found in the modeled subject area, and can be represented as a set of *topics* linked by *associations*. Therefore the Topic Maps technology is well suited for structuring learning material around subject ontologies.

As part of the TM4L Editor evaluation we conducted a study to find out what are the major difficulties that authors of educational topic maps face. In this paper we propose strategies for overcoming some of the problems in educational topic maps authoring and discuss their use for re-designing TM4L.

Support for Ontological Classifications

It is tempting to think that the classification schemes utilized by the libraries can be adapted in a simple way for use in e-learning systems. However, the neutral and objective classification used in the libraries might be in conflict with the user centric considerations in an e-learning environment. In such an environment the classification structure should account for multiple perspectives depending on the context in which the information is being sought. For a conceptual structure to be interoperable and resources to be sharable between different groups of learners, this contextual information must be explicit. An implied goal in re-designing TM4L was to produce an authoring environment enabling the representation of different *perspectives* of a domain instead of just capturing and representing an objective reality.

An important requirement was derived from the observation that a *practically useful* educational topic map cannot solely rely on a single standardized classification. A shareable conceptual structure of an educational topic map normally can not be derived from a particular textbook or a course syllabus. Besides the fact that textbooks and courses (modules) are changed frequently, their structures are often founded on a subjective categorization. The resulting topical structure needs also to conceptualize the instructor's view of the learning collection and the knowledge level of learners for which the classification is designed and constructed. To insure better support for educational TM authors, TM4L allows them to incorporate domain structures and course structures. Since the act of classification is inherently contextual, TM4L supports incorporating *perspectives* in classification construction.

On the other hand a shared objective classification scheme can function as a means for communication within and among domains by standardizing a vocabulary and meanings in the domains. A classification is objective when the criteria used for classifying are not subject to the whim of the person doing it. The lack of a shared

understanding and consistency in using concepts on a textbook and course level might be compensated by using objective domain conceptual structures (domain ontologies). The purpose of ontologies is not to replace the contextual viewpoints but to coexist with them complementing and enhancing the informational support. Intuitively ontological classification enables us to organize a set of entities into groups, based on their ultimate nature and existing relations. From this viewpoint we conceive domain ontologies as a *conceptual reference system*, with a collection of concepts, relations between concepts and classification hierarchies. The resulting conceptual schema could serve as an aid for integrating related concepts (terms) from different repositories.

TM4L provides support for authors that want to create ontological, objective classification: they can create light-weight domain ontologies represented as topic maps. Differently from typical ontology editors, our model includes three basic concept hierarchies: “superclass-subclass”, “whole-part” and “class-instance”. In this way authors are able to create more easily expressive conceptual structures that include various classifications of certain concepts. Authors can create typed concepts (with variant names) and an arbitrary number of resources linked to them. They can add any number of relations including transitive and symmetric relations in conjunction with the predefined classification hierarchies. By limiting the predefined relations and meta-level concepts we wanted to avoid unintended deviation between the learner's expectations and the produced output. The degree of formality of ontology is not necessarily a measure of quality in an e-learning environment: the focus should be on capturing the intended use of the ontology.

However, there are some challenges involved in the domain ontology development process: it is difficult and costly. Arriving at a representation of a domain requires deep knowledge of that domain and involves identifying its boundaries, selecting which concepts to define and at what level of detail and deciding how these concepts should be related. For example, programming languages are often classified between being interpreted or compiled. However these parameters do not reveal the essence of programming languages; neither semantics nor syntax depends significantly on whether code is compiled or interpreted.

Further on, concepts should account for multiple perspectives depending on the context in which the ontology is being used. All these assume support of a sufficient range of operations in the ontology development process, such as ontology design, implementation and browsing, merging of ontologies, searching for resources, multiple perspectives, etc.

3. Reusing Ontologies

One way to minimize the cost of creating ontologies is to make them reusable. The domain ontology component of a TM whose development is costly is *more stable* and therefore *reusable*. It is stable because classification based on a domain ontology is objective. The fact that the Procedural Languages are subclass of the Imperative Languages and that C is an instance of the Procedural languages is independent of human judgment or interpretations. Our framework supports reuse not only of learning resources but also of domain knowledge and instructional knowledge.

Figure 1. The Java Portal: exploring the Java ontology.

Figure 2. The Java Portal: exploring the Java course handouts.

We can view the domain ontology conceptual structures as independent information resources in their own right. An existing conceptual structure can then be overlaid on different information pools (using the TM4L editing

functionality), or merged with other topic maps (using its merge functionality) capturing different viewpoints on the same collection. An existing structure can be a source for new perspectives or used for interchange.

TM4L supports ontology reuse which is essential for effective information integration. In addition to promoting common understanding of concepts and avoiding redefinition of terms it allows adding new meaning to preexisting course structures as it is illustrated below.

TM4L supports the extension of conceptual strictures through merging. For example, an existing conceptual stricture (ontology) implemented as a topic map can be integrated with a course representing topic map by merging them with additional mapping when necessary. We used this approach to create the topic map, driving the Java Portal (see Fig. 1 and 2): by merging the separately created Java ontology and a particular course structure (developed for the CSC 1311 Programming 1; using Tony Gaddis' book 'Starting Out with Java').

4. "One Ontology-Many Courses" Model

The conceptual representation of the domain is a key factor in our framework, which is built on the assumption that a rational structure will facilitate a correct representation of a collection of learning resources as well as the process of seeking them. The intuition supporting this assumption is that a conceptual representation can facilitate prediction and thus the chances of success of a specific navigational strategy.

We argue that a single standard classification of learning materials in a subject area is a weak strategy from a practical perspective. In general, a classification is simply one particular rationalization of the relationships in a given domain that satisfies a group of individuals at a certain point in time. Any classification, no matter how it is constructed, represents one possible view on a domain and thus one potential way to organize a learning collection. Any field of knowledge may be classified from different viewpoints or contexts resulting into different classifications. For each particular task the authors choose to represent one particular view of the domain, therefore, a classification of a knowledge area is aimed to support a given viewpoint at the expense of other views.

We can think of two basic perspectives on a course Topic Map in a particular subject domain: a *structure of concepts* describing the subject area and a *collection of text units* (chapters, subchapters, sections, segments) storing the corresponding information in a specific format. A common mistake is to confuse the container for the concept(s) contained. Figure 3 illustrates these two viewpoints on the Prolog domain. The first one is a conceptual classification; the second one represents Prolog from a course structure perspective. The suffix LN attached to each term stands for Lecture Notes.

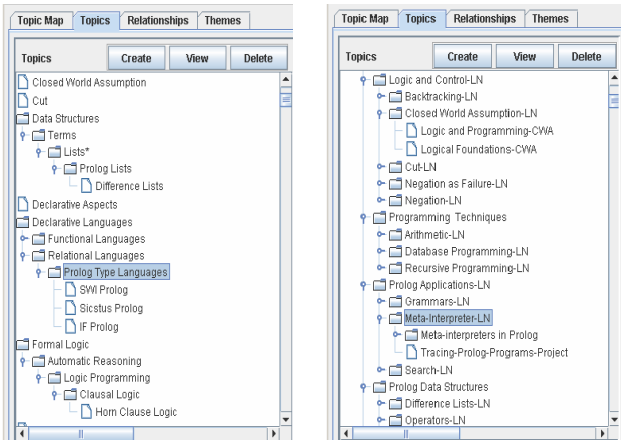


Figure 3. Two perspectives on Prolog learning resources.

It would be difficult to argue that only one of these classifications is an accurate representation of the domain and the other is not. The selection of categories in a classification depends on the viewpoint/context of the community of individuals in which the classification is created and used: in our case instructors and learners. In this case a classification based on a modules/lessons structure is not less significant than its conceptual counterpart.

The second classification supports learners that like to explore resources clarifying particular fragments from the studied material (lecture), the first one supports learners that like to explore resources related to individual concepts in the domain of their study. Therefore we found it beneficial to incorporate both of these classifications in the educational Topic Maps that we are developing, to serve as bases of different views.

This approach however posed a new problem. By extending the domain ontology conceptual structure with the specific course structure the resulting topic map turned to be course specific. This of course is highly undesirable with regard to the topic map reusability in a different course in the same subject area, for example, a course using a different textbook or instructor's handouts, etc. Hence we choose to create two independent topic maps – an

ontology topic map and a *course structure topic map*. This separation of the content structure into "public" and "private" along with an efficient support for merge is the basis of our proposed model of educational topic maps.

To create a practical information support for users one needs to know the community's vocabulary since the classification must reflect and respond to this particular community. This consideration shifted the focus of our work on enhancing TM4L from the system's issues and techniques to the domains and contexts in which classifications function.

5. Building TM-driven Course Portals

Topic maps provide a suitable model for building portals and other forms of Web-based information delivery. The function of portals to provide a single point of access to a broad array of resources and a rich navigational structure are a fundamental characteristic of topic maps. Topic map-driven Web portals create dynamically the frame structures and content from the underlying topic map.

For displaying/using our developed educational topic maps it was feasible and convenient to relate the portal interface to the course structure components. However, the short life-cycle of course structures is quite common in the learning domain. In addition, we wanted to reuse the portal interface for a number of courses. Therefore the conventional methods for portal constructions were not suitable.

To overcome these problems, in TM4L we devised a generic topic map-driven portal and a tool for *automatic generation of portals* from it. The generic portal provides the interface, the topic filtering functionality, and the topic map-directed browsing functionality. The latter supports two different patterns of browsing based on the structure of the underlying topic map. By providing a specific description of the desired presentation categories (indexes) and a specific topic map (an XTM file) to the tool, the author may generate a specific portal from the generic model. Different portal versions may be generated by specifying different categories that should be explicated by the portal, such as course units, glossary, examples, tests, online resources, etc. The specified categories must be resource or topic types in the corresponding topic map. Figures 1 and 2 are screenshots of such a portal – the Java Portal (P4J), available at http://iiscs.wssu.edu/p4j/.

The resources in P4J are linked to course topics that are organized in a hierarchy, and there is an "uplink" to the much richer conceptual level. Therefore the learning resources can be viewed in the context of the related topics along with the relevant conceptual components of the learning collection. In a similar manner topics can be inspected in a drill down fashion. This means that the learner always sees the concepts to which the inspected resource is linked to and can always switch back and forth between the resource and concept domains. The P4J portal navigation is context-driven and view-driven. The learner can inspect the learning collection from a selected perspective (domain, course structure, etc.) or choose to examine a contextually related topic (e.g. to switch from "Lists" to the related topic "Recursion"). This kind of contextual visualization gives the learner a feeling of being in a "relational space".

6. Conclusion

In an ideal world there would be a single conceptual structure for describing each subject area and a single standard for structuring course resources. However the reality of e-learning combined with the legacy of traditional learning is a mixture of different conceptualizations, viewpoints and subjective logic. Rather than preaching how to reach the absolute consistency, we address the reality by providing means to divide the learning resource structure into "public" and "private", that is, into durable and unstable. While many works have reported on particular conceptual frameworks and implementations, the impact and practical implications of various aspects of the traditional learning support on e-learning tools have not been discussed in great detail. In this paper we outline the evolution of the TM4L Environment based on some practical observations. The focus is on the difficulties authors face when creating ontology-based courseware with TM4L, on the lessons learned and on the proposed strategies for overcoming some of the problems. Many of our findings may pertain to other ontology-based editing environments.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. DUE-0442702 "CCLI-EMD: Topic Maps-based courseware to Support Undergraduate Computer Science Courses."

References

1. Dicheva, D. & Dichev, C. (2006) TM4L: Creating and Browsing Educational Topic Maps, British Journal of Educational Technology, 37(3) (in press).
2. Park, J. & Hunting, S. (2002) XML Topic Maps: Creating and Using Topic Maps for the Web, Addison-Wesley.

Authors' Information

Darina Dicheva – Winston-Salem State University, N.C. 27110, USA; e-mail: dichevad@wssu.edu

Christo Dichev – Winston-Salem State University, N.C. 27110, USA; e-mail: dichevc@wssu.edu

AN ONTOLOGY-BASED APPROACH TO STUDENT SKILLS IN MULTIAGENT E-LEARNING SYSTEMS

Anatoly Gladun, Julia Rogushina

Abstract: *The main idea of our approach is that the domain ontology is not only the instrument of learning but an object of examining student skills. We propose for students to build the domain ontology of examine discipline and then compare it with etalon one. Analysis of student mistakes allows to propose them personalized recommendations and to improve the course materials in general. For knowledge interoperability we apply Semantic Web technologies. Application of agent-based technologies in e-learning provides the personification of students and tutors and saved all users from the routine operations.*

Keywords: *e-learning, ontology, software agents, multiagent systems, Semantic Web technologies*

ACM Classification Keywords: *I.2.11 Distributed Artificial Intelligence: Multiagent systems*

Introduction

The growth of the information society provides a way for fast data access and information exchange all over the world. Now, a lot of materials that used for education process exist in electronic form. In many domains the results of work are plain test materials. However, human readable data resources (like student control works, reports etc.) have serious problems for achieving machine analyses.

Computer technologies have been significantly changing the content and practice of education. The consequent applications of multimedia, simulation, computer-mediated communication and communities, and internet-based support for individual and distance learning all have the potential for revolutionary improvements in education [1, 2].

Online learning (e-learning) offers new possibilities in learning: a student can et immediate feedback on solutions to problems, learning paths can be individualized, etc. Online learning is a growing business: the number of organizations working on online learning and the number of courses available on the Internet is growing rapidly. Now a lot of e-learning tools with varying functionality and purpose exist [3,4].

E-learning is a contrary concept to the traditional human tutoring system. The course tutor in a software tutoring system controls learners relatively weaker than in the traditional tutoring where it is the tutor who is charge the contents and sequence of instructions. Therefore, in order to obtain better tutoring outcomes, a software tutoring system should emphasize engaging students in the learning process and be adaptive to each individual learner.

The goal of the early software tutoring systems was to build user interfaces that provide the efficient access to knowledge for the individual learners. Recent and emerging work focuses on the learner control over the learning process such as learner exploring, designing, constructing, and using adaptive systems as tools [5].

For example, 61% of Helmut Schmidt University students that use ILIAS (www.ilias.uni-koeln.de) prefer electronic forms of learning.

With the application of more computer techniques in education and the involvement of more adults in software tutoring systems, the learner control strategy has become more appreciated than tutor control or program control.

Learning control needs the comparison means of learner's knowledge base that forms (and modifies) in learning process with the course domain knowledge base. It requires the powerful and interoperable tools of knowledge representation and analysis.

A structured information representing is required and ontologies (machine processable representations containing the semantic information of a domain) can be very useful. The ontology systems serve as a flexible and extendable platform for knowledge management. The inspiring idea to develop reusable atomic learning components and to capture their characteristics in widely-accepted, formal metadata descriptions will most probably attract learning object providers to annotate their products with the accepted standards.

The experience of the developed countries shows the technological achievement of remote training – e-learning - that opens many new opportunities in expansion of student's number with the same number of the tutors and in improving of education quality. In recent years, e-learning has been widespread, especially since standardizing initiatives for learning technologies [6] have begun.

For distant learning where the tutor works with many students without direct contacts, it is very important to provide the objectivity and automatization of knowledge examination.

An important component of e-learning is testing of student's skills and knowledge. One of the main problems arising during creation of testing systems is interoperability of created tests, opportunity to reuse these tests in different testing systems. To organize test exchange between various systems it is necessary to create some universal format of tests preservation and their processing instructions. And an important condition for this format should be its independence from the platform. Standardization of educational technologies and, in particular formats of test data preservation are working out all over the world. Now Ministry of Education and Science of Ukraine realize the Program of On-line Education Development. According to these activities the development of projects of standards for systems, methods and technologies standards of on-line education in educational institutions taking into account international standards was provided. But different test formats such as Instructional Management Systems (IMS) Question and Test Interoperability (QTI) of Global Learning Consortium are not adequate to reflect all relationships of different applied domains.

Many authors [7,8] utilize the ontology's semantic data to improve the analyses of information in unstructured documents. The domain ontology plays central role as a resource structuring the learning content [9].

One of the key challenges of the course construction process is to identify the abstract domain of information within which this course will exist. Tutor has to describe the main terms and concepts from which a course is to be constructed.

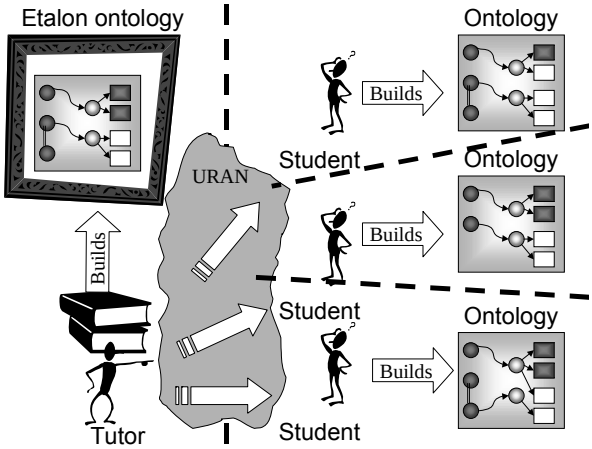
The main idea of our approach is that the domain ontology is not only the instrument of learning but an object of examining and forms by students. We propose for students to build the domain ontology of examine discipline and then compare it with etalon one.

We work on information content of Ukrainian Research and Academic Network (URAN) that is oriented in consolidation if education and scientific organizations of Ukraine [10]. Network is used to perform access to information sources. URAN provides opportunities for conducting scientific researches, distance learning, teleconferencing, etc.

Methodology for Building of Ontology

The students have to make some steps to design the ontology of domain:

1. Define the main classes and terms of domain and describe their meaning;
2. Construct the taxonomy of domain terms;
3. Define synonymy and other relations between these terms;
4. Describe the instances of constructed classes.



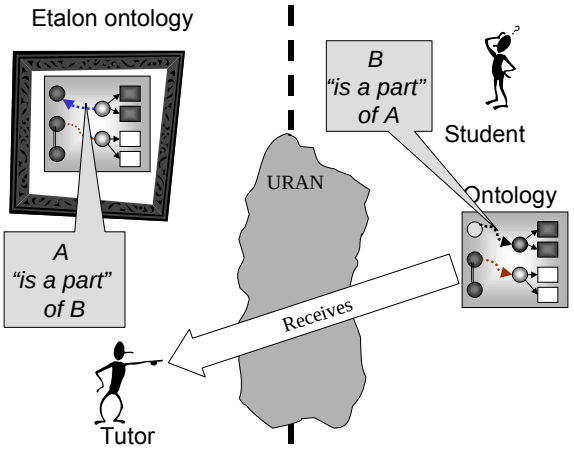
Pic.1. Ontology building process as a result of learning

For knowledge interoperability we apply Semantic Web technologies [11]. The Semantic Web project provides a solid basis for future progress in e-learning.

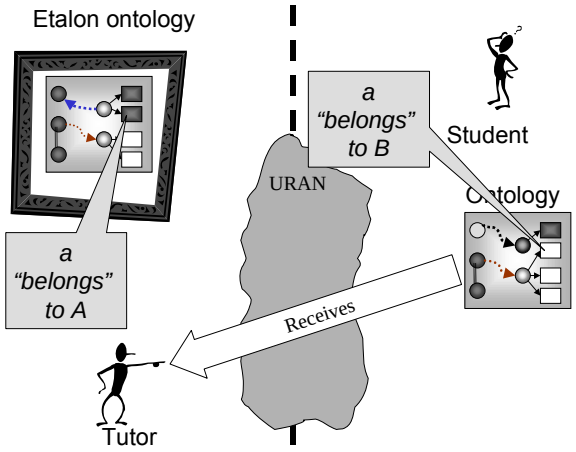
The ontologies are stored in a semantic markup language OWL [12] that is developed as a vocabulary extension of RDF [13] for applications that process the content of information. OWL is supported by many ontology visualizes and editors, like Protégé 2.0 [14]. Protégé is an integrated software tool used by system developers and domain experts to develop knowledge-based systems. Ontology in Protégé is a model of a particular field of knowledge - the concepts and their attributes, as well as the relationships between the concepts. It is represented as a set of classes with their associated slots.

Formal model of ontology O is ordered triple of finite sets O = < T, R, F > [15], where T - the domain terms of which is described by ontology O; R - finite set of the relations between terms of domain; F – the domain interpretation functions on the terms and the relations of ontology O. In process of ontology building students use relations from the fixed set that contains the most widely used relations: R={"is a subclass of", "is a part of", "is a synonym", "has attributes", "has elements"}. It simplifies the ontology building and analyses processes.

When student builds the domain ontology the tutor can compare it with etalon ontology, constructed by tutor. We use the original algorithm for automatically comparing of ontologies that provides correspondence of hierarchical levels in term taxonomy (if class A is a subclass of B in etalon taxonomy and B is a subclass of A in students taxonomy there is a mistake - pic.2) and controls affiliation of instances with classes (if instance a belongs to class A in etalon taxonomy and student describe instance a that belongs to class B is a mistake - pic.3).



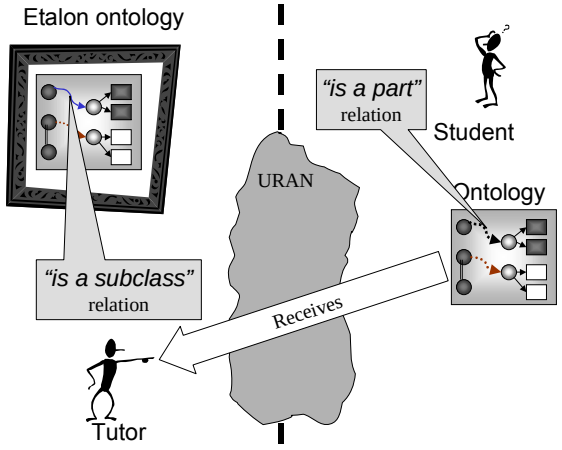
Pic.2. Hierarchical direction class error



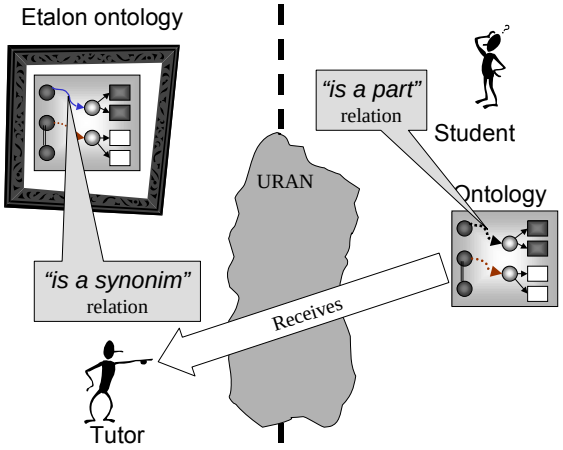
Pic.3. Instance classification error

We distinguish the mistakes of different gravity. If student uses improper relation but from group of hierarchical relations (for example, A is a part of B instead of A is a subclass of B - pic.4) it is not so important as if she or he

uses hierarchical relation instead of synonymic relation (for example, A is a part of B instead of A is a synonym of B - pic.5). More serious mistake is improper direction of hierarchical relations (for example, A is a part of B instead of B is a part of A - pic.2).

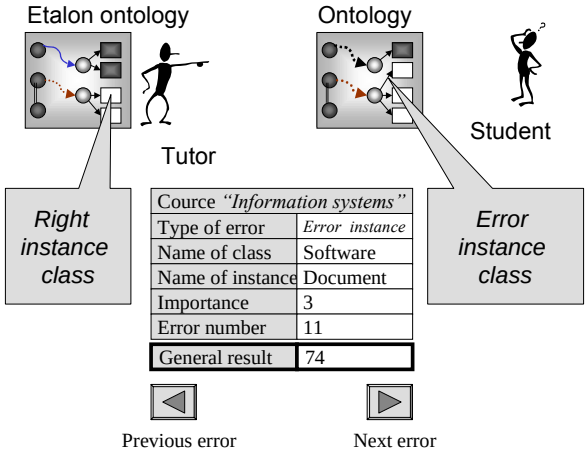


Pic.4. Class name error



Pic.5. Hierarchical class error (instead of single-level)

On base of this algorithm we grade the results of students work with 100-ball system. The experimental prototype of system that controls student's knowledge by means of ontological analyses in URAN network was developed by Java.



Pic.6. Rating of the student ontology correctness.

Intelligent Software Agents in e-Learning

Ontological representation of student domain skills can be automatically processed by intelligent software agents [16]. It is appropriate to use software agents for e-learning because they work efficiently in dynamic heterogeneous distributed environment [17].

Agent works autonomously without human direction in an actual or simulated environment. An intelligent agent may be defined as a computational process that can perform tasks autonomously. It inhabits a complex and dynamic environment with which it may interact to accomplish a given set of goals. Intelligent agents can reason in a rational manner and report back result to humans.

One of the main properties of an intelligent agent is sociability. Agents are able to communicate between themselves, using some kind of agent communication language, in order to exchange any kind of information. In that way they can engage in complex dialogues, in which they can negotiate, coordinate their actions and collaborate in the solution of a problem.

A set of agents that communicate among themselves to solve problems by using cooperation, coordination and negotiation techniques compose a multi-agent system (MAS). A lot of researchers use MAS for e-learning and e-coaching tasks.

Personalized e-learning employs an active learning strategy which empowers the learner to be in control of the context, pace and scope of their learning experience. It supports the learner by providing tools and mechanisms through which they can personalize their learning experience. This learner empowerment and shift in learning responsibility can help to improve learner satisfaction with the received learning experience.

The aims of personal e-learning agents are at increasing of information dissemination of existing courses through delivering the relevant course information offer to the right student at the right moment. For example, students of different specialties learn on different programs and in many cases have different theoretical and practical background. Their personal agents can consider it and propose them not only the universal course program but additional facts and references from allied courses that they didn't learn.

Application of agent-based technologies in e-learning provides the personification of students and tutors and saved all users from the routine operations.

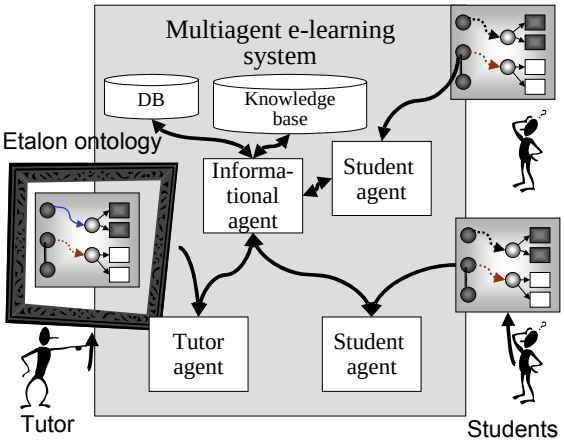
Results and Future Work

The main contribution of this paper is our architecture of e-learning MAS and methods of ontology comparing.

We use the multiagent system of e-learning. It includes personal agents of students and course tutors. Use of some agents-facilitators raises the efficiency of this system and helps to users in search of required information. Agents of students and tutors don't communicate directly. They send ontological information to informational agent that analyses them and returns the results to students and tutor.

We develop a prototype of a multiagent ontology-based e-learning system that produces automatically semantic control of student domain beliefs of learned course.

The focus of ontology analysis is on knowledge structuring (of main domain terms and their relations). We use ontologies to describe learning materials and to represent student belief about course domain. Etalon ontology is sent to MAS knowledge base by tutor personal agent.



Pic.7. Architecture of e-learning MAS.

When student forms the domain ontology in OWL format her/his personal agent connects not with course tutor personal agent but with informational agent and sends this ontology for comparing with etalon one (its last version). After the comparing, informational agent sends these results to the student and tutor personal agent. If some student or tutor usually prefers some manner of learning and information presentation then the personal agent has to provide all these requirements for new course without direct instructions of student. Student receives information in appropriate to her/him form and taking into account previous results of examination. For example, if student makes the same mistakes in some ontologies she/he receives the notification about it and advice that links with suitable course materials.

In future we plan to use the inductive inference methods to form the most appropriate personal strategy of learning for every student (for example, some students profit by theoretical materials and some other ones - from examples or practical tasks, somebody prefers graphical or text representation of information, etc.).

Use of student personal agent allows finding the situation where student makes mistakes of the same type in ontologies of different courses. She/he receives the notification about it and advice that links with suitable logical course materials. Other important advantage of multiagent technology use is deals with course tutor personal agent. If the big part of students make the same mistakes course tutor receives the notification about it and can change suitable course materials

We have been experimenting with the challenges of e-learning MAS exploiting for student beliefs control in course "Visual C++ System Programming" (European University, Kiev, www.eufimb.edu.ua) and "Modern Internet Technologies" (Kiev Slavistic University, www.mgi_ksu .edu.ua).

We create detailed domain models of these courses. Model of course "Visual C++ System Programming" contains more then 150 terms and uses 8 relations between these terms. Course is accompanied by 16 online lectures and 10 practical exercises. Model of course "Modern Internet Technologies" contains more then 68 terms and uses 10 relations between these terms. Course is accompanied by 18 online lectures and 6 practical exercises.

Table 1. Average measures of student ontology parameters

Course title	"Visual C++ System Programming"	"Modern Internet Technologies"
Number of students	22	16
Number of terms	153	68
Number of relations	8	10
Terms correctly used in student ontology K_{term}	94.2%	91.6%
Relations between terms correctly used in student ontology K_{rel}	72.0%	66,3%
Type of relations between terms correctly used in student ontology K_{type}	89.1%	81.5%
m_{term}	0.9	0.7
m_{rel}	0.3	0.5
m_{type}	0.7	0.8
Overall rating of student ontology correctness K	88,51%	80,14%

$$K = (K_{term} * m_{term} + K_{rel} * m_{rel} + K_{type} * m_{type}) / (m_{term} + m_{rel} + m_{type})$$

Conclusion

The main features of our approach to knowledge control are the following:

- all results are analyzed automatically without tutor;
- results are analyzed objectively;
- students can work with knowledge base;
- a structurization of domain knowledge simplifies the learning process;
- tutors can exchange their knowledge based on etalon ontologies.

One of the essential elements needed for effective learning is feedback. In the current generation of e-learning systems automatically produced feedback is almost only used in question-answer situation. Valuable feedback, for example produced by a human tutor via e-mail, is often possible but this introduces delays and is time consuming. We want to develop ontology-based mechanisms of feedback that use the context of education. Different student errors need different methodologies of tutor to describe their causes.

Feedback is used in many learning paradigms. The concept of feedback is very important in educational psychology. It is one of the main psychological principles that one of the essential elements needed for effective learning is feedback. Information about examining results is required to assess progress, correct errors and improve performance. Feedback describes any communication or procedure given to inform a student of the accuracy of a response, usually to an instructional question. Feedback allows the comparison of actual performance with some set standard of performance. Information that is acquired by student from feedback instruction includes not only answer correction but other information such as precision, timeliness, learning guidance, motivational messages, background material, sequence advisement, critical comparisons, and learning focus. In traditional learning students and tutors can interact directly and students can freely ask questions and tutors usually know whether their students understand concepts or problem solving techniques and relations between them. Feedback is an important component of this interaction. In e-learning systems feedback problem is much more difficult and has a lot of technological and social aspects.

In future, we plan to develop more powerful algorithms of ontology analyses that consider ontology integration and their distributed upgrade on base of multiagent technologies. Application of student and tutor agents will provide the personalization of distributed learning process. These agents will use the history of learning for feedback between student and tutor.

Acknowledgements

This work was supported in part by the Grant NATO NIG 971779. "National Telecommunication Networks for Scientific and Educational Institutions of Ukraine with Access to Internet - URAN" and EU INCO-Copernicus Project 960114 – EXPERNET "A distributed Expert System for the Management of a National Network of Ukraine".

Bibliography

1. McArthur, D., Lewis, M., and Bishay, M. The Roles of Artificial Intelligence in Education: Current Progress and Future Prospects. 1993. - http://www.rand.org/hot/mcarthur/ Papers/role.html.
2. Forbus, K. D. and Feltovich, P. J. The Coming Revolution in Educational Technology. Smart Machines in Education, eds. Forbus, K. D. and Feltovich, P. J., pages 3-5, AAAI Press/MIT Press, 2001.
3. Aroyo, L., Dicheva, D., Courseware authoring tasks ontology // Proceedings of the international conference in education, IEEE, 2002.
4. T. Murray, S. Blessing, S. Ainsworth, Authoring tools for advanced technology learning environments: towards cost-effective adaptive, interactive, and intelligent educational software. - http://helios.hampshire.edu/~tjmCCS/atoolsbook/chaptersV2/ChapterList.html.
5. Kay, J. Learner Control. User Modeling and User-Adapted Interaction, 11:111-127, 2001.
6. IEEE Learning Technology Standards Committee (LTSC): http://ltsc.ieee.org/
7. Magnin L., Snoussi H., Nie J., Toward an Ontology–based Web Extraction, The Fifteenth Canadian Conference on Artificial Intelligence, 2002.
8. Bredeweg B., Forbus K. D. Qualitative Modeling in Education. AI Magazine, V. 24, No. 4.
9. Angelova A., Kalaydjiev G., Strupchanska A. Domain Ontology as a Resource Providing Adaptivity in e-learning. - lml.bas.bg/~albena/publications/wose_paper5.pdf.
10. Ukrainian Research and Academic Network (URAN). - http://www.cei.uran.net.ua/eng/uran.htm.
11. W3C Semantic Web Activity. - http://www.w3.org/2001/sw/Activity,
12. OWL. - http://www.w3c.org/TR/owl-features/.
13. Resource Description Framework (RDF) Model and Syntax Specification, W3C Recommendation. - http://www.w3.org/TR/REC-rdf-syntax/.
14. Protege. - http://protege.stanford.edu/ontologies/ ontologyOfScience.
15. Gladun A., Rogushina J. Multiagent ontology-based intelligent system of e-commerce. The International Conf. TAAPSD'2004.
16. Wooldrige M., Jennings N.R. Intelligent Agents: Theory and Practice / Knowledge Engineering Review, Vol.10, No.2, 1995.P.115-152.
17. Gritsenko V., Gladun A., Rogushina J. Agent-oriented Programming as a Tool of the Intellectual Distributed Applications Development. Proc. of International Scientific Conference "Software Design: Challenge of Time and Role in Informational Society", 2005.

Authors' Information

Anatoly Gladun - PhD**,** Senior Researcher, International Research and Training Centre of Information Technologies and Systems, National Academy of Sciences and Ministry of Education of Ukraine, 44 Glushkov Pr., Kiev, 03680, Ukraine.

Julia Rogushina - PhD**,** Senior Researcher, Institute of Software Systems, National Academy of Sciences of Ukraine, 44 Glushkov Pr., Kiev, 03680, Ukraine; e-mail: jjj@cybergal.com

SELF-ORGANIZATION OF FUNCTIONING PROCESS OF EXPERT SYSTEM WITH THE USE OF SUBJECT DOMAIN ONTOLOGY

Elena Netavskaya

***Annotation**: Optimization of design, creation, functioning and accompaniment processes of expert system is the important problem of artificial intelligence theory and decisions making methods techniques. In this paper the approach to its solving with the use of technology, being based on methodology of systems analysis, ontology of subject domain, principles and methods of self-organisation, is offered. The aspects of such approach realization, being based on construction of accordance between the ontology hierarchical structure and sequence of questions in automated systems for examination, are expounded.*

***Keywords**: Expert systems, ontology, self-organisation*

***ACM Classification Keywords**: I.2.1. Application and Expert Systems*

Introduction

Expert Systems (ES) relate to basic and most early directions of artificial intelligence. The first ES, being based on the use of the computing engineering, are known from the 60-th years of past century. Today it is difficult to name the branch of science or production, where they were not used. At the same time, swift motion on stages

<society without boundaries>→<information society>→<society, being based on knowledge>

and like avalanche growth of quantity information considerably complicate the adequate use ES, that it is related to the informative redundancy, incompleteness and illegible information, its subjectivity. We have the knowledge representation problems of experts, optimizations for receipt processes of conclusions in ES, determinations of plenitude of information bases.

The automated systems for study examination (ASSE) related to information technologies are the typical representative ES. Not deciding on the tasks of tests verification and results estimation we will mark that all foregoing problems are reflected in them to a full degree. In further exposition we will follow a conclusion [Feigenbaum, 1963] about that is "basic principle of knowledge engineering, that possibilities of Problems Solver by intelligence agent above all things are determined by its informative base and only in the second – by the used method of conclusion".

Analysis, description and component parts of expert systems creation process

An idea or necessity is pre-condition and incentive reason for new knowledge receipt. In the ES case such necessity consists in the receipt of some conclusion, being one of determinative features at decision making. With ASSE the level of knowledge is estimated. We will illustrate the process of the ES creation through three methodological structural units which are ontology (O), system analysis (SA) and self-organisation (SO).

Following the stages of SA algorithm, in accordance with [Plotinskiy, 1992], [Zgurovskiy, 1997], we will define the goal of the ES creation as a mean the attribute of which is feature to accumulate expert knowledge and afterwards possibility to replace experts in the decision making processes. The problems decided for goal achievement are determined by a subject domain. In particular, for ASSE are such the knowledge examination, their level determination and, partly, acquisition. The ES functioning is determining by an external environment, which information as themes, questions, possible answers acts from, inference rules and in which a result is passed – estimation of knowledge level.

Be complex system, ES it is possible to present by three models [Timchenko, 1991]: structure, functioning and development. The structure model is a theoretical-plural model:

$$M_b = < I_t, I_q, I_a, R_o >, \quad (1)$$

which the element base of the system is represented in: I_t – informative table for the themes examination; I_q – tables of the given questions (depending on the questions types of them can be some); I_a – tables of possible

answers with pointing of their rightness gradation; R_o – aggregate of inference rules, sometimes presented by some procedure or algorithm. The functioning model determines the process of goal achievement by the system, which is carried out by its elements, subsystems, integral ES:

$$M_f =< O_t, O_q, Q_a, P_1, P_2, ..., P_n, A >, \quad (2)$$

where O_t – dynamic operations accompanying the process of theme choice; Q_q – dynamic operations of questions set choice; Q_a – dynamic operations for forming the possible answers set and their estimation; $P_i, i = \overline{1,n}$ – aggregate of operations realizing the sequence of transitions at examination; A – algorithm for knowledge evaluation. The development model reflects the motion of ES, possessing the attributes of openness, mobility, system and informative unity, complexity on the stages of its life cycle:

$$M_d =< A_d \Theta A_m \Theta A_u, B_r >, \quad (3)$$

where Θ – logical disyunction or konyunction, A_d – adaptation procedures to the change of external terms; A_m – modernization procedures and use of new technologies; A_u – procedures of partial or complete utilization; B_r – machineries of feed-back, allowing taking into account future processes to produce the changes in ES on all stages of its life cycle.

Models (1)-(3) are an additional informative feature allowing to carry out structurisation of process of ES creation and functioning. We will remark that ES will be effectively functionate, as already it is indicated higher, if it has the adequately formed informative base. In the its capacity rationally to use ontology – the enough complex organized structure of knowledge about a subject domain from one side, with other – initial material for the receipt of new knowledge [Grechko, 2005]. Construction and use of ontology in ES is based on that:

- ontology at that rate is jointly used by the collectives of agents;
- knowledge about a subject domain are used repeatedly ;
- knowledge about a subject domain are separated from a process and algorithm of examination;
- it is needed for the analysis of knowledge about a subject domain.

Development and use of ontology on today not formalities, for their construction there are only some fundamental rules. At the same time, the application domains of ontology and aspects of their development are enough various, about what the papers analysis in 2005 year is testifies. So, in [Grechko, 2005] ontology of analysis method for the Saati hierarchies is built; in the paper [Darevich, 2005] for the rise of intellectual analysis of text it is offered to use weighing of notions in the ontology model; authors [Kucherenko, 2005] consider the questions related to presentations in ontology of fuzzy notions and relations; the paper [Shalfeeva, 2005] is devoted to research of evaluation in the process of ontology creation. The algorithm of comparison of interests for salespeople and buyers in Internet-shops is offered in paper [Gladun, 2005]; in the paper [Gribova, 2005] the problem of development of user interface with the use of different types of dialog, being based on the use of ontologies, is considered; procedure of structurisation of knowledge in an application area is studied in [Palagin, 2005]. The paper analysis shows that they can be divided into two types: the articles which the aspects of creation and improvement of ontologies are studied in behave to one, to other – decisions of the applied tasks with their use.

As, as stated above, knowledge about a subject domain exist separately from the methods of analysis, inevitably there is the problem of their composition. Such co-operation in the conditions of informative redundancy must be enough effective on the criterion of time minimization. However the problem solving comes into collision with the problem of synergetic effect [Haken, 1985], in consequence of what structures, uniting data, questions, answers and inference rules possessing some time stability, are appear. Such structures aspire in time to position with "minimum energy". But achievement of the indicated state takes place in the conditions of incomplete information. Thus, from one side there is the informative redundancy of initial data and shortage of information about process optimization and results of knowledge control (in ASSE) – with other. Solve such contradiction is suggested with the use of self-organisation of transitions process on the ES levels (sequences of questions for the knowledge estimation in ASSE). Self-organisation procedure of the ES functioning will allow to rationalize and intellectualize process of information treatment necessary for the receipt and knowledge estimation.

Aspects of ontologies creation for expert systems

The ontologies are admitting multiple use in different application, they can be complemented and modified. For many subject domains, especially in the field of commodities of consumption, to ontologies are already built. We will consider the features of their creation for the knowledge estimation. We will assume that a course on information technologies is the studied subject. Coming from the problem-oriented exposition of subject, discipline contains lectures, practical and laboratory studies, that corresponds to acquisition by studying knowledge Z, abilities U and skills N. Thus, the subject study determines realization of reflection:

$$K \to < Z,U,N > . \quad (4)$$

It is known, that ontology – it is the attempt of all-embracing and detailed formalization of some domain of knowledge by a conceptual chart, which consists of hierarchical data structure, containing all relevant classes of objects, their communications and rules accepted in this domain. Formally, ontology presents as a three elements $\{X,R,F\}$, where X – finite set of concepts (elements of knowledge), R – set of relations between concepts, F – set of interpretative functions for concepts and / or relations. Construction of ontology for ES, which is ASSE, begins from determination of subject domain limits, being based on presentation of place and role of educational subject in the general flow diagram of subjects and, as a result, finding of a priori information being the base of subject study. A subject domain includes basic concepts, and also list of problems which outline the informative field of course (fig.1).

In accordance with the list of problems it is necessary to define concepts and build a hierarchical structure. So, in a subject on information technologies concepts of the first level would be: “System”, ”Information”, ”Intelligence”, ”Technology”, Model”, ”Planning” and some other. They are determined by expert assumptions as above all, base notions. Attributive nouns or attributes of concepts of higher level will be concepts of the second level, for example, “interface design “ or “mathematical model”. Such construction proceeds to achievement of base elements – structural units possessing an independent value for the subject study.

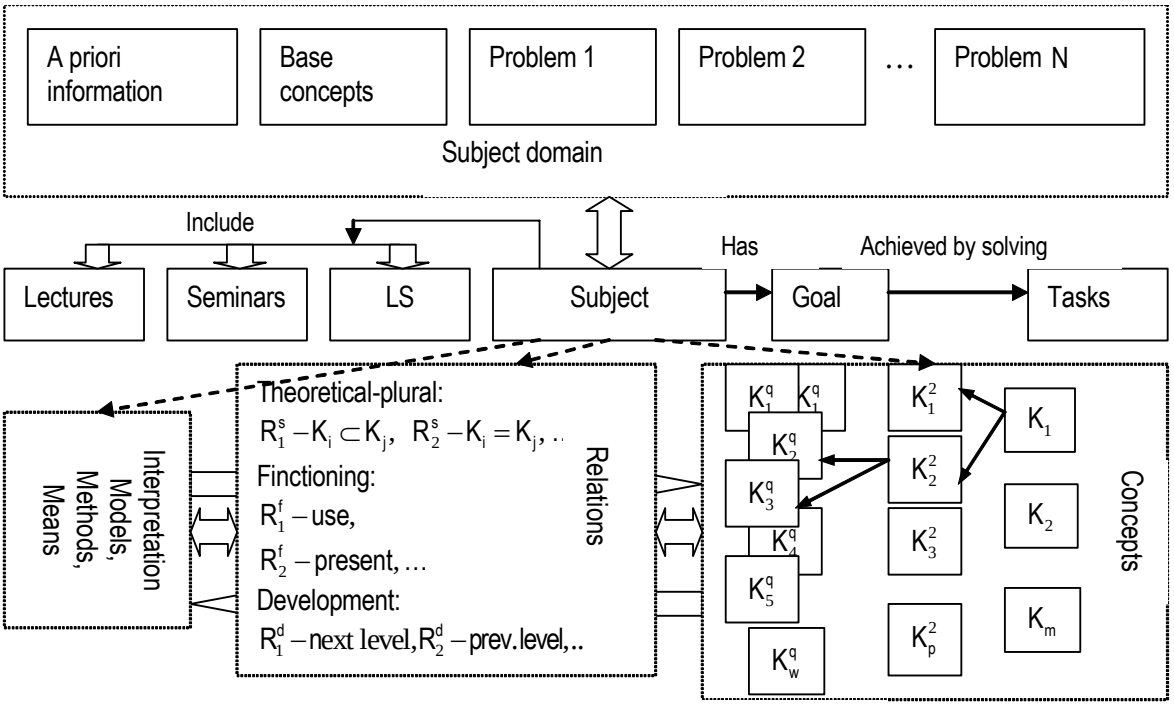


Fig. 1 – ASEE on the ontology base structural chart

Some objectivisation in the process of ontology construction can be brought the automated analysis of electronic summary of lectures text. For this goal on the first stage it is necessary to count up the number of the most often meeting words-concepts of the first level. On the second stage the number of the most often meeting words-attributes (adjectives) or attributive notions (nouns) of the first level concepts is calculated. The third stage (can be inverted with the second) is dedicated to determination of often meeting relations of concepts functioning of

the first level (sometimes jointly with concepts of the second level), expressed by the verbs of action and etc. Automated treatment allows to define concepts of low levels being attributes, by determining concepts for higher levels. Resulting ontology is got, carrying out composition of the data receipt it is automated, and expert conclusions.

Ontology and questions superposition technology in expert system

Technology of interactive intercourse, supposing the presence of decision making person and user (expert) answering questions of the system, lies in the ES basis. Rational conducting of such dialog which on an eventual stage is conducted between a user and computer is possible at implementation of such terms: a questions and possible answers data-base is enough complete for acceptance of decisions; procedure of examination is verification; the algorithm of examination must be time optimized.

Traditionally, conducting procedure of examination was fully determined by a decision making person, investigation of what was the decision being based exceptionally on the subjective preferences. In the ASSE lecturer made a list of questions from different theme of subject. For the examination their definite number was offered. For an answer for a question student got marks and their summary number made estimation, which, in the general case, did not guarantee subject knowledge, as a random selection of questions did not reflect the complete picture of the studied problems and tasks. Material of some sections remained out of examination, and some questions with insignificant modifications repeated oneself.

We suggest to use by examination principles of self-organisation [Ivakhnenko, 1975], [Molchanov, 1988]: "multiplicity" of models and choice freedom. Their interpretation in relation to our problem consists that there is the set of well-organized sequences of test questions resulting in a correct result, and also that on any stage of testing a few variants of its continuation can be chosen.

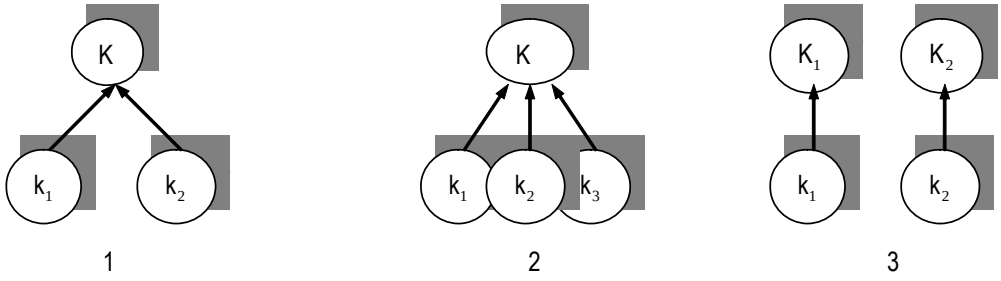


Fig.2 – Possible constituents of ontology construction

Making the examination sequence, we will follow a structure definite by the graph of ontology. The variants of accordance choice, resulted on the fig. 2, are thus possible. We will remark that graph presenting ontology, has the descending orientation. At the same time, the examination process can have as descending (deductive – from general to partial), so ascending (inductive – from partial to general) character. On the fig. 2 fragments proper to ascending approach are shown. Forming of questions sequence at that rate answers a "And-Or" structure. Concepts of low level k_1 and k_2 (see fig. 2.1.) make concept K and between them it is possible to put logical "OR". Then in the questions sequence there must be necessarily questions on essence of both concepts and right answers on them suppose knowledge of essence K and deliver from questions on its maintenance. We will remark that concepts of low level for k_1 and k_2 have not general elements, notions, functions. If this is not so (see fig. 2.2), rationally to set questions on concept essence, or its constituent, being general for one level concepts. If concepts are at one level (see fig. 2.3) and have not crossing at more low level, right answers for questions an incident to all one level concepts are the necessary condition of examination passing.

A expert system can function in the "active" and "passive" modes. In the "passive" mode a questions sequence by the system are determined beforehand and written in a data-base, in "active" mode – the sequence of questions is formed in the process of student answers. In the first case the time on the question generation is minimized, but the adequate reaction on the rightness of answers is absent, in the second – if power of ontology is enough large, time of determination of a next question can be considerable. Advantage of the "active" mode consists that

there is possibility of flexible reaction and determination of sequence of next questions depending on previous answers.

Problems and prospects of functioning process optimization for expert system

The process of extraction of knowledge with the use of ES is based at function of three subsystems [Luger, 2003]: editor of knowledge base, conclusion machine and subsystem of explanations. Optimization of their functioning requires the decision of such tasks:

- formal presentation of ontology in the element bases of knowledge base;
- providing of determination possibility of accordance between presentation of ontology and table containing thematic questions;
- developments of examination algorithm (knowledge examination) foreseeing possibility of the flexible tuning as a result of self-organization of questions base in real-time mode;
- development of models and methods of examination, the initial stage of which is formalization of questions depending on the types of answers;
- consideration of possibility of fuzzy presentation of subjective judgement;
- development of the reporting system and interpretation of the ES functioning results, foreseeing explanation of examination logic.

As a result of indicated problems solving the prospects of systems approach to the ES creation in different domain of knowledge are opened. The considerable degree of standardization of process of their creation and planning optimizes the process of receipt of expert conclusions. Forming of the ontologies proper subject domains is the necessary condition of this, a sufficient condition consists in realization of technology of superposition of test elements and elements of ontology.

Conclusion

The offered technology of expert system creation, being based on methodology of system analysis, ontology of subject domain, and also principles and methods of self-organisation is another step in the direction of effective expert systems creation. Efficiency consists in time minimization of examination conducting; in decisions objectivisation, being based on automation of expert analysis process, more complete scope of subject domain and reduction of informative redundancy of test questions and their sequences; indirect forming at experts and student conception of subject domain structure, its base elements and their functional intercommunications.

Swift development of distance learning is another arguments in behalf on creation and use of the ASSE, being based on the use of subject domains ontologies being the basis of educational courses. We will remark that development of ontology is an enough difficult and labour-intensive process, therefore rationally this process within bounds of educational establishment, and afterwards in more wide scales unify, for what to develop the program-methodical providing. Their use as base platforms for development of distance courses, integrating in itself the subsystems of learning, reference information, methodical literature, test examples, receptions of reports and control tasks, will be another application for developed ontologies.

Bibliography

[Feigenbaum, 1963] E.A. Feigenbaum. The simulation of verbal learning behavior. In Feigenbaum and Feldman. – New York: McGraw-Hill, 1963.

[Plotinskiy, 1992] Yu.M. Plotinskiy. Mathematical design of dynamics of social processes. – Moscow: MSU, 1992. – 133 p.

[Zgurovskiy, 1997] M.Z. Zgurovskiy, A.V. Dobronogov, T.N. Pomerantseva. Research of social processes on the basis of systems analysis methodology. – Kiev: Naukova dumka, 1997. – 222 p.

[Timchenko, 1991] A.A. Timchenko, A.A. Rodionov. Informatics bases of the system design of new technique objects. – Kiev: Naukova dumka, 1991. – 152 p.

[Grechko, 2005] A.V. Grechko. Ontology of analysis method for the Saati hierarchies // Artificial intelligence. – 2005. – № 3. – Pp. 746-757.

[Darevich, 2005] R.R. Darevich. Rise of efficiency of test intellectual analysis by weighing of notions of ontology model // Artificial intelligence. – 2005. – № 3. – Pp. 571-577.

[Kucherenko, 2005] E.I. Kucherenko, D.A. Pavlov. Some aspects of development analysis of fuzzy ontology // Artificial intelligence. – 2005. – № 3.– Pp. 162-169.

[Shalfeeva, 2005] E.A. Shalfeeva. Classification of structural properties of ontology // Artificial intelligence – 2005. – № 3. – Pp. 67-77.

[Gladun, 2005] A.Ya. Gladun, Yu.V. Rogushina. Ontology how perspective direction of intellectualisation information retrieval in the multiagent systems E-commerce // In Proc. XI-th Int. Conf. "KDS-2005". – Varna.– Vol. 1. – Pp. 158-165.

[Gribova, 2005] V. Gribova. Implementation of various dialog types using an ontology-based arproach to user interface development // In Proc. XI-th Int. Conf. "KDS-2005". – Varna.– Vol. 1. – Pp. 153-158.

[Palagin, 2005] A. Palagin, V. Peretyatko. Development of procedures of objects recognition with usage multisensor ontology controlled instrumental complex // In Proc. XI-th Int. Conf. "KDS-2005". – Varna.– Vol. 1. – Pp. 140-147.

[Haken, 1985] H. Haken. Synergetycs. – M.: Science, 1985. – 320 p.

[Ivakhnenko, 1975] A.G. Ivakhnenko. Long-term prognostication and management by the difficult systems. – Kiev: Naukova dumka, 1975. – 311 p.

[Molchanov, 1988] A.A. Molchanov. Design and modeling of the difficult systems. – Kiev: Visha shkola, 1988. – 359 p.

[Luger, 2002] G.F. Luger. Artificial intelligence. Structures and strategies for complex problem solving. – Addison Wesley: Boston, 2002. – 864 p.

Information about an author

Elena Netavskaya – Cherkasy state technological university; Av. Shevchenko 460/501, Cherkasy, Ukraine; e-mail: neelena@list.ru

PROTOTYPING ADAPTIVE ONLINE LEARNING COURSES[1]

Tim Geleverya, Olga Malinovskaya,
Tatiana Gavrilova, Vladimir Golenkov

Abstract: *This article describes the process of prototyping adaptive online learning using the authoring tool for developers, which is based on ontologies. The article also gives a brief overview of contemporary situation and describes modern trends of evolution e-learning courses and present standards in this area. It also describes architecture of system VITA II.*

Keywords*: E-learning, ontologies, user modeling, LOM.*

ACM Classification Keywords*: K.3.1 Computer Uses in Education – Distance learning, K.3.2 Computer and Information Science Education – Computer science education, E.1 Data Structures – Trees, I.2.6 Learning – Concept learning (Knowledge acquisition).*

Introduction

Using new computer technologies modern teachers can easily create online-courses accessible through the Internet.

However there are a number of problems with the creation of e-learning courses. The first one is that not every teacher posses an equal good knowledge in computer technologies, which are necessary for creation and management of e-learning courses.

The second problem is that users of e-learning courses do not posses an equal knowledge level. They also have different psychological, social and other personal characteristics. (All these facts, defining personal characteristics of users, create so-called "user model"). That is why the data, assignable to different users, must be different. As a result, the e-learning system is needed of some measure to be adaptive for ensuring the best process of education, or in other words it must be slaved for exact user model.

[1] This work is supported by grant Bel-RFFI 06-01-81005.

In this connection, the main goal is to create universal authoring tool for developers of e-learning systems with user-friendly interface, which can help teacher with minimum set of knowledge in informational technological area develop adaptive e-learning course based on user model, suitable world standards in e-learning.

State-of-the-art standards of e-learning courses

Today there are plenty of organizations, developing standards for e-learning systems. Each of these organization popularizes its own standard. Some of the leading organizations with the most widely accepted standards are [Hage H., Aimeur E.]:

- IEEE Learning Technology Standards Committee [http://ltsc.ieee.org/];
- ADL Initiative (Advanced Distributed Leaning) [http://www.adlnet.org/];
- IMS Project (Instructional Management System) [http://www.imsproject.org/].

According to ADL (Advanced Distributed Learning) [http://www.adlnet.org/] standards materials, designed for e-learning must correspond to following requirements:

- *reuse* – materials must be easily designing and using with different tools;
- *accessibility* – materials must be easily assessable both for students and developers of curses;
- *interoperability (интерпорабельность)* – materials must posses an ability to share upon different programming basics;
- *extendibility* – materials mast be easily modified and tuned for new versions of software.

Today with the appearance and spreading of e-learning standards, learning materials and tests should correspond to some abilities: to be reusable and to be shared upon various e-learning basics (platforms).

Furthermore ADL introduces standard ADL SCORM (Shareable Content Object Reference Model) [SCORM]. SCORM is an XML-based scheme, which is used for describing and getting access to learning object. The SCORM meta-data application profiles defined in this section directly references the IEEE LTSC Learning Object Meta-data (LOM) standard [LOM].

The data model “Learning Object Metadata” is described with the «IEEE 1484.12.1 – 2002, Learning Object Metadata. Standard» from 15/07/2002. This standard is developed by IEEE Review Committee.

The 1484.12.1-2002 Information Model is broken up into nine categories. These categories are based on the definitions found in the 1484.12.1-2002. The nine categories of meta-data elements are [LOM]:

- The *General* category groups the general information that describes the resource as a whole.
- The *Lifecycle* category groups the features related to the history and current state of this resource and those who have affected this resource during its evolution.
- The *Meta-metadata* category groups information about the meta-data record itself (rather than the resource that the record describes).
- The *Technical* category groups the technical requirements and characteristics of the resource.
- The *Educational* category groups the educational and pedagogic characteristics of the resource.
- The *Rights* category groups the intellectual property rights and conditions of use for the resource.
- The *Relation* category groups features that define the relationship between this resource and other targeted resources.
- The *Annotation* category provides comments on the educational use of the resource and information on when and by whom the comments were created.
- The *Classification* category describes where this resource falls within a particular classification system

The LOM data model is an hierarchy of data elements, including aggregate data elements and simple data elements (leaf nodes of the hierarchy).

For each data element, the LOMv1.0 Base Schema defines:

- *name:* the name by which the data element is referenced;
- *explanation:* the definition of the data element;
- *size:* the number of values allowed;
- *order:* whether the order of the values is significant (only applicable for data elements with list values);
- *example*: an illustrative example.

For simple data elements, the LOMv1.0 Base Schema also defines:

- *value space:* the set of allowed values for the data element – typically in the form of a vocabulary or a reference to another standard;
- *datatype:* indicates whether the values are LangString, DateTime, Duration, Vocabulary, CharacterString or Undefined.

Both the size and datatype information may include smallest permitted maximum values.

Authoring tool VITA II

Describing authoring tool for fast prototyping is based on existing authoring tool VITA, designed and described in [Gavrilova, Geleverya, 2002].

While the designing authoring tool VITA two types of learning were taken into account:

- Learning from course content (learning on basis of content). The idea is to learn material in series. This type of learning is most widely distributed. The content of the material is set by the teacher. (The trainee is given a framework of field of knowledge of e-learning course).
- Learning from conceptual definition of knowledge domain given as ontologies and links of concepts with learning materials describing these concepts. The idea is to localize student on tight terms of knowledge domain and to give more detailed information about this terms. In this way it is possible to change the way of navigation for different users. In other words different user models can be compared with own navigation style in hypertext environment. This style is shown by the way of traveling through the given ontologies and reviewing those hypertext pages that they really needs and not all pages one after another as in first case.

In authoring tool «VITA» Groups are entered to realize the first type of learning. ***Group*** is a course, consisting of paragraphs, which puts together basis[2]. These groups are necessary for student with different levels of training.

The Ontologies are entered to reale the second type of learning. ***Ontology*** is a coupled concepts of knowledge domain with which the paragraphs consisting basis are associated.

The authoring tool gives the author of e-learning course following base opportunities [Gavrilova, Geleverya, 2002]:

1. Building course on basis of prepared source material – giving a link between paragraph of basis and source file, consisting information.
2. Storing information in united depository.
3. Managing course with the repository (content of course).
4. Editing and review of paragraphs.
5. Editing of course structure.

The language HTML is used for creating e-learning course. The HTML provides:

- Multilogical features for creating and work with the course;
- Presence of Web Browser;
- Opportunities of creating local and remote courses.

The tool haves following features [Gavrilova, Geleverya, 2002]:

- Simple and user-friendly interface;
- Visual tools;
- Feature of introduction of extension modules.

Learning course consist of paragraphs. Paragraph is an indivisible element of course. Relations of paragraphs organize pseudo tree-like structure of course and define its content.

Attribute content of paragraph contains:

- Name – defines name of paragraph;
- Identification of element in paragraph space, including in course;

[2]Basis is a source structure on course's content.

- Name of the file, containing information about this paragraph;
- List of child elements (paragraphs).

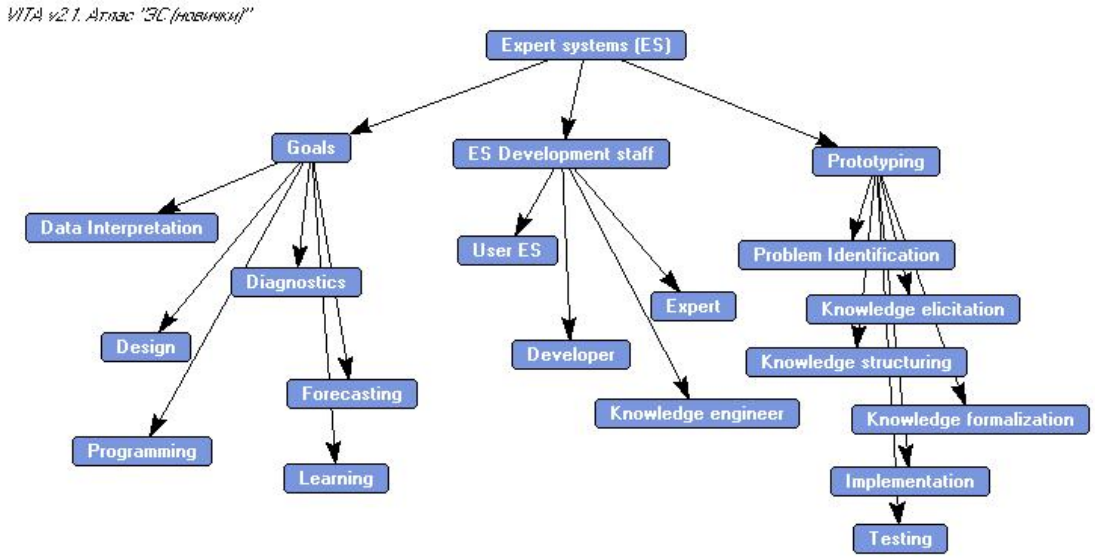
Based on listed above it can be said that using of this authoring tool helps the course' developer quickly atomize the creating and support process of learning course, and to facilitate debugging and searching errors.

Authoring tool VITA was learned in detail and on its base adaptive learning course based on ontologies was created.

Adaptive online course of artificial intelligence

Adaptive learning course was built upon the textbook «Base of knowledge of intellectual systems" [Gavrilova, Khoroshevskiy, 2001].

The three main concepts of knowledge domain were extracted from course "Base of knowledge of intellectual systems": artificial intelligence, expert system, knowledge engineering. The three models of users were defined: beginner, developer, knowledge engineer. The ontology was built both for each user mode and each main concept of course, so nine ontologies were built. The example of this ontology is given in Picture 1.



Picture 1. Ontology of concept «Expert system» for user model «Beginner»

The process of ontology development was divided into several subtasks. At first, the detailed glossary of used in course terms was built (**Glossary of Terms**), The next, **concept classification trees** were built.

As one can see on Picture 1, the main drawback of such ontology is that it is very bulky. So some troubles may cause during work with it.

Every element of ontology can be open as hypertext page for more detailed examining of presented information. Moreover every concept has brief description.

The following drawbacks were revealed during work with authoring tool VITA:

- necessity of availability specific software for setup authoring tool (authoring tool VITA is developed on language Python);
- absence of automatic test generator;
- absence of possibility to give user models and links them with ontologies;
- uncomfortable ontology editor;
- absence of possibility of viewing ontologies in Web;
- absence of possibility of creating multi-level ontologies (desiring ontologies are huge and the work with them is very uncomfortable).

In connection with listed above drawbacks the requirements to developing authoring tool (preliminary named VITA II) were revealed:

- availability of automatic test generator;

- availability of possibility links user models and ontologies;
- possibility of creating learning course without programming;
- user-friendly visualization of ontologies;
- support of multilevel ontologies;
- absence of necessity of additional software for setup authoring tool;
- support of graphic templates, corresponding to contemporary requirements of Web-usability;
- capability of loading Web-pages, created in another editors.

Besides, corresponding to standards of learning systems, the requirements to learning materials, using in developing tool, were picked out:

- possibility of reusing learning materials;
- accessibility of materials to users of other systems and learning courses;
- data generality (compliance Learning Object Metadata and Shareable Content Object Reference Model);
- extensibility (availability of features of downloading of Learning objects from other programs and systems);
- application of international standards of e-learning;
- using contemporary semantic markup languages (RDFS, OWL).

Architecture of system

System VITA II inherits base advantages of tool VITA, particularly availability of presentation data in two views – in form of groups and ontologies.

As was stated above, one of the main drawbacks of tool VITA is that ontologies, giving learning resources to end user, are very bulky and hard to percept. That is why in system VITA II introduce the support of multilevel ontologies.

The next scheme offers the division of multi-lavel ontologies:

1. Content fragment (CF) [Jovanovic' J., Gasevic' D, 2005] – elements of lower level that can't be opened as ontologies (content units in their most basic form, like text, audio, video, images, tables);
2. Content Object (CO) [Jovanovic' J., Gasevic' D, 2005] – ontologies of middle level (contains set of CF, CO and navigation between them);
3. Learning object (LO) [Jovanovic' J., Gasevic' D, 2005] – ontologies of higher level (collections of CO and links between them).

At the same time ontologies of higher level (learning objects) should be described according to standard IEEE 1484.12.1 – 2002 и ADL SCORM Version 1.3, i.e. should be simple extended, accessible for other learning systems and independent from software platform.

Developing software tool supposes a possibility of user's work with different types of information, for example like text, html-pages, images, tables, diagrams, ontologies, multimedia and etc. At the time, to different resource type corresponds own software module, realizing work with exact resource.

As system VITA II is directed on learning that is why source course basis, designing with html-language facilities, is paramount. Hence all material, existing in system, can be divided on:

- initial learning data (course basis, representing group of paragraphs),
- resources (html-pages, images, tables, diagrams, presentations, multimedia and etc.)

Materials, representing in course basis, should be structured in a view of tree, namely broken in accordance with structure of learning course, the recourses should be divided in accordance of data types.

Learning data is supreme aspect and should be linked with the resources by associating every recourse with an element of learning material. The final data representing for the course user is a number of multilevel ontologies, described according to the LOM and SCORM standards.

Basic constructive elements (modules) of system, which quickly prototype learning courses, are:

- information recourse manager,
- tools of examination and editing of each type of information data,
- the LOM description module (generating and recognizing).

The system is realized upon the NET platform.

Information recourse manager.

Information recourse manager is a tree-like structure, where the "root" elements are the courses parts. This structure of data presentation helps to achieve clear and exact results. Each manager's element is an offspring of basic class, in which main parts and elements are given.

For each type of recourse means and methodology of examination and operation of its elements are also determined.

The LOM description module.

The LOM description module is used to describe different information recourses according with international LOM and SCORM standards. These standards are realized in filling in fixed number of fields.

The main functions of this module are

- generating of meta-description of information resources,
- recornizing of resource's type according to it meta-description.

Organization of work in system.

The basic problem while the organization of work process is an operation of learning material elements, because it consists of a number of html-pages with different recourses (texts, multimedia, images). System does not create the course' design, so it is necessary to give an opportunity to course user to chouse his own strategy of data's editing. System only represents to the user the necessary course part, which contains elements of learning data. Then user himself decides how to design this material.

Conclusion

As a result of operation with designed system, the developer of on-line courses possesses an adaptive learning course, which contains sum of ontologies. The lowest ontologies level is informational data of different types (text, images, multimedia, etc.). These materials correspond to international standards and can be borrowed from existing learning courses of others developers. So developers can simply generate there own courses. The adopting process under the needs of certain user occurs throe the correlation of each ontology to the certain user's model, which is given in meta-description.

Bibliography

[Гаврилова, Хорошевский, 2001] Гаврилова Т.А., Хорошевский В.Ф. Базы знаний интеллектуальных систем: учебник. СПб.: Питер, 2001.

[Gavrilova, Geleverya, 2002]. Гаврилова Т.А., Гелеверя Т.Е. Программный инструментарий Vita. Версия 2.1. Техническая документация, 2002.

[Brusilovsky, Sosnovsky, 2005] Brusilovsky P., Sosnovsky S., etc. Interactive Authoring Support for Adaptive Educational Systems, 2005.

[Clarebout, Elen, 2005] Clarebout G., Elen J. The Effects of Pedagogical Agent in an Open Learning Environment, 2005.

[Hage, Aimeur, 2005] Hage H., Aimeur E. Exam Question Recommender System, 2005.

[Jovanovic', Gasevic', 2005] Jovanovic' J., Gasevic' D. Ontology of Learning Object Content Structure, 2005.

[LOM] LOM v1.0 Base Schema Draft Standard for Learning Object Metadata, IEEE 1484.12.1-2002, 15 July 2002.

[SCORM] ADL SCORM Version 1.3 WORKING DRAFT 0.9, November 27, 2002.

http://ltsc.ieee.org/

http://www.adlnet.org/

http://www.imsproject.org/

Authors' Information

Olga Malinovskaya – graduate student of Saint-Petersburg State Polytechnical University, Politechnicheskaya 29, 195251, St. Petersburg, Russia, e-mail: malinol@mail.ru

Tim Geleverya – senior developer, "GRANIT" company; e-mail: gt@csa.ru

Tatiana Gavrilova – Professor Saint-Petersburg State Polytechnical University, Politechnicheskaya 29, 195251, St. Petersburg, Russia, e-mail: gavr_csa@rambler.ru

Vladimir Golenkov – Head of AI Department of Belorussian University of Informatics and Radio-electronics, 220027, Brovki str.,6, Minsk,Belorussia, e-mail: golen@bsuir.unibel.by

A DOMAIN-INDEPENDENT OTOLOGY-BASED APPROACH TO REPRESENTATION OF COURSEWARE KNOWLEDGE

Eleonora Stoyanova, Polina Valkova, Irina Zheliazkova

***Abstract**: This paper proposes an ontology-based approach to representation of courseware knowledge in different domains. The focus is on a three-level semantic graph, modeling respectively the course as a whole, its structure, and domain contents itself. The authors plan to use this representation for flexibie e- learning and generation of different study plans for the learners.*

***Keywords**: domain-independent, ontology-based approach, semantic graph, courseware, knowledge representation*

***ACM Classification Keywords**: Computer Uses in Education, Knowledge Representation Formalisms and Methods*

Introduction

The term “ontology” has been put in use in knowledge engineering (KE) since the beginning of the 90’s in order to stress on the need and benefits of an explicit conceptual structuring of domain knowledge [Kellett & Marshall, 1990], [Decker et al., 1999]. In its simplest form this key idea was proposed earlier by [Dovgialo & Visotcki, 1983] for the needs of Computer-Aided Learning (CAL). The authors suggested a top-down strategy for development and implementation of courseware including mainly background material, structuring in topics with corresponding pretests and posttests. According to the semantic networks formalism the type of the relationships between topics, e.g. on the level of course is "a-part-of". At the same time and independently [Sato & Chimura, 1984] proposed to represent the structure of the taught course by means of a binary matrix, corresponding to a directed graph whose nodes stand for the topics and arcs for their input/output relationships. The idea of structuring the domain knowledge on two levels respectively course (global) and topic (local) belongs to [Peachey & MacCalla, 1983]. They proposed to represent the input ("preconditions") and output ("expected results") relationships of the topics by means of AND/OR/NOT operators.

The ontology-based approach was transferred from CAL systems to Knowledge Management CAL (KMCAL*)* systems where it was extended using mainly different kinds of networks (feature, semantic, augmented, and so on). The term KMCAL was used for the first time by [Webb, 1988] as a result of a comparative analysis of the frame-oriented CAL systems and Intelligent Tutoring Systems (ITS). He introduced the feature networks (FN’s) for knowledge representation on the topic level. Webb is a co-developer of the KMCAL system ECCLES based on this intelligent formalism and aimed to enhance English language learning. He pointed out seven teaching modes (interrogative, test, declarative, immediate, constructive, reference, and mixed) derived from the FN’s representation. The author also pointed out the restrictions of the FN’s structure making this formalism not so appropriate on the course level. More recently a domain-independent KMCAL system MICROKURS had been developed for teaching/learning declarative knowledge in different domains. On the local level this system is based on priority semantic networks [Marinov & Zheliazkova, 2005] and on the global level - on a set of input/output relationships between the course topics.

The ontology-based approach to modern domain-oriented design environments (DODEs) follows the contributions in cognitive and pedagogy sciences. The researchers in this area are interested in methods for solving practical tasks and in specialized tools facilitating the capturing, structuring, sharing and reusing of domain knowledge. This complex problem has been partly solved by development of knowledge repositories called ontological servers (*Ontolingva, Ontobroker, KA2* and other projects) with the static domain knowledge for reuse [Decker et al., 1999]. The efficiency of this kind of servers is more perceptible in case of theoretical courses with complex learning objects such as algorithms, context-free grammars, neural networks, final state-machines, and so on. DODEs are seen by [Fisher & Schvarff, 1998] as environments for support of self-directed learning. Gamelan, Educational Object Economy, and Agent Sheets Behaviour Exchange are three web-based DODEs supporting the learning needs of a community of software developers. Short analyses of the tools for support of

the domain ontology can be found in [Stefanov, 2003], where the computing education ontology for the needs of the project DIOGENE is also described. This ontology consists of one class (computing education field), a lot of instances (terms from the computing domain) and a predefined set of relations, called respectively "*has-part*", "*requires*", and "*suggested-order*".

Nowadays the course author has to work as a knowledge analyst and engineer making visible the course curriculum and conceptual structure of the taught domain. For the needs of representation and delivering the curriculum contents through INTERNET a hybrid approach has been proposed by [Fung, 2002]. Although it is demonstrated on the basis of an introductory course in Mathematical Statistics the approach is domain-independent. It is based on a combination of standard directed graphs and modified state-charts. Except the hierarchical relationships, e.g. type "a-part-of" and "type-of", two other types can be used, called respectively "*contained-in*" and "*applied-to*". The author focuses on the opportunity for generating different study plans with depth-first, top-down, bottom-up, and breath-first strategy to suit the different students and delivering the appropriate teaching multimedia material.

In task-oriented environments (TODEs) the ontology-based approach using a visual language facilitates and enhances both adaptive individualized planning and flexible self-directed learning. Once a course is represented in such a way different study plans can be generated to suit the different learning styles and flexible learning can be ensured that suits the learner's interests, needs and preferences. Additionally in TODEs the ontology design is also used as a procedure for precise assessment and diagnostics of the learner's knowledge relatively to the author's one [Zheliazkova, 2005].

In this paper a domain-independent three-level semantic graph is proposed to model courseware knowledge. This model is close to the multi-hierarchical directed graph model described by [Otzuki, 1990] for the needs of KE and to the discrete Markov chain model of [Jelev, 2004] for the e-learning assessment in higher education. Each level of the proposed model is commented and demonstrated with elements of an example courseware in a separate section. The courseware specially created for research purposes by means of MACROMEDIA FLASH 5.0 is in the domain of Programming Environments (PEs) and the working language is Bulgarian. The Conclusion outlines the authors' contributions and future intentions.

First Level Semantic Graph Model

The first level of the semantic graph model concerns the courseware as a whole (fig.1).

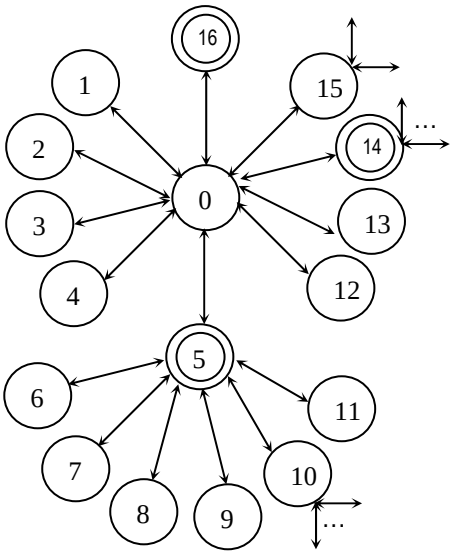


Fig. 1. First level semantic graph model

Index	Dialog state
0	Courseware home page
1	Institution home page
2	Department home page
3	Author's home page
4	Teacher's home page
5	Initial course page
6	Curriculum timetable
7	Curriculum annotation
8	Curriculum contents
9	Technology of teaching
10	References
11	Discussion forum
12	Search machine
13	Frequently asked questions
14	Dictionary
15	Courseware map
16	Home pages of similar courses

Table 1. Decoding the node

The model has the form of two connected stars, where all arcs are type ‘a-navigation-link” and the node indexes are decoded in table 1. The frequently used links are to/from the pages of: the members of the authors’ team, elements of the curriculum, standard tools for human INTERNET communication, similar courses, and so on. Node 10 (References) has a set of external links from/to some nodes on the third level of the model representing the domain knowledge itself. Node 15 (Courseware map) has links to/from nodes belonging to the three levels as well as opportunities for downloading the background material, pretests, and exercises (fig. 2). Double-circle nodes (5, 14,16) stand for subgraphs. The need of a terminological dictionary (node 14) can be motivated by the different meanings one and the same word may have in different domains. For example, the word “frame” in Psychology means a kind of spirit, in Mechanics skeleton of construction, in CAL screen content, in Artificial Intelligence an information structure for knowledge representation, and so on. Although English has become the main official language all over the world it’s better for non-native speaking English learners to use their native language as working one. That’s why the dictionary has to be multilingual.

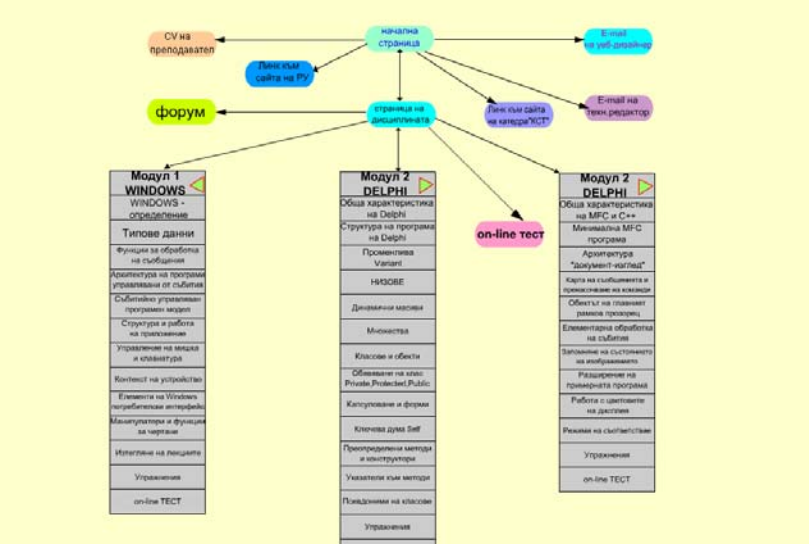


Fig. 2. PEs courseware map together with modules and topics

Second Level Semantic Graph Model

A detailed analysis of the standard curriculum for engineering education at the University of Rousse has shown that besides explicit hierarchical links of type “topic-subtopic” such a curriculum can contain several types of implicit links, such as: “topic-test”, “topic-exercise”, “test-test”, and “exercise-exercise”. In the most common case the cardinality of these types of relationships is 1:N. That means more than one taught topic can be tested with one test or applied to one exercise. Also more than one exercise (test) can serve as subexercise (subtest) in another exercise (test). The above-listed types of links can be found on fig. 3, presenting the subgraph node 5 from fig. 1. Here the following graphical primitives are used in addition: an ellipse for a subtopic, a triangle for a test, and a square for an exercise. Fig. 4 presents a part of the PEs hierarchical structure viewed in Notepad as a standard text file.

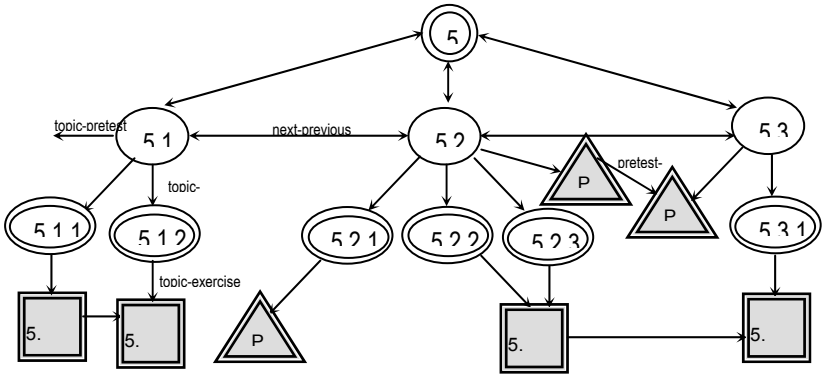


Fig.3. Second level semantic graph model

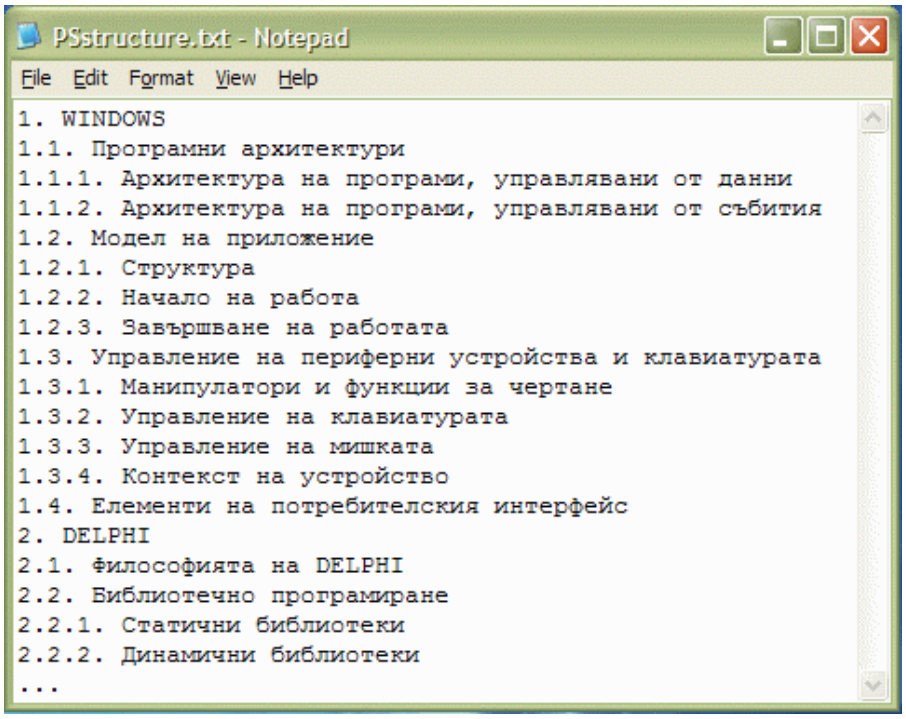


Fig. 4. A part of the PEs hierarchical structure viewed in Notepad

Third Level Semantic Graph Model

On the third level the semantic graph model represents the domain contents itself. Actually, three disconnected semantic graphs respectively for a topic (lecture, lesson), test, and exercise serve as formalism for this purpose.

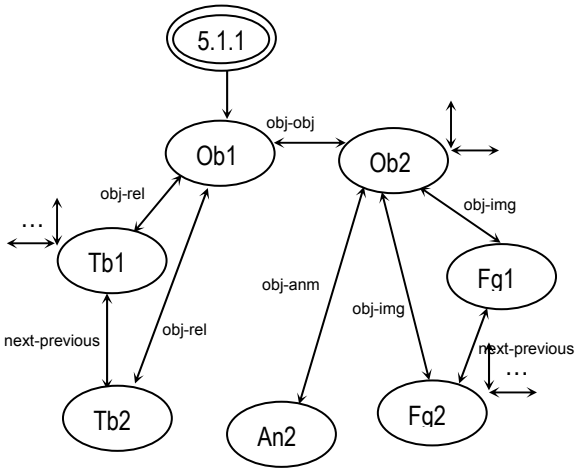


Fig. 5. An example semantic graph of a topic

On fig. 5 an example semantic graph corresponding to the ellipse terminal node 5.1.1 from fig. 3 is shown. Here nodes Ob1 and Ob2 stand for the text description of two learning objects, nodes Fg1 and Fg2 for two different images of Ob2, node An1 for an animation of the same object, and nodes Tb1 and Tb2 for two different relations of Ob1. A text description page from a topic from the PEs lecture is shown on fig. 6 together with links to several nodes on the third and first levels.

Fig.6. A text description page with links to 3 nodes on third level and 8 nodes on first level

The semantic graph on fig. 7 corresponds to the triangle subgraph node P1 from fig. 3. Note, that the goal of the test is not only to assess the learner's understanding of the background knowledge, but also to fill in the missing knowledge and correct wrong one, e.g. the test plays a teaching function too.

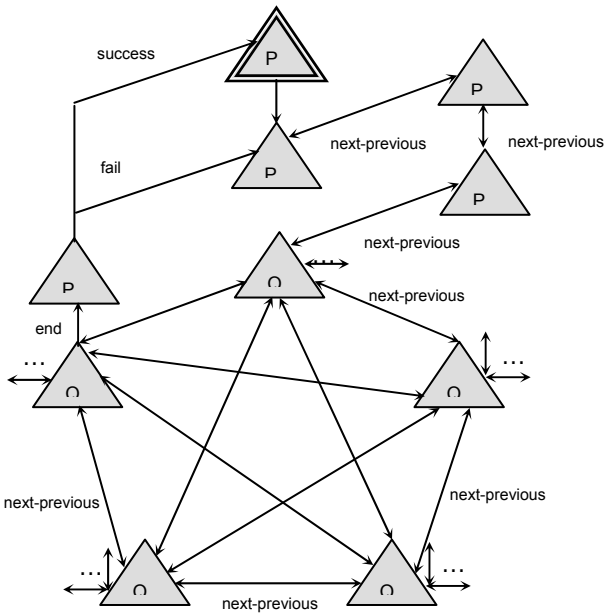


Fig. 7. The semantic graph model of an on-line test

Here the node AP represents the administrative information about the test as a whole, e. g. organization, department, author, goal; DP - key directives for the teacher's intervention; CP - criteria for the learner's assessment; Qi (i =1,…,N) - questions. The bi-directional links between nodes-questions of type "next-previous" signify that the learner has free access from the current question to all other questions. External links from/to each question correspond to the related pages with the topic material accessible by means of the key teacher's directive "Help". A fragment of an on-line test covering the module WINDOWS of the PEs courseware is given on fig. 8.

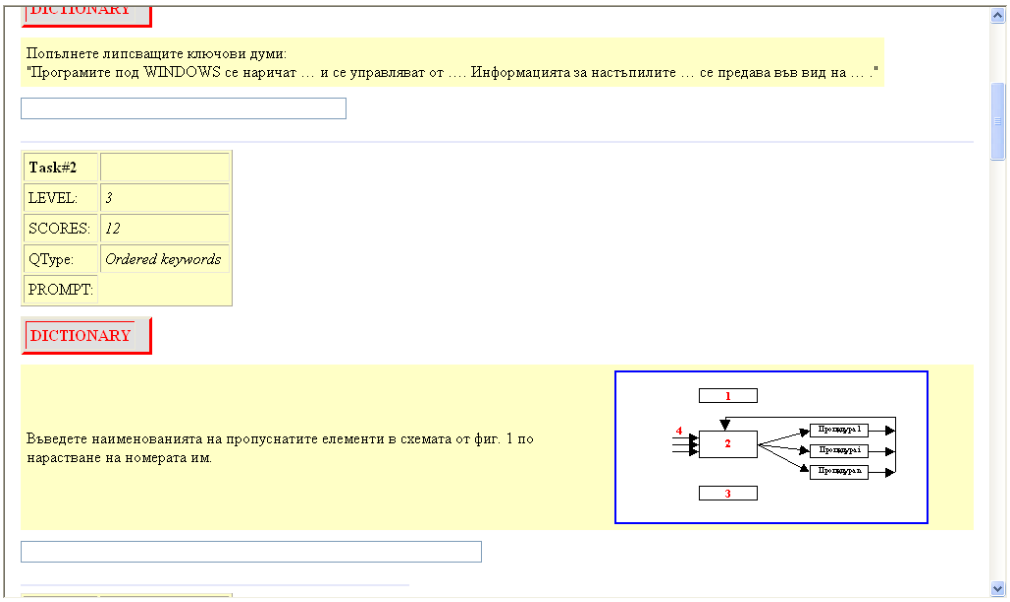


Fig. 8. A fragment of the on-line test covering the module WINDOWS

When the learner's test result is lower than a given threshold he/she has to perform the test again (the feedback "fail"). Otherwise the learner is allowed to start the related practical exercise (the feedback "success").

The semantic graph model corresponding to the square subgraph E1 from fig. 3 is presented on fig. 9. Fig. 10 presents a fragment of downloaded WORD document with exercise instructions. In on-line mode the learner can read the exercise information about: administrative data (EA), directives for the teacher's intervention (ED), and the assessment criteria (EC). Node TFi corresponds to the i-th task formulation, and the next several nodes TSj (j=1,M*)* for skills for solving the task, extracted step by step. Links "next-previous" between nodes stand for the suggested order of the corresponding pages, and the external links mean the learner has access to the related pages with the topic material.

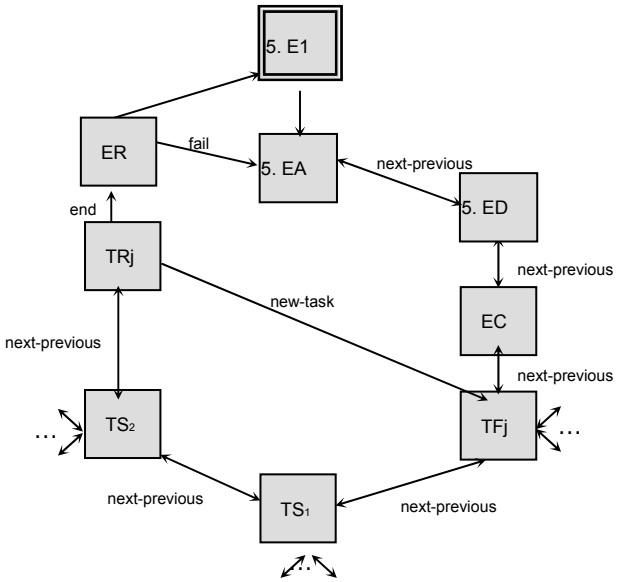


Fig. 9. The semantic graph model of an on-line exercise

A report with a final mark about the learner's task performance is generated (node RTi). The feedback to TFi corresponds to the learner's moving on to the next task. An exercise report (node ER) is received. If the learner's mark is lower then a given threshold (link "success') he/she has to perform the exercise again (the feedback "fail"). Otherwise he/she is allowed to continue with the next topic.

Fig. 10. A fragment of a downloaded WORD document with exercise instructions

Conclusion

A domain-independent ontology-based approach to courseware knowledge representing has been proposed. The three levels of the semantic graph model correspond to the course as a whole, structure and domain knowledge itself. There are three disconnected semantic graphs respectively for a topic, test, and exercise on the third level. The practical usefulness of the approach has been demonstrated on a courseware from the field of software engineering education. The future authors' efforts will be focused on implementation an on-line environment for courseware generation based on the described semantic graph model.

Bibliography

[Chimura & Sato, 1984] Chimura H., Sato T., Computer-Assisted Analysis and Determination of Instructional Sequences, *IEEE Trans.*, Vol. 3, 1984, pp.136-144.

[Decker et al., 1999] Decker S., Erdmann M., Fensel D., Studer R. ONTOBROKER: Ontology Based Access to Distributed and Semi-structured Information, *Proceedings of Eighth Working Conference on Database Semantics, Semantic Issues in Multimedia Systems,* Meersman R., Tari Z., Stevens S., (Eds.), Kluwer Academic Publishers, Boston, 1999.

[Dovgyalo & Visotski, 1983] Dovgyalo A. M. & Visotski I. I. A Top-Down Methodology for Development and Implementation of Courseware. Journal of Cybernetics, 1983, pp. 103-108 (in Russian)

[Fisher & Scharff, 1998] Fisher G., Scharff E. Learning Technologies in Support Self-Directed Learning, *Journal of Interactive Media in Education*, Vol. 98, No. 4, October, 1998, www-jime.open.ac.uk/98/4.

[Fung, 2002] Fung I. P. A Hybrid Approach to Represent and Deliver Curriculum Contents, http://www10.org/cdrom/papers/207/node3.html

[Jelev, 2004] Jelev G. A. Strategy for E-learning in the Faculty of Computer Systems and Control in TU-Sofia. Proceedings of the International Conference on Computer Science, 7 December, Sofia, Bulgaria, pp. 263-267.

[Kellett & Marshall, 1990] Kellett J. M., Marshall G. Knowledge Engineering: Tools and Techniques. Artificial Intelligence in Engineering. Ed. G. Winstanley. John Wiley & Sons, England, 1990, pp.119-149.

[Marinov & Zheliazkova, 1992] Marinov M. T., Zheliazkova I. I. An Interactive Tool Based on Priority Semantic Networks, *Int. J. of Knowledge-Based Systems,* No.18, 2005, pp.71-77.

[Osuga & Saeki, 1990] Osuga S., Saeki I. Knowledge Acquisition, (Eds.), Moscow, Mir, 1990 (in Russian).

[Stefanov, 2003] Stefanov K. Computing Ontology Creation. Proceedings of the International Conference on Computer Systems and Technologies. Sofia, Bulgaria, 14-16 June 2003.

[Peachey & McCalla, 1986] Peachey D. R., and McCalla G. I. Using Planning Techniques in Intelligent Tutoring Systems. *Int. J. Man-Machine Studies*, Vol.24, 1986, pp.77-98.

[Webb, 1988] Webb G. I. A Knowledge-Based Approach to Computer-Aided Learning, *Int. L. Man-Machine Studies*, Vol. 29, 1988, pp. 257-285.

[Zheliazkova, 2005] Zheliazkova I. I. From ITSs Through VLEs Towards DODEs and TODEs, Proceedings of the *National Conference on E-learning in Higher Education,* 3rd –5th June, Kitten, Bulgaria, 2004, pp. 21-25.

Authors' Information

Eleonora Stoyanova – e-mail: eleonorarymenova@abv.bg

Polina Valkova – e-mail: valkova_99@yahoo.com

Irina Zheliazkova – e-mail: irina@ecs.ru.acad.bg

University of Rousse, Rousse – 7017, Studentska street 8, Bulgaria

MODELS, TECHNIQUES AND APPLICATIONS OF E-LEARNING PERSONALIZATION

Desislava Paneva, Yanislav Zhelev

***Abstract**: In recent years Web has become mainstream medium for communication and information dissemination. This paper presents approaches and methods for adaptive learning implementation, which are used in some contemporary web-interfaced Learning Management Systems (LMSs). The problem is not how to create electronic learning materials, but how to locate and utilize the available information in personalized way. Different attitudes to personalization are briefly described in section 1. The real personalization requires a user profile containing information about preferences, aims, and educational history to be stored and used by the system. These issues are considered in section 2. A method for development and design of adaptive learning content in terms of learning strategy system support is represented in section 3. Section 4 includes a set of innovative personalization services that are suggested by several very important research projects (SeLeNe project, ELENA project, etc.) dated from the last few years. This section also describes a model for role- and competency-based learning customization that uses Web Services approach. The last part presents how personalization techniques are implemented in Learning Grid-driven applications.*

***Keywords**: Adaptive Learning Content, Customized learning, Grid technologies, Learning Management Systems, Personalization, Simple Sequencing, Web Services.*

***ACM Classification Keywords:** H.3.7 Digital Libraries- User issues, H.3.4 Systems and Software- User profiles and alert services, H.3.5 Online Information Services- Web-based services, K.3.1 Computer Uses in Education – Distance learning.*

Introduction

Personalised learning presupposes high quality teaching that is adaptive to the different ways students achieve their knowledge and skills. Therefore, the teaching courses, curricula, and school organisations have to be designed in a way to reach as many students as possible with diverse needs and experiences for as much of the time as possible. Personalised courses actively engage the learners by providing teaching strategies and materials that appeal to the learners' knowledge and preferences etc. Since it would be costly and unfeasible for teachers to produce personalised courses that meet all of these requirements, the LMSs are of prime importance for education. Such systems allow for delivering information outside the traditional bound of a classroom situation, where learners are taught by a static one-fits-all approach. An educational system that responds to individual needs by creating a personal learning path enables individual students to experience excellence in his or her learning. Analytical study of key functional LMSs requirements such as adaptability, personalization, modality, possibility for record-keeping on student's performance, and usage statistics for the system as a whole has been done in [Pavlov et al.'04].

The personalization includes how to find and filter the learning information that fits the user's preferences and needs, how to represent it and how to give the user tools to reconfiguration the systems, in consequence, reconfiguration system could be part of personalized environment in some systems. The user modelling is the process of constructing (often computer-based) users models, background knowledge and users behaviour

representation, while the user model means all the information collected about a user that logs to a web site, in order to take into account her needs, wishes, and interests. Every LMS has its techniques to modelling his users so as to construct the user model or profile.

In general the personalization might be examined in the following aspects [Zheleva'05, Graziano et al.'03, Gibson et al.'02]:

- Personalization of the learning content, based on learner's preferences, educational background and experience, learning content tailored to individual learning style of the user;
- Personalization of the representation manner and the form of the learning content (for example, learning content in the form of the adaptive learning sequences of learning objects).
- Full personalization, which is a combination of the previous two types.

The following approaches can be used to apply the learning personalization:

- Personalization, controlled by the learner – It requires direct input of the learner's needs and preferences by filling question forms or by choosing options and alternatives.
- Personalization, based upon an existing user profile and meta-descriptions of the information content - In this case, the learners' preferences are stored in their profile.
- Personalization via searching for a correlation between the learners - Correlation is through the values of the attributes, describing the learner's profile. If there is a strong correlation, there is a possibility that the content for a given profile is suitable for applying to its close (adjacent) profiles.

Personalization in current Learning Management Systems tends to be concerned with remembering which courses the user is allowed to view and how they like their pages to be presented. In some cases users are able to edit their own profile; to maintain their personal calendar which keeps track of their event transactions; to subscribe to forums, etc. Observing the educational process as a whole, learners are very rarely allowed to get access to learning objects which are conditioned on a wide range of personal data including achievement, date/time and class code. In [Paneva'05] the author gives an overview of several methods for implementing personalization, which are exploited in several widely used LMSs in the recent years.

Learning Modelling and Profiling

The student model enables the system to provide individualised course contents and study guidance, to suggest optimal learning objectives, to determine students' profiles and the actual knowledge they have acquired, to dynamically assemble courses based on individual training needs and learning styles, and to join teachers able to provide support in terms of guidance and motivation and therefore to help the students with different backgrounds and knowledge levels to achieve their learning goals effectively on the Web.

The software developers face a number of challenges and difficulties when trying to model student profile and activities on real eLearning systems. The process of collecting student modelling data is time-consuming and requires the development of complex data structures to represent student's personal information, knowledge and behaviour in the learning domain. Once student data is collected, it must be converted into a format compatible with knowledge representation and reasoning systems to function as the input for the adaptive systems. Faced with these requirements, student modelling data is often stored in proprietary, hard-to-access formats that don't encourage reuse or distribution. In addition, in most cases the student models can only be used with the learning application, which it was developed for and when the application is changed or replaced they will be useless.

The student model needs to cover a certain amount of information that can be divided into two main groups:

- general student information such as learning goals, cognitive aptitudes, measures for motivation state, preferences about the presentation method, factual and historic data (personal information), etc.,
- information about student's behaviour in the learning domain such as overall competence level for the course, module competence level, concept competence level, module study time, test solving status, etc.

Naturally, student models "do not have to fully account for all aspects of student behaviour. In fact, we are interested in computational utility rather than in cognitive fidelity" (Self, 1990).

Learner model standards

The standards related to user model definition and representations are two:

- IEEE Public And Private Information (PAPI) **-** It specifies both the syntax and semantics of a 'Learner Model,' which will characterize a learner and his or her knowledge/abilities. [PAPI '02]
- **I**MS Learner Information Package (LIP) **-** It can hold information about the learner, including his progress and received awards. [IMS LIP '01]

In [Paneva'05] the author gives an overview of these main learning model standards.

Methodology for Development and Design of Adaptive Learning Content

The term "adaptive learning" means the capability to modify any individual student's learning experience as a function of information obtained through their performance on situated tasks or assessments. With the integration of the IMS Simple Sequencing Specification [IMS SS'03], SCORM [SCORM] allows the learning strategies to be translated into sequencing rules and actions, which are associated with the activities a learning experience consists of. The sequencing rules are based on learner's progress and performance and affect the availability of the learner is allowed to experience.

All learning activities can be associated with sequencing information defined by the content author. In run time, each activity experienced by the learner is associated with tracking status data, which may affect the overall sequencing process. This means that learners with difficulties in satisfying the learning objective should be able to experience additional activities (or repeat some of the activities) to improve their knowledge level and skills. Some restrictions concerning number of attempts and/or period of time for any activity could be set by the content author.

The process of defining a specific sequence of learning activities begins with the creation of a learning strategy for the achievement of the determined pedagogical aim/s. Learning strategy specifies types of learning activities and their logical organization (the activity tree) as well as the prerequisites and expected results for each activity. The rules for managing the instructional flow are the other important part of the strategy. Describing the rules by means of IMS SS elements and attributes the content author transforms the sequencing strategy into strategy for the activity tree traversal management. The author establishes an aggregation of learning objects associating leafs of the activity tree with appropriate Sharable Content Objects (SCOs). The outcome of this process is a content package. The imsmanifest.xml file of the package describes SCOs organization and their sequencing. The implementation of adaptive learning in given eLearning environment could be promoted and facilitated by providing of sequencing templates for the development and design of instructional flows.

The sequencing template describes the conceptual organization of the learning content as a sequence of template pages and provides the learning strategy implementation translating it into sequencing strategy. Such sequencing template can be used in different knowledge domains from different instructors who want to follow the described in the package content organization and the implemented learning strategy. In this case, instructor is responsible only to identify (or create) and then to incorporate the relevant multimedia content in each of the template pages accordingly the subject matter of the course taking into consideration the concrete learning objectives and context.

The main advantage of the Simple Sequencing approach is that the sequencing rules are described outside the learning objects' content. In this way, the instructional designer can change the rules (i.e. the learning strategy) without any changes in the content or its organization. Nesting manifests of the developed sample packages the content author can developed more complex strategies and content structures. The main disadvantage of the methodology is that selected strategy cannot be changed dynamically in time of learning.

Innovative Services for Learning Personalization and Customization

The future trends in ubiquitous learning point to the investigation and development of specialized learning services, methods and instruments allowing wide range of learners to access and to follow courses by web-based tools and Digital Video Broadcast tools at training institution and/or on workplace, or at home, combined with the

practically ubiquitous connectivity of the mobile devices. Following this idea we extract several innovative personalization services that are suggested by several very important research projects (ELENA project [ELENA], LOGOS project [LOGOS], SeLeNe project [SELENE], etc.) dated from the last few years.

Personalization service - It can be recognized as a functionality, which customises access to learning services and learning resources (in the context of the delivery of a learning service) based on learner profiles (career development plans can even be part of such a profile). The result of the personalization service is usually a customized view on a learning repository or a learning management network (connecting various educational nodes that facilitate the provision of additional educational services). The customization can be performed in many ways using techniques such as collaborative filtering or rule-based personalization in order to modify a user's query or to reduce the results produced by the query. A special personal learning assistant (PLA) can support learners in searching for, selecting and contracting learning services. PLAs can also trigger the delivery of the following services:

- *Query rewriting service* **-** The query rewriting service extends a user query by additional restrictions, joins, and variables based on various profiles. This extension is performed based on heuristic rules/functions maintained by the query rewriting service. Query rewriting services can be asked for adding additional constraints to user queries-based on user preferences and language capabilities. They can also be asked to extend a user query based on previous learner performance maintained in learner profiles, if a query is constructed in the context of improving skills. Query rewriting services can also be asked to rewrite a user query based on information the connected services need.
- *Recommendation service* **-** The recommendation service provides annotations for learning resources in accordance with the information in a learner's profile. These annotations can refer to the educational state of a learning resource, the processing state of a learning resource, etc. The service holds heuristic rules for deriving recommendations based on learner profile information. Recommendation services can be asked to add recommendation information to existing instances based on learner profile information.
- *Link generation service* **-** A link generation service provides (personalized) semantic relations for a learning resource in accordance with the information in a learner's profile. These relations can show the context of a resource (e.g. a course in which this learning resource is included), or they can show other learning resources related to this resource (e.g., examples for this learning resource, alternative explanations, exercises). The link generation service holds heuristic rules for creating semantic hypertext links. Some of the rules refer to information from the learner profile, in absence of learner profile information the service can at least provide some, not optimized, hypertext links. Link generation services can be asked for adding links and link type annotations to a given learning resource. They can be asked to generate a context for a given learning resource, or to generate a context for several learning resources by adding hyperlinks between them. They can be asked also to generate a learning path.

Customized learning, presenting just the right material to the learner on demand, can be described using data representations from learning technology standards (learner profiles, competency definitions, sequencing rules, learning objects). William Blackmon and Daniel Rehak [Blacmon et al.'03] offer a web services-based methodology for customization by profile, specifically one of eliminating LOs from a course because either:

- Learner's current role does not require the learning objective taught by the LO, or
- Learner's profile indicates the learner has already achieved the objective taught by a LO.

The learning content and data used in customization are represented in a set of standards-based data models. These are used in a content authoring and delivery process that customizes the activities delivered to the learner based on the learner's role and competencies [IMS Competency Definition, IMS LIP].

For content and learning activity customization are used six sets of data elements:

- Learning Objects - the collection of content and learning resources maintained in a content repository.
- Content Structure - the organization of learning objects in a tree or hierarchical structure.
- Roles - definitions of the job roles of a learner.
- Competency Definitions - definitions of the skills and knowledge acquired by a learner.

- Learner Information Package - the collection of stored profile information about a learner.
- Sequencing -- rules used to select content and sequence the learner through a content structure.

The major steps for a customized course preparation and delivering are[3]:

- Create Course and Content Description -- describe the course (content structure and set of LOs) and behaviour rules used to express the progression of the learner through the content:
  - Associate role and competency definitions with each learning object by mapping a sequencing objective id (used to label the objective) to a competency definition id or to a role id.
  - Specify the conditional rules used to customize the course by eliminating learning objects from the activity sequence.
- Establish Learner Profiles -- specify the role of the learner (which in turn may yield a set of competencies required to perform the role), and contain data on the learner's record relative to each of the specified competencies.
- Register Learners -- register the learner for the course.
- Deliver Course -- deliver the course, matching the course description to the learner's profile to select content. As the learner completes instruction, the profile may be updated to include mastery of subject matter. Delivery and customization continues until all required activities have been completed.

The customization process has been implemented through a set of web services. Rather than building large, closed systems, the focus is on flexible architectures that provide interoperability of components and learning content, and that rely on open standards for information exchange and component integration. The overall web services architecture for learning is divided into layered services. The layers from top to bottom in this services stack are:

- User Agents -- provide interfaces between users (both end user applications and program agents) and the learning services. Agents provide the major elements of learning technology systems: authoring of content, management of learning, and actual delivery of instruction to learners.
- Learning Services -- collection of (many small, simple) data models and independent behaviours. Service components are characterized as providing a single function that implements a particular behaviour. Each service is identifiable, discoverable, (de)referenceable, and interoperable. They include built-in security and rights management, and assume an unreliable underlying network. Services are grouped into logical collections, where upper-level services rely on the support from the lower-level services:
  - Tool Layer - Tools provide high-level, integrated server applications. Accessed via known, published interfaces, they provide the public interface to the learning tools (tutors, simulators, assessment engines, collaboration tools, registration tools, etc.). User agents and end user applications are built using collections of tool services.
  - Common Applications Layer - These are services that provide the commonly used learning functions and application support behaviours used by tools and agents (sequencing, managing learner profiles, learner tracking, content management, competency management, etc.).
  - Basic Services Layer - Basic services provide core features and functionality that are not necessarily specific to learning, but which may need to be adapted for learning (storage management, workflow, rights management, authentication, query/data interfaces, etc.).

All services are built on and use a common infrastructure model. The infrastructure layer relies on basic Internet technologies (e.g., HTTP, TCP/IP) to connect service components over the network. The services themselves are implemented using web services bindings. Messaging is done with SOAP; service descriptions are catalogued with UDDI, and described in WSDL - all are XML representations [Samtani et al.'02]. Overall service coordination is expressed in a workflow or choreography language. These standard technologies permit the upper-level services to be implemented in a platform-neutral manner, and provide interoperability across different implementations of the actual learning services.

[3] Assuming there is a globally defined set of learner job roles and competency definitions

Grid technologies. Personalization in Learning Grid-driven applications

The utilisation of currently available communication and information technologies has turned traditional location based education into location independent one. Nowadays, learning is equivalent to searching for sources and selecting the appropriate source to study from. The multitude of sources available on the Internet makes the selection of the appropriate source a rather difficult task. Learners need to access large volumes of data, most times distributed in many locations. Learners also need a variety of services available on demand that can be used and accessed from their environment to satisfy their learning needs. All of the above can be enabled by the utilisation of grid technologies.

Grid is a modern technology for the flexible, secure and coordinated sharing of distributed resources and data. Grid technologies define a new powerful computing paradigm where the customer of the grid will be able to use his or her private work place (Workstation, PC, UMTS phone....) to invoke any application from a remote system, use the system best suited for executing that particular application, access data securely and consistently from remote sites, exploit multiple systems to complete complex tasks, or use multiple systems to solve large problems that exceed the capacity of a single one. Another interesting aspect of grid technologies is their support for resource sharing and problem solving in dynamic, multi-institutional virtual organizations. In this vision, the sharing does not mean simply exchange of data or files but rather a concrete access to resources. This “sharing capability” imposes the definition and implementation of well-defined resource management policies to specify what is accessible, from whom and under which conditions.

The philosophy and approach behind Grid technologies [Hsing-Chuan et al.’04] show the right characteristics for achieving an effective learning. Indeed, they allow to access and integrate the different technologies, resources and contents that are required in order to realise new paradigms in eLearning. They are the most promising approach to realise an infrastructure that will allow learning process actors to collaborate, to take part in realistic simulations, to use and share personaliselly high quality learning data and to innovate solutions of learning and training. Grid will be able to support learning processes allowing each learner to use, in a transparent and collaborative manner, the resources already existing on-line, by facilitating and managing dynamic conversations with other human and artificial actors available on the grid, etc.

A high quality example of personalization techniques implementation, based on grids is demonstrated in SeLeNe (Self eLearning Networks). This project was funded as an EU FP5 Accompanying Measure (IST-2001-39045) running from 1st November 2002 to 31st January 2004. SeLeNe was part of action line V.1.9 CPA9 of the IST 2002 Work Programme, contributing to the objectives of Information and Knowledge Grids by allowing access to widespread information and knowledge, with eLearning as the test-bed application. The developers conducted a feasibility study into using Semantic Web technology for syndicating knowledge-intensive resources (such as learning objects) and for creating personalized views over such a Knowledge Grid.

A self e-learning network consists of web-based learning LOs that have been made available to the network by its users, along with metadata descriptions of these learning objects and of the network’s users. The architecture of the network is distributed and service-oriented. The personalization facilities include: querying learning object descriptions to return results tailored towards users’ individual goals and preferences; the ability to define views over the learning object metadata; facilities for defining new composite learning objects; and facilities for subscribing to personalised event and change notification services.

Summary

One of the main goals of contemporary eLearning is the possibility for learning adaptation to be assured for each learner in respect to her/his necessities, preferences, needs, performance, and progress. The achievement of interoperability and content reusability in the existing diversity of software and hardware platforms is a real challenge. One big limitation of the web-based interaction is the smaller communication bandwidth than traditional face-to-face interaction. The term bandwidth represents the amount of information that can be transferred in a unit of time through any means possible. In the face-to-face communication mode, if a verbal instruction is not understood, the clue can be available to the counterpart through gestures, group dynamics and other such means, but the clues in the web-based mode are not always so clear and in many cases not available at all. Therefore, tailoring the information to the right-level for the receiver to understand and integration of different appropriate methods for learning adaptation are crucial factors for the success of any LMS.

Bibliography

[Pavlov et al.'04] Pavlov R., Dochev D. (2004), New Information Technologies and Interactive Environments for Vocational and Life-long Learning, Analytical study, ICT Development Agency, Sofia.

[Zheleva'05] Zheleva M. (2005), Design and development of Intended Instructional Flows in Web-based Learning Environments, In: I.Simonics, R.Pavlov, T. Urbanova (Eds.) "Technology-enhanced Learning with Ubiquitous Applications of Integrated Web, Digital TV and Mobile Technologies", Proceedings of HUBISKA Open Workshop, 6th eLearning Forum, 9-10 June 2005, Budapest.

[Graziano et al.'03] Graziano A., Russo S., Vecchio V. (2003), Metadata-based Distributed Architecture for Personalized Information Access, In: Proceedings of the European Distance and E-Learning Network /EDEN/ Annual Conference "Integrating Quality Cultures in Flexible, Distance and eLearning", June15-18, Rhodes, Greece

[Gibson et al.'02] Gibson, D., Knapp, M., & Kurowski, B. (2002) Building responsive dissemination systems for education with the semantic web: Using the new open-source "liber" application, In:Proceedings of EdMedia 2002 conference, Montreal, Quebec

[Paneva'05] Paneva D., Some Approaches for Personalization in Learning Management Systems, In D.Dochev, R. Pavlov (Eds.) "e-Learning solutions – On the Way to Ubiquitous Applications", Proceedings of Joint KNOSOS-CHIRON Open Workshop, Sandanski, 26-27 May 2005.

[IMS SS'03] IMS Simple Sequencing Information and Behavior Model, Version 1.0, 2003.

[SCORM] Sharable Content Object Reference Model; http://www.adlnet.org.

[PAPI '02] PAPI Learner Specification, 2002, Available online: http://edutool.com/papi/

[IMS LIP '01] IMS Learner Information Package, 2001, Available online: http://www.imsproject.org/profiles/

[Blackmon et al.'03] Blackmon W. H., Rehak D. R, Customized Learning: A Web Services Approach, In Proceedings: Ed-Media 2003, Honolulu, Hawaii, USA, 2003

[IMS Competency Definition'02] IMS Reusable Definition of Competency or Educational Objective Information Model, Version 1.0 Final Specification, IMS Global Learning Consortium, 2002

[IMS LIP'02] IMS Learner Information Packaging Information Model Specification, Version 1.0 Final Specification, IMS Global Learning Consortium, 2002

[Samtani et al.'02] Samtani G., Sadhwani D. (2002), Web services and application framework working together, In: Journal "Web Services Architect", March, http://www.webservicesarchitect.com/content/articles/samtani04.asp

[Hsing-Chuan et al.'04] Hsing-Chuan Ho, Chao-Tung Yang, Chi-Chung Chang (2004), Building an E-Learning Platform by Access Grid and Data Grid Technologies, In: Proceedings of IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE'04).

[SELENE] SeLeNe project: Self e-Learning Networks, Available online: http://www.dcs.bbk.ac.uk/selene/

[ELENA] Elena project (final report), Available online: http://www.elena-project.org/images/other/D73FinalReport.pdf

[LOGOS] LOGOS Project, Available online: http://logosproject.com/

Authors' Information

Desislava Paneva - Institute of Mathematics and Informatics, BAS, Bulgaria, 8, Acad. G. Bonchev str.; e-mail: dessi@cc.bas.bg

Yanislav Zhelev – Burgas Free University, Center of Informatics and Technical Sciences, 62 San Stefano str., Burgas 8001, Bulgaria; e-mail: jelev@bfu.bg

PRODUCTION-FRAME MODEL OF REPRESENTATION OF DATA DOMAIN

Alexander Kuzemin, Darya Fastova, Olesya Dyachenko

***Annotation**: One of the most important problems of e-learning system is studied in given paper. This problem is building of data domain model. Data domain model is based on usage of correct organizing knowledge base. In this paper production-frame model is offered, which allows structuring data domain and building flexible and understandable inference system, residing in production system.*

Introduction

Learning as regulable process represents controlled knowledge caring from tutor (teacher, adviser, expert) to trainee (to the pupil, student or listener). There are a number of the theoretical and practical questions concerning adequacy of these technologies of learning, and also knowledge and skills control in conditions of organization of educational process with usage of training means, in particular, computer information technologies.

The great demand of new modern forms and transfer methods and control of knowledge is the practical proof of necessity of creation and introduction of corresponding intellectual computer training systems. Therefore, questions of development and perfection of their methodical, mathematical, algorithmic, program and software for realization of required functionalities, adaptations to specific trainee features take on vital importance.

Native and foreign scientists, such as D. A. Pospelov, V. V. Popov, I. N. Kuznetsov, B.A.Sazonov, I.V.Makarova, V.P.Tikhomirov, M.Minsky, D. Blum and others were engaged in problems of usage and development of technologies of representation and controlling of knowledge transfer. The analysis of works of listed authors on a researched question has revealed insufficient workup of mathematical models, methods and representation algorithms and knowledge transfer in the computer systems, interfering creation of the universal program systems based on knowledge.

Problem statement

E-learning systems consist of combination of three components: trainee model, learning process model and data domain model (a course, area of learning) [1].

The trainee model contains the information of his personalities, preferred to him strategies of learning, typical errors. The learning process model ensures formation of information model, presentation of the information and evaluation test of trainee potential. This model contains knowledge about planning and organization of learning process, about general and individual strategy of learning.

Data domain model (DDM) is defined by set of entities. Concepts or themes, to each of them corresponds unit of a teaching material which is not requiring (from the point of view of the teacher) partition on subtheme can be considered as entities of data domain (DD) in learning systems. Each theme is described by a set of parameters (attributes), which are important for learning process controlling. Thus, following demands are made to data domain model:

1. Possibility to reflect relationships between elements

2. Possibility of obtaining of a holistic image of knowledge

Creature of such data domain model presents a development of knowledge base (KB) - formal warehouse of addresses and types of entities. KB contains information reflecting regularities of given DD and permitting to predict and deduce new facts, which are not presented in DB.

Selection of knowledge representation language is the main problem in creature of KB. In this case frame system will be knowledge representation language, which meets the requirements made to DDM.

Object of the paper: creature of a data domain model which would allow forming the curriculum, taking into account personal trainee requirements.

Urgency. The given approach is based on usage of production-frame DDM in learning system. Such system can be considered as unified hierarchy of frames

The offered approach is self-organizing i.e. components of system set structure of data domain. The hierarchic combination of frames can be naturally extended by production rules for assigning of data domain dynamics.

Hierarchic frame knowledge representation like object-oriented approach gives an opportunity of frequentative usage of knowledge due to inheritance. Inheritance from frame hierarchies allows extending bases, and efficient knowledge clusterization around slots of frames allows reducing to minimum problems of the concordance of knowledge bases.

1 Background

The structure for the description of the theoretical unit consisting of characteristic of their values is called as frame. Characteristic is called as slots, and values - fillers of slots. Slot can contain not only concrete value, but also a procedure name, permitting to calculate it by the given algorithm, and also one or several rules with which help this value can be found.

Set of frames modeling any data domain, represents hierarchic structure in which frames are connected. There is a frame keeping the most general information, true for all remaining frames on the top level of hierarchy.

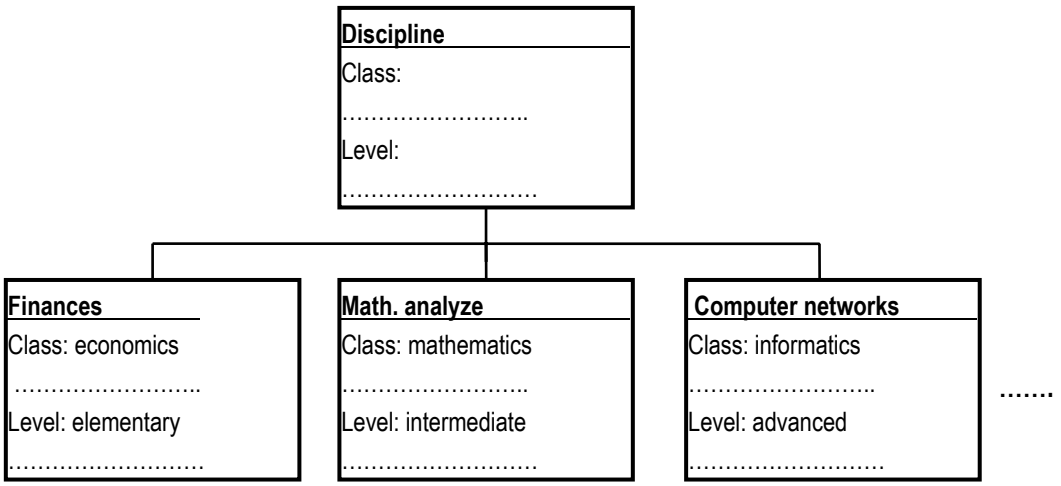


Fig.1 – Fragment of knowledge base describing data domain

Frames have ability to inherit characteristic values from the parents located on a higher hierarchy level. Characteristics values of frames can be transmitted on default to the frames located below of them in hierarchy, but if last contain eigenvalues of characteristic data they are considered as true data. This circumstance allows easily taking into account different elimination in frame systems [4].

DDM has static character, has few eliminations and connections between objects change infrequently. Described properties are realized in frame systems of knowledge representation in the best way. Values of slots are represented in the single copy in the system, as it is included only in one frame describing the most general concept from all those which contain the slot with data a name. Such property of frames systems enables to reduce a memory capacity, which is necessary for their placement in the computer. Besides economies of memory, presentation in KB the relationships existing between concepts of data domain concern to advantage of such model.

Under "Class" we understand scientific area, "Level" - a knowledge level of trainee, "Volume" - volume and the form of stated materials. Given attributes can be obligatory or not obligatory.

2 System of inference

Flexible and simple system of inference is necessary for representation DDM. Data domain model will interact with the trainee model by means of inference of production rules. Inferences of production system have search functions in KB, follow-up of operations above knowledge and obtaining of the conclusions. They describe knowledge in the form IF - THAT

There is a problem of mixing of rules with different properties in production system, at a solution of a problem of inference in several application areas. That fact essentially decrease efficacy of machining. However usage of a frame model in DDM in given model of learning process structures data domain that excludes mixing rules with different properties.

3 Production-frame knowledge representations

The architecture of production-frame hierarchy bases on frame knowledge representation in which frame hierarchy with relation of representation and active slots is taken for the base, and the inference is made by production rules. Such approach allows combining naturally in one model static knowledge about solving problems in the form of values of slots, structural knowledge of data domain by the way of inheritance hierarchies.

Thus, the frame system can be shown as:

$$W : S \to I$$

where I — set of frames, $S = \{S_i\}$, $i = \overline{1,n}$ — finite set of slots is presented as follows $\langle v, d, \{D_j\} \rangle$, including the current value of the slot $v = \langle v_1, v_2, ... v_l \rangle \in T$ and a default value $d = \langle d_1, d_2, ..., d_k \rangle \in T$, demons procedures {Dj}.

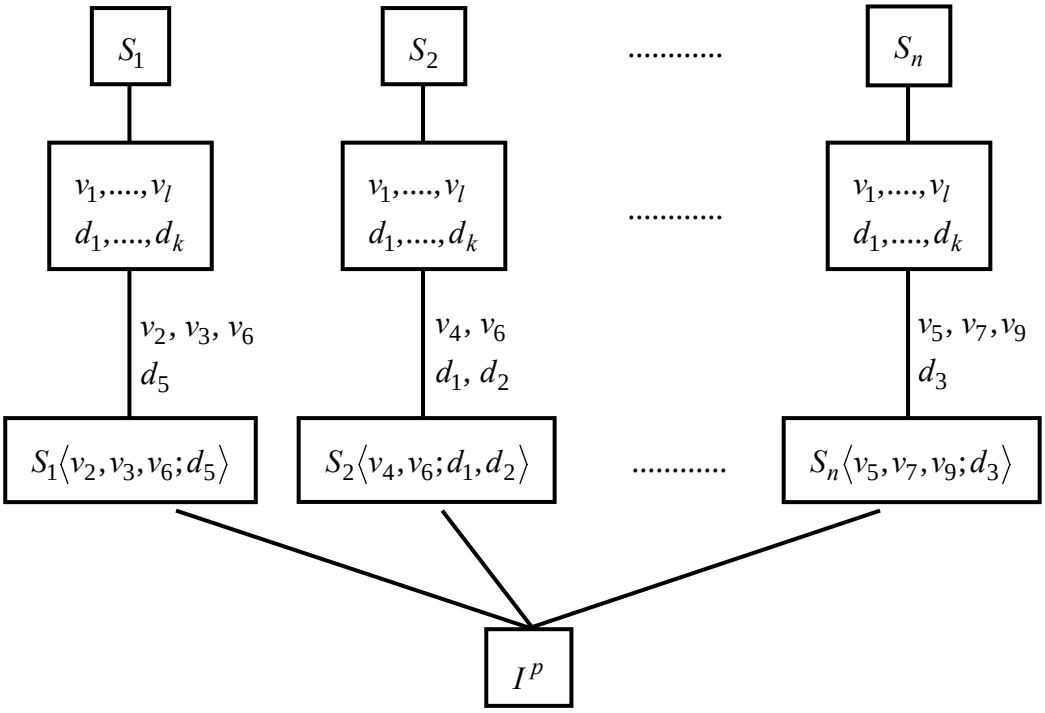


Fig. 2 – Example of inference of production system

Relation of inheritance is induced by the slot with the reserved name *parent*: $F : G \Leftrightarrow \|F(parent)\| = G$. Typical for frame systems the operation of the specification of the frame on a sample can be realized by implicit

inclusion in a model rule $F(parent) \leftarrow match(F,G)$. By consideration of multiple inheritance *parent slot* is supposed list type, and $F:G \Leftrightarrow G \in \|F(parent)\|$.

The Inference about values of output parameter (resultant frame I_P) implements under condition of accurate values of data-in (the frame of the query $I_з$), consisting of a subset $v \times d$) and is presented in such way:

$$L:\begin{cases} IF\ \left\langle v_1^1, ..., v_l^1; d_1^1, ...d_k^1 \right\rangle\ THEN\ S_1^p \\ IF\ \left\langle v_1^2, ..., v_l^2; d_1^2, ...d_k^2 \right\rangle\ THEN\ S_2^p \\ ..... \\ IF\ \left\langle v_1^n, ..., v_l^n; d_1^n, ...d_k^n \right\rangle\ THEN\ S_n^p \end{cases}$$

The resultant frame I_p represents combination v and d, which belong to different frames. Thus I_p - inference of production system can be presented as:

$$\left\langle S_1^p, S_2^p, ..., S_n^p \right\rangle \rightarrow I^p$$

Conclusions

In given paper the approach which represents creature of DDM of a learning system is developed. DDM is a knowledge base organized on the basis of production-frame knowledge representation language. Creature of frame hierarchy of knowledge base allows to present knowledge in the structured form with preservation of property of inheritance. That is actual at construction of DD of adjoining subjects. The production system ensures a flexible and understandable inference which allows generating a inference proceeding from personal trainee requirements. Representation DDM by a production-frame model allows reducing memory capacity necessary for filing.

Bibliography

1. Soshnikov D. Software Toolkit for Buiding Embedded and Distributed Knowledge-Based Systems. Proceedings of the 2nd International Workshop on Computer Science and Information Technologies, Vol.1, USATU Publishing, Ufa, 2000. pp. 103-111.
2. Петрушин В.А. Экспертно-обучающие системы. – Киев:Наукова думка, 1991. –196 с.
3. Попов Э.В. Общение с ЭВМ на естественном языке. М.: Наука, 1982. – 360 с.
4. Поспелов Д.А. Семиотические модели в управлении. / В кн. "Кибернетика. Дела практические". – М.: Наука, 1984. – с. 70-87.
5. Представление и использование знаний: Пер. с япон. / Под ред. Х. Уэно, М. Исидзука. – М.: Мир, 1989. – 220с.
6. Приобретение знаний: Пер. с япон. / Под ред. С. Осуги, Ю. Саэки. – М.: Мир, 1990. – 304 с.
7. Шеховцов Б.Г., Шкиль А.С., Пиженко И.Н., Шмаин Д.Ю. Концепция программно-информационной поддержки гипертекстового учебного материала для дистанционного обучения // АСУ и приборы автоматики. 2001. Вып.114. С.77-81.

Authors' Information

Alexander Ya. Kuzemin – Prof. of Information Department, Kharkov National University of Radio Electronics, Head of IMD, Ukraine, e-mail: kuzy@kture.kharkov.ua

Darya Fastova – phd student, Kharkov National University of Radio Electronics, Ukraine

Olesya Dyachenko – phd student, Kharkov National University of Radio Electronics, Ukraine

RESEARCH OF THE INFLUENCE OF METHODS OF COMPRESSION ON THE CO-EFFICIENT OF INFORMATION SECURITY OF OBJECTS UNDER ATTACKS

Dimitrina Polimirova, Eugene Nickolov, Cecko Nikolov

***Abstract**: In this paper a possibility for quantitative measuring of information security of objects, exposed to information attacks and processed with methods of compression, is represented. A co-efficient of information security, which reflects the influence of the level of compression obtained after applying methods of compression to objects and the time, required by the attack to get access to the corresponding object, is proposed. Methods' groups with the highest and respectively the lowest values of the co-efficient of information security for all methods' groups in relation to all attacks' groups are determined. Assessments and conclusions for future investigations are proposed.*

***Keywords**: Information Flows, Information Security, Information Attacks, Methods of Compression, Compressed Objects, Coefficient of Information Security, Level of Compression.*

***ACM Classification Keywords**: D.4.6 Security and Protection: information flow controls*

The Situation

Since the 70-th of XX century the problem for security and protection of information flows has drawn developers' and constructors' attention in the area if information technology [1]. With the first malattack in the 60th of last century [2], a progress in the area of object protection is observed and requirements for information security of objects are increased. Later the problem for creation of maximum protection for information flows arisen.

This cause the necessity to solve various problems associated with heightening the information flow security in the processes of transfer, processing and storage of different types of information flows. The researches on various methods for heightening the information security of different types of file objects became actual.

The Problem

The information flows, exposed to different attacks, are characterized with their large volume. Different methods for compression were developed and their use became necessity to reduce the volume of information flows. Compression can be used as a means of heightening the information security of information flows, especially when a password is applied during the compression.

We shell understand as a compression the transformation of the input data flow into codes. The decision for the correspondence "input data — output codes" is based on a preliminarily chosen model [3]. In case of effective compression, the obtained flow of codes is smaller in volume than the input data, but even though the compression is not effective, the file object will have a better protection against different attacks, because the output data will be presented by codes.

The purpose of research is to determine those compression methods whose application will give the object the highest security and protection against different attacks.

The Experiment

The experiments which were carried out were related to the necessity for a quantitative measure of information security. The following limitations have to be taken into consideration: 1) only the sets of potential attacks' groups (A_{pot}), methods' groups (M_{pot}) and objects' groups (O_{pot}), described respectively in Tables 1a, 1b, 1c will be analyzed (the groups' formation is described in [4]); 2) the experiments will be carried out according to standard user, not corporate or governmental, requirements.

Table 1a Attacks' groups

ATTACKS' GROUPS
I. Advertisements
V. Chat
VI. Criminal Investigations
VII. Cracking
VIII. Spying
X. Exploits
XIII. Scanners
XIV. Keyboard Modifiers
XVII. Computer Trojan Horses
XVIII. Computer Backdoors
XIX. Computer Worms
XX. Computer Viruses
XXI. Accessible information
XXIV. Content
XXV. Data Encapsulation
XXVII. Spoofing
XXXI. Social Engineering
XXXIII. Zombie Computers

Table 1b Methods' groups

METHODS' GROUPS
I. Statistical lossless methods
II. Dictionary lossless methods
III. Image lossless methods
IV. Audio lossless methods
V. Other lossless methods
VI. Dictionary lossy methods
VII. Image lossy methods
X. Audio lossy methods

Table 1c Objects' groups

OBJECTS' GROUPS
II. Scientific file formats
III. Data file formats
IV. Graphics
V. Sound
VII. Internet related
VIII. Binaries
X. Miscellaneous

A research will be conducted according to the influence of the level of compression on the time required by the attack, representative of an attacks' group from the potential set of attacks' groups, to get access ρ to the object, representative of an objects' group from the potential set of objects' groups, where ρ will present the different type of an access (read, write, execute and delete – to simplify the investigations will not make difference between its functionality). The following tasks were posed:

1) To determine the co-efficient of information security of the object.

The co-efficient of information security as a quantitative value, which depends on several fundamental parameters, can be determined. The co-efficient of information security can be presented as a value, dependent on several main parameters. These parameters can be represented as ratios of separate values before and after certain impact. The parameter *TIME* can be represented as a ratio of *time-prim* for the attack of the object before impact of the method of compression to the *time-second* for the attack after impact of the method. The parameter *SIZE* can be represented as a ratio of *size-prim* of the object before impact of the methods of compression to the *size-second* of the object after the compression. In this way the rest parameters as password, encryption etc. can be dynamically described in the necessary depth and concreteness.

1.1) The first stage of this task is to determine the *LEVEL OF COMPRESSION* of the objects processed with different methods of compression.

Three typical representatives for each objects' group, from the set of potential objects' groups (totally 21 representatives) and three typical representative for each methods' group, from the set of potential methods' groups (totally 24 representatives) are chosen. Each object is processed with the different methods and results are obtained for the corresponding level of compression (*L*) (totally 504 levels of compression, 9 for each relation object's group—method's group). The level of compression (*L*) can be defined as [5]:

$$L = \left(1 - \frac{\text{uncompressed size}}{\text{compressed size}}\right).100 = [\%]$$

To improve the visualization in Table 2 are shown the average values of the level of compression for each relation method's group—object's group. Maximum and minimum values are marked.

Table 2 Average values of the level of compression (*L*).

		METHODS' GROUPS									
		I	***II***	***III***	***IV***	***V***	***VI***	***VII***	***VIII***	***IX***	***X***
OBJECTS' GROUPS	***I***	-	-	-	-	-	-	-	-	-	-
	II	78%	32%	-	-	62%	-	-	-	-	-
	III	64%	92%	-	-	78%	-	-	-	-	-
	IV	-	-	81%	-	-	**96%**	94%	-	-	-
	V	-	-	-	73%	-	-	-	-	-	86%
	VI	-	-	-	-	-	-	-	-	-	-
	VII	56%	53%	-	-	48%	-	-	-	-	-
	VIII	**5%**	9%	-	-	7%	-	-	-	-	-
	IX	-	-	-	-	-	-	-	-	-	-
	X	6%	8%	-	-	-	-	-	-	-	-

1.2) The second stage is connected to the experiments which have to be carried out measuring the *TIME* required by an attack $a_i \in A_{pot}$ to get access ρ to the object $o_f \in O_{pot}$ in the case when the object is processed with method of compression $m_j \in M_{pot}$, and in the case when the object is not processed with method of compression.

The equipment used for the experiments includes: two server configurations with two workstations each. The first server configuration (in conjunction with two workstations) is used for research of "attacking behavior", another server configuration (in conjunction with another two workstations) is used for research of "protective behavior".

The workstations in the respective server configurations assume the role of "managing station" and "model station". The research techniques are based on the utilization of program systems for scientific research of the firm "The MathWorks", including the *family Matlab 7.2* and the *family Simulink*.

Three typical representatives are chosen for: 1) each of 18-th attacks' groups; 2) each of 8-th methods' groups; 3) each of 7-th objects' groups from the potential sets of groups, shown in Tables 1a, 1b, 1c.

An estimation is done with respect to the time required by the attack (totally 54 attacks) to get access to each object (totally 21) in these cases: 1) objects are not processed with methods of compression; 2) objects are processed with methods of compression (as the time for applying of each one from 24th methods to the corresponding object is investigated).

Results are obtained for totally 27216 triple relations attack—methods—object and the obtained values are compared with the values of relations attack—object. Thus the time growth for realizing an attack to an object, after its processing with the corresponding method, can be determined (in percentage). To improve the visualization in Table 3 are shown only the average values for obtained time growth in percentages, required by the attacks to be realized to the objects in case when the objects are processed with methods of compression. Maximum and minimum values are marked.

Table 3 Average values for obtained time growth in percentages, required by the attacks to get access to the objects processed with methods of compression.

		METHODS' GROUPS									
		I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*
OBJECTS' GROUPS	*I*	-	-	-	-	-	-	-	-	-	-
	II	27,90%	**24,53%**	-	-	60,33%	-	-	-	-	-
	III	41,23%	38,29%	-	-	38,13%	-	-	-	-	-
	IV	-	-	51,86%	-	-	43,64%	44,28%	-	-	-
	V	-	-	-	38,29%	-	-	-	-	-	36,05%
	VI	-	-	-	-	-	-	-	-	-	-
	VII	46,76%	44,36%	-	-	**85,88%**	-	-	-	-	-
	VIII	34,17%	31,10%	-	-	30,93%	-	-	-	-	-
	IX	-	-	-	-	-	-	-	-	-	-
	X	43,29%	39,65%	-	-	-	-	-	-	-	-

1.3) The third stage is to determine the *CO-EFFICIENT OF INFORMATION SE*CURITY (K_{INF}).

The coefficient of information security can be determined as a combination of the coefficients of time (K_T) and size (K_S). K_T is the ratio of the time, required by the attack to get access to the object before impact of the method, to the time, required by the attack to get access to the object after impact of the method.

$$TIME = \frac{\text{time required by the attack before impact of the method } (t')}{\text{time required by the attack after impact of the method } (t'')}$$

K_S is the ratio of the uncompressed object's size to the object's size processed with method.

$$K_S = \frac{\text{uncompressed object's size } (s')}{\text{compressed object's size } (s'')}$$

Then K_{INF} can be denoted as:

$$K_{INF} = K_T + K_S = \frac{t'}{t''} + \frac{s'}{s''}$$

After the investigations, described in 1.1) and 1.2), which were carried out, the co-efficient of information security can be determined for the representatives of corresponding objects. To improve the visualization in Table 4 are shown only the average values for obtained coefficients of information security from applying methods to objects, exposed to attacks. Maximum and minimum values are marked.

Table 4 Average values for obtained coefficients of information security from applying methods to objects, exposed to attacks.

		METHODS' GROUPS									
		I	*II*	*III*	*IV*	*V*	*VI*	*VII*	*VIII*	*IX*	*X*
OBJECTS' GROUPS	*I*	-	-	-	-	-	-	-	-	-	-
	II	5,27	2,23	-	-	3,03	-	-	-	-	-
	III	3,37	13,12	-	-	5,16	-	-	-	-	-
	IV	-	-	5,74	-	-	**25,56**	17,22	-	-	-
	V	-	-	-	4,32	-	-	-	-	-	7,78
	VI	-	-	-	-	-	-	-	-	-	-
	VII	2,81	2,68	-	-	2,06	-	-	-	-	-
	VIII	1,71	1,79	-	-	1,77	-	-	-	-	-
	IX	-	-	-	-	-	-	-	-	-	-
	X	**1,63**	1,69	-	-	-	-	-	-	-	-

2) To determine the methods with the highest coefficients of information security

A database can be build after experiments which were carried out during the first posed task. The base includes matrices with the values of obtained coefficients of information security when methods of compression were applied to objects. The number of matrices is 18 (one for each group of attacks). Each matrix includes the average values of the obtained K_{INF} from the applying 8 methods' groups to 7 objects' groups. Thus for each matrix the methods' group with the highest co-efficient of information security for every objects' group for the respective attacks' group can be determined. In Table 5 are shown the average values for the best methods' group, after applying the highest values of the co-efficient of information security for all objects' groups, exposed to the corresponding attacks' group, are obtained.

Table 5 Methods' groups after applying the highest value of the co-efficient of information security for all objects' groups, exposed to the corresponding attacks' group, are obtained

Attacks' groups	Group of methods with the highest values of K_{INF}
I. Advertisements	Dictionary lossless methods
V. Chat	Dictionary lossy methods
VI. Criminal Investigations	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
VII. Cracking	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
VIII. Spying	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
X. Exploits	Statistical, dictionary lossless, methods
XIII. Scanners	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XIV. Keyboard Modifiers	Image lossless
XVII. Computer Trojan Horses	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XVIII. Computer Backdoors	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XIX. Computer Worms	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XX. Computer Viruses	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XXI. Accessible information	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XXIV. Content	Statistical, dictionary lossless, dictionary lossy methods
XXV. Data Encapsulation	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XXVII. Spoofing	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XXXI. Social Engineering	Statistical, dictionary lossless, dictionary lossy, audio lossy methods
XXXIII. Zombie Computers	Statistical, dictionary lossless, dictionary lossy, audio lossy methods

From the experiments which were carried out and after the obtained results methods with the highest co-efficients of information security in relation to all attacks for every object can be determined. Results are shown in Table 6.

Table 6 Methods' groups with the highest coefficients of information security in relation to all attacks' groups for every objects' group

Objects' group	Groups of methods with the highest values of K_{INF} in relation to all attacks' groups
II. Scientific file formats	Statistical lossless methods
III. Data file formats	Dictionary lossless methods
IV. Graphic file formats	Dictionary lossy methods
V. Sound formats	Audio lossy methods
VII. Internet related	Statistical lossless methods
VIII. Binaries	Dictionary lossless methods
X. Other	Dictionary lossless methods

The following assessments could be made from the experiments which were carried out:

1) With respect to the LEVEL OF COMPRESSION:

a) With respect to the selected objects which will be processed by methods of compression, the assessment is positive, and the assumptions don't impact the obtained results. With respect to the chosen methods of compression, the assessment is also positive, and the conducted experiments can be applied successfully enough for the other methods as well.

b) The best results are shown with objects from the graphics group (*IV*), processed with a method of compression belonging to the group of dictionary lossy methods (*VI*). The worst results are shown with objects from the group of other file formats (*X*), processed with a method of compression belonging to the group of statistical methods (*I*).

2) With respect to the TIME, required by an attack to get access to an object:

a) The selected server configurations, defined as "high-level professional server", and the workstations, defined as "middle-level professional station", are sufficient for making the necessary conclusions and recommendations. The conducted experiments on these configurations are also valid for other standard user not corporate (government) configurations as well.

b) The best results are shown with objects from the group of “Scientific file formats”, processed with a method of compression belonging to the group of Statistical methods with respect to the attack belonging to the group of Spying. The worst results are shown with objects from the groups of Data file formats and graphics file formats, processed with a method of compression belonging to the group of “Other” lossless methods with respect to the attack belonging to the groups of Criminal Investigations and Computer Viruses.

c) Based on the conducted experiments, we can make the conclusion that increasing the level of compression leads to increasing the time required by an attack to get an access to an object.

3) With respect to the CO-EFFICIENT OF INFORMATION SECURITY:

a) The best results are shown with objects from the group of graphics file formats (*IV*), processed with a method of compression belonging to the group of dictionary lossy methods (*VI*) with respect to the attack belonging to the group of Zombie Computers (*XXXIII*). The worst results are shown with objects from the group of other file formats (*X*), processed with a method of compression belonging to the group of dictionary lossless methods (*I*) with respect to the attack belonging to the group of Criminal Investigations (*VI*).

b) It is necessary to conduct research by using a password in determining the coefficient of information security.

Conclusion

The influence of compression methods on different objects is substantial and can be utilized for making decisions with regards to the information security of these objects in relation to different attacks, and for evaluation of methods with the lowest risk with respect to the coefficient of information security.

Bibliography

[1] Dorothy E. Denning, A lattice model of secure information flow, Communications of the ACM, v. 19 n. 5, p. 236-243, May 1976.

[2] http://www.sptimes.com/Hackers/history.hacking.html

[3] Тимоти Стенли, Компресиране на данни, изд. “Интерфейс България”, София, 1998

[4] Dimitrina Polimirova, Eugene Nickolov, Cecko Nikolov, Investigating The Relations Of Attacks, Methods And Objects In Regard To Information Security In Network TCP/IP Environment, Proceedings of International Conference Information Theories and Applications: Cyber Security, 28-29 June 2006, Varna, to be published.

[5] David Salomon, Data Compression: The Complete Reference, Springer Verlag New York, Inc., 2004.

Authors' Information

Dimitrina Polimirova – PhD Student, Research Associate, National Laboratory of Computer Virology, Bulgarian Academy of Sciences, Phone: +359-2-9733398, e-mail: polimira@nlcv.bas.bg

Eugene Nickolov – Professor, DSc, PhD, Eng, National Laboratory of Computer Virology, Bulgarian Academy of Sciences, Phone: +359-2-9733398, e-mail: eugene@nlcv.bas.bg

Cecko Nikolov – PhD Student, Research Associate, National Laboratory of Computer Virology, Bulgarian Academy of Sciences, Phone: +359-2-9733398, e-mail: nikolov@nlcv.bas.bg

CODE ACCESS SECURITY IN MICROSOFT'S .NET

Petar Atanasov

***Abstract**: This paper introduces basic concepts of code access security, using and implementing security features, as well as types of security syntax and mechanism of checking and requesting specific permissions.*

Preface

Role-Based security is at the heart of Microsoft Windows 2000/XP operating systems, but it isn't enough to depend on the code itself and to neglect the skills and awareness of the user.

This security model cares about user access secure resources and any code usually runs under the credentials of the logged on user.

Common scenario for Windows users:

1. Installed ActiveX runs under user's security permission set or it can do pretty much anything with the system that the user can. (Delete, update files, etc.)
2. The user has no idea of what the ActiveX does and most likely it doesn't cross his/her mind, but what matters to the user is that the computer remains secure.

The recent virus "*Sasser*" does not require user interaction for the computer to get infected; simply plug an unprotected machine to the net and in a matter of minutes it becomes infected.

Code Access Security reveals where Role-Based security needs assistance for adequate performance. Code Access Security builds upon Role-Based security.

It provides it with the mechanism of securing the code based on who wrote it and where it came from, or where it is executed (evidences).

These evidences are mapped to the permissions (rights), which can be administered by four different policies, which correspond to the role user represents:

1. Domain Administrator – Enterprise Policy
2. Machine Administrator – Machine Policy
3. Actual User of the machine - User Policy
4. Developer - Application domain Policy

These policies are configurable after the application is deployed and can be modified at any point in time.

One major concept was introduced with CAS - Partially trusted code is code that has been granted only access to the resources it needs to execute successfully and no more.

Code Access Security Basics

Every application that targets the common language runtime must interact with the runtime's security system. When an application executes, it is automatically evaluated and given a set of permissions by the runtime. Depending on the permissions that the application receives, it either runs properly or generates a security exception. The local security settings on a particular computer ultimately decide which permissions code receives. Because these settings can change from computer to computer, a developer can never be sure that his code will receive sufficient permissions to run. This is in contrast to the world of unmanaged development, in which it's not necessary to worry about code's permission to run.

Every developer must be familiar with the following code access security concepts in order to write effective applications targeting the common language runtime:

- Writing type-safe code: To enable code to benefit from code access security, must be used a compiler that generates verifiably type-safe code.
- Imperative and declarative syntax: Interaction with the runtime security system is performed using imperative and declarative security calls. Declarative calls are performed using attributes; imperative calls are performed

using new instances of classes within code. Some calls can be performed only imperatively, while others can be performed only declaratively. Some calls can be performed in either manner.

- Requesting permissions for the code: Requests are applied to the assembly scope, where code informs the runtime about permissions that it either needs to run or specifically does not want. Security requests are evaluated by the runtime when code is loaded into memory. Requests cannot influence the runtime to give code more permissions than the runtime would have given to the code and had the request not been made. However, requests are what code uses to inform the runtime about the permissions it requires in order to run.
- Using secure class libraries: Class libraries use code access security to specify the permissions they require in order to be accessed. The developer should be aware of the permissions required to access any library that code uses and make appropriate requests in itself.

Type-safe code

Type-safe code is code that accesses types only in well-defined, allowable ways. For example, given a valid object reference, type-safe code can access memory at fixed offsets corresponding to actual field members. However, if the code accesses memory at arbitrary offsets outside the range of memory that belongs to that object's publicly exposed fields, it is not type-safe.

JIT compilation performs a process called verification that examines code and attempts to determine whether the code is type-safe. Code that is proven during verification to be type-safe is called verifiably type-safe code. Code can be type-safe, yet not be verifiably type-safe, due to the limitations of the verification process or of the compiler. Not all languages are type-safe, and some language compilers, such as Microsoft Visual C++, cannot generate verifiably type-safe managed code. To determine whether the language compiler is used generates verifiably type-safe code, should be consulted the compiler's documentation. If is used a language compiler that generates verifiably type-safe code only when developer avoids certain language constructs, it might be useful to be used the .NET Framework SDK PEVerify tool to determine whether code is verifiably type-safe.

Code that is not verifiably type-safe can attempt to execute if security policy allows the code to bypass verification. However, because type safety is an essential part of the runtime's mechanism for isolating assemblies, security cannot be reliably enforced if code violates the rules of type safety. By default, code that is not type-safe is allowed to run only if it originates from the local computer. Therefore, mobile code should be type-safe.

Security Syntax

Code that targets the common language runtime can interact with the security system by requesting permissions, demanding that callers have specified permissions, and overriding certain security settings (given enough privileges). There are two different forms of syntax to programmatically interact with the .NET Framework security system: declarative syntax and imperative syntax. Some operations can be done using both forms of syntax while other operations can be performed using only declarative syntax. A developer should be familiar with both forms

Declarative Security

Declarative security syntax uses attributes to place security information into the metadata of code. Attributes can be placed at the assembly, class, or member level, to indicate the type of request, demand, or override developer want to use. Requests are used in applications that target the common language runtime to inform the runtime security system about the permissions that the application needs or does not want. Demands and overrides are used in libraries to help protect resources from callers or to override default security behavior.

In order to use declarative security calls, developer must initialize the state data of the permission object so that it represents the particular form of permission he needs. Every built-in permission has an attribute that is passed a SecurityAction enumeration to describe the type of security operation wanted to perform. However, permissions also accept their own parameters that are exclusive to them.

The following code fragment shows declarative syntax for requesting that code's callers have a custom permission called `MyPermission`. This permission is a hypothetical custom permission and does not exist in the .NET Framework. In this example, the declarative call is placed directly before the class definition, specifying that this permission be applied to the class level. The attribute is passed a **SecurityAction.Demand** structure to specify that callers must have this permission in order to run.

Visual Basic

```
<MyPermission(SecurityAction.Demand, Unrestricted = True)> Public Class MyClass1

   Public Sub New()
      'The constructor is protected by the security call.
   End Sub
   Public Sub MyMethod()
      'This method is protected by the security call.
   End Sub
   Public Sub YourMethod()
      'This method is protected by the security call.
   End Sub
End Class
```

C#

```
[MyPermission(SecurityAction.Demand, Unrestricted = true)]
public class MyClass
{
   public MyClass()
   {
      //The constructor is protected by the security call.
   }
   public void MyMethod()
   {
      //This method is protected by the security call.
   }
   public void YourMethod()
   {
      //This method is protected by the security call.
   }
}
```

Imperative Security

Imperative security syntax issues a security call by creating a new instance of the permission object wanted to invoke. Developer can use imperative syntax to perform demands and overrides, but not requests.

Before making the security call, developer must initialize the state data of the permission object so that it represents the particular form of the permission he need. For example, when creating a FileIOPermission object, can be used the constructor to initialize the **FileIOPermission** object so that it represents either unrestricted access to all files or no access to files. Or, developer can use a different **FileIOPermission** object, passing parameters that indicate the type of access he wants the object to represent (that is, read, append, or write) and what files he wants the object to protect.

In addition to using imperative security syntax to invoke a single security object, developer can use it to initialize a group of permissions called a permission set. For example, this technique is the only way to reliably perform assert calls on multiple permissions in one method. Use the PermissionSet and NamedPermissionSet classes to create a group of permissions and then call the appropriate method to invoke the desired security call.

Developer can use imperative syntax to perform demands and overrides, but not requests. It might be useful to use imperative syntax for demands and overrides instead of declarative syntax when information that is needed in order to initialize the permission state becomes known only at run time. For example, if developer wants to ensure that callers have permission to read a certain file, but he does not know the name of that file until run time, use an imperative demand. Developer might also choose to use imperative checks instead of declarative checks when he needs to determine at run time whether a condition holds and, based on the result of the test, make a security demand (or not).

The following code fragment shows imperative syntax for requesting that code's callers have a custom permission called `MyPermission`. This permission is a hypothetical custom permission and does not exist in the .NET Framework. A new instance of `MyPermision` is created in `MyMethod`, guarding only this method with the security call.

Visual Basic

```
Public Class MyClass1

   Public Sub New()
   End Sub
   Public Sub MyMethod()
      'MyPermission is demanded using imperative syntax.
      Dim Perm As New MyPermission()
      Perm.Demand()
      'This method is protected by the security call.
   End Sub
   Public Sub YourMethod()
      'YourMethod 'This method is not protected by the security call.
   End Sub
End Class
```

C#

```
public class MyClass {
   public MyClass(){
   }
   public void MyMethod() {
       //MyPermission is demanded using imperative syntax.
       MyPermission Perm = new MyPermission();
       Perm.Demand();
       //This method is protected by the security call.
   }
   public void YourMethod() {
       //This method is not protected by the security call.
   }
}
```

Manifest stores Metadata information that can be read without running the assembly; therefore if Developer was using Declarative security to enforce security than Administrator can simply run command-line utility (Permview.exe) to view what Permission is needed to have to run produced code.

In comparison, Imperative is more flexible and is stored as MSIL code, which will be compiled in JIT and given a Security Exception at run-time.

Requesting Permissions

Requesting permissions is the way the developers let the runtime know what code needs to be allowed to do. He request permissions for an assembly by placing attributes (declarative syntax) in the assembly scope of code. When the assembly is created, the language compiler stores the requested permissions in the assembly manifest. At load time, the runtime examines the permission requests, and applies security policy rules to determine which permissions to grant to the assembly. Requests only influence the runtime to deny permissions to code and never influence the runtime to give more permissions to the code. The local administration policy always has final control over the maximum permissions code is granted.

Although code does not have to request permissions in order to compile, there are important reasons the code should always request permissions:

- Requesting permissions increases the likelihood that code will run properly if it is allowed to execute. Code that request a minimal set of permissions will not run unless it receives those permissions. If developer does not identify a minimum set of permissions, code must gracefully handle any and all situations where not being granted some permission might prevent it from executing properly.
- Requesting permissions helps ensure that code is granted only the permissions it needs. If code is not granted extra permissions, it cannot damage the resources protected by those extra permissions, even if it is exploited by malicious code or has bugs that can be leveraged to damage resources. Developer should request only those permissions that his code needs, and no more.
- Requesting permissions lets administrators know the minimum permissions that the application needs so that they can adjust security policy accordingly. Administrators use the Permission View Tool (Permview.exe) to examine assemblies and set up security policy to issue required permissions. If developer does not explicitly

request the permissions that application requires, the Permission View tool cannot return any information about the permissions that produced application requires. If an administrator does not know this information, the application is difficult to administer.

Requesting permissions informs the runtime which permissions the application needs to function or specifically does not want. For example, if the application writes to the local hard disk without using isolated storage, application must have FileIOPermission. If the developer does not request **FileIOPermission** and the local security settings do not allow the application to have this permission, a security exception is raised when the application attempts to write to the disk. Even if the application can handle the exception, it will not be allowed to write to the disk. This behavior might be frustrating to users if the application is a text-editing program that they have been using for an extended period of time. On the other hand, if the application requests **FileIOPermission** and the local security settings do not allow the application to have **FileIOPermission**, the application will generate the exception when it starts and the user will not face the problem of losing any work. Additionally, if the application requests **FileIOPermission** and if it is a trusted application, the administrator can adjust security policy to allow it to execute from the remote share.

If the code does not access protected resources or perform protected operations, developer does not need to request any permissions. For example, a permission request might not be necessary if the code simply computes a result based on inputs passed to it, without using any resources. If code does access protected resources but does not request the necessary permissions, it might still be allowed to execute, but it could fail at some point during execution if it attempts to access a resource for which it does not have the necessary permission.

To request permissions, developer must know which resources and protected operations code uses, and he must also know which permissions protect those resources and operations. In addition, he needs to keep track of any resources accessed by any class library methods that are called by the solution components.

The following table describes the types of permission requests.

Permission request	**Description**
Minimum permissions (RequestMinimum)	Permissions code must have in order to run.
Optional permissions (RequestOptional)	Permissions code can use but can run effectively without. This request implicitly refuses all other permissions not specifically requested.
Refused permissions (RequestRefuse)	Permissions that is wanted to ensure will never be granted to the code, even if security policy allows them to be granted.
Perform any of the above requests on built-in permission sets (Requesting Built-in Permission Sets).	Built-in permission sets, including **Nothing**, **Execution**, **FullTrust**, **Internet**, **LocalIntranet**, and **SkipVerification**.
Perform any of the above requests on XML-encoded permission sets (Requesting XML-Encoded Permissions).	XML representation (either a string containing the XML-encoded permission set or the location of an XML file containing the encoded permission set) of a desired permission set.

If developer specifies required permissions (using **RequestMinimum**), the code will be granted each required permission that security policy allows. The code will be allowed to run only if it is granted all the permissions it requires.

Requesting optional permissions without also requesting required permissions can, in some cases, severely restrict the permissions granted to an assembly. For example, suppose security policy normally grants Assembly A the permissions associated with the **Everything** named permission set. If the developer of Assembly A requests Permission A as optional and does not request any required permissions, Assembly A will be granted either Permission A (if security policy allows it) or no permissions at all.

Using Secure Class Libraries

Secure library is a class library that uses security demands to ensure that the library's callers have permission to access the resources that the library exposes. For example, a secure class library might have a method for creating files that would demand that its callers have permissions to create files. The .NET Framework comprises secure class libraries.

If code requests and is granted the permissions required by the class library, it will be allowed to access the library and the resource will be protected from unauthorized access; if code does not have the appropriate permissions, it will not be allowed to access the class library, and malicious code will not be able to use developer's code to indirectly access the resource. Even if code has permission to access a library, it will not be allowed to run if code that calls the code does not also have permission to access the library.

Code access security does not eliminate the possibility of human error in writing code; however, if applications use secure class libraries to access protected resources, the security risk for application code is decreased because class libraries are closely scrutinized for potential security problems.

Java – Security Manager

In JDK 1.1, local applications and correctly digitally signed applets were generally trusted to have full access to vital system resources, such as the file system, while unsigned applets were not trusted and could access only limited resources.

A security manager was responsible for determining which resource accesses were allowed.

The Java 2 SDK security architecture is policy-based, and allows for fine-grained access control. When code is loaded, it is assigned "permissions" based on the security policy currently in effect.

Each permission specifies a permitted access to a particular resource, such as "read" and "write" access to a specified file or directory, or "connect" access to a given host and port.

The policy, specifying which permissions are available for code from various signers/locations, can be initialized from an external configurable policy file.

Unless permission is explicitly granted to code, it cannot access the resource that is guarded by that permission.

These new concepts of permission and policy enable the SDK to offer fine-grain, highly configurable, flexible, and extensible access control.

Such access control can now not only be specified for applets, but also for all Java code, including applications, beans, and servlets.

Conclusion

Even with these powerful features code access security doesn't implement the process of creating and envisioning the whole solution, nether in the web nor in the software solutions. The weak implementation of security in the sense of too much references (and invocations) to objects in the security stack could result in performance penalty, while not securing the application could lead to unpredictable results. Code access security gives only a well defined, structured and clean way of performing security, while the security itself is still a priority task number one for the software/web developer.

Bibliography

MSDN

Code Access Security (CAS) and Design Patterns

Security Managers and the Java™ 2 SDK

Author's Information

Petar P. Atanasov – Software developer at Unicontsoft (http://www.unicontsoft.com); phone: ++359-88-7208893; e-mail: ppa@unicontsoft.com

VIRTUAL ENCYCLOPAEDIA OF THE BULGARIAN ICONOGRAPHY

Lilia Pavlova-Draganova, Vladimir Georgiev, Lubomil Draganov

***Abstract**: East-Christian icon art is recognised as one of the most significant areas of the art of painting. Regrettably, it is still being neglected in the digital documentation and the registry of the art of painting. The accessibility to that large part of mankind's cultural and historical ancestry would be enhanced greatly if icons of all possible kinds and origins were digitised, classified, and „exhibited" in the Internet. That would allow the preservation and even the future digital restoration of a large number of rare specimens of the East-Christian art of painting. This article aims to introduce how modern techniques from the area of digital libraries can be used for implementing the demonstrative multimedia library "Virtual encyclopaedia of the Bulgarian iconography " [4], containing a large number of Bulgarian iconic art masterpieces and iconography of various authors, periods and schools.*

***Keywords**: Digital libraries, Multimedia content, Cultural heritage.*

***ACM Classification Keywords:** H.3.7 Digital Libraries – Collection, Dissemination, System issues.*

Basic concepts and characteristics of digital libraries with multimedia content

Digital libraries (DL) are a contemporary conceptual solution for access to information archives. The "i2010: Digital Libraries" initiative aims at making European information resources easier and more interesting to use in an on-line environment. The Commission adopted on 30/09/2005 the "i2010: Digital Libraries" communication outlining the vision of this initiative and addressing in particular the issues of digitisation, on-line accessibility and digital preservation of our cultural heritage. [i2010: Digital Libraries, 2005].

According to an informal definition of digital libraries, they are managed collections of information, with associated services, where the information is stored in digital formats and accessible over a network. Digital libraries contain diverse hypertext-organized collections of information (digital objects such as text, images, and media objects) for use by many different users. The collected information is organized thematically and uses hyperlinks that allow the connection between any piece of data and additional data on the same topic. As an addition to the digital objects collection, there are many levels of metadata, indexes, hierarchical links, etc. [Krastev' '05]

The main characteristics of digital libraries are the following:

- Ability to share information;
- New forms and formats for information presentation;
- Easy information update;
- Accessibility from anywhere, at any time;
- Services available for searching, selecting, grouping and presenting digital information, extracted from a number of locations. Using these services depends on the user preferences, needs and wishes of the users, i.e. there is personalization available;
- Contemporary methods and tools for digital information protection and preservation;
- Ability to use different types of computer equipment and software;
- No limitations related to the size of content to be presented.

[4] The digital library "Virtual encyclopaedia of the Bulgarian iconography" is part of the project "Digital libraries with multimedia content and its application in Bulgarian cultural heritage" by contract 8/21.07.2005 between IMI-BAS and State agency for information technologies and communications.

In the past digital libraries were isolated and monolithic systems limited to access to content of a single provider. The development of the technologies during the last years provides new functionalities and advanced services to contemporary digital libraries such as specialized services for:

- Multi-layer and personalized search, context-based search, relevance feedback, etc.
- Resource and collection management;
- Metadata management;
- Indexing;
- Semantic annotation of digital resources and collection etc.

Three types of digital library architectures are described in [Pavlov at al. '05] - hypermedia digital library, Grid-based infrastructure and hyperdatabase infrastructure. They have different complexity. Considering the specific needs and requirements of different cultural and historical heritage projects some of them could be chosen and implemented.

The new digital libraries will provide and manage complex services, processes and workflows on the basis of existing services. It is expected that these services be heterogeneous, autonomous and distributed. The flexibility, the automatic adaptation, the access anywhere and anytime, the decentralization, the wide variety of digital objects and collections, the information security, etc. will be of the some requirements. [Kiernan at al. '03] [Pavlov, Paneva '05]

"Virtual encyclopaedia of the Bulgarian iconography" - a demonstrator of digital library with multimedia content

Information and multimedia technologies have the potential to make national cultural and historical heritage visible and available for present and future use. The goal of the project "Virtual encyclopaedia of the Bulgarian iconography " is to develop the information content, structure, and the realisation of a digital library with multimedia content as a demonstrator of "Virtual encyclopaedia of the Bulgarian iconography". That library includes several hundred specimens of Bulgarian icons from different artists, historical periods, and schools. The chosen architecture represents a hypermedia digital library [Pavlov at al. 05], which means that presentation of a complex multimedia content in the Internet is simplified. The resources are digital objects of different formats – text, graphics, and other media. They are structured in a hypermedia manner, i.e., some digital objects point to other ones. In this way the user can navigate quickly, in a non-linear fashion, within areas of related topics, using the hyperlinks. The digital objects are grouped according to their topics into thematic collections [Pavlov at al. 05]. For each object and collection, special meta-descriptions are created. They include data about the title, the artist, the period (in years and centuries), the school, the dimensions (width/height/thickness), the technique, the base material (type of wood, ground coat, etc) the category, the location, the title description, the author description (biographic data), the comment (icon features such as state, founder's and other signatures, previous restorations), etc. Also, they contain links to other digital objects and collections, keywords, and so on. That information is used for the semantic annotation and indexing of the digital objects, which facilitate their locating during search requests, and their web-based representation [Paneva at al. '05]. A multitude of specialized services for metadata management, content management, indexing, metadata annotation of digital objects and collections, creation of requested document, context-based search, multi-object, multi-feature search, etc, are presented. The "Search" service in the demonstrator aids the visitors in finding a certain object by the following criteria: icon title, author, period, type, school, region and location. The search can be conducted by one as well as by more criteria. Figure 1 depicts the architecture of hypermedia digital library.

The organisation of the media databases, the representation and description of the digital objects, and the classification of the artefacts, are developed according to the recommendations of the international group of museum experts of East-Christian Art (UNESCO/I.DB.I) and the standards of CIDOC/MICMO. A lot of work is done on that project for the creation of a high quality interface to a developed multimedia database of several hundred pieces of art of different authors, periods, and schools of the Bulgarian Christian iconography. That group includes painted icons and icons built with mosaics that are located in European museums, churches, monasteries, and private collections.

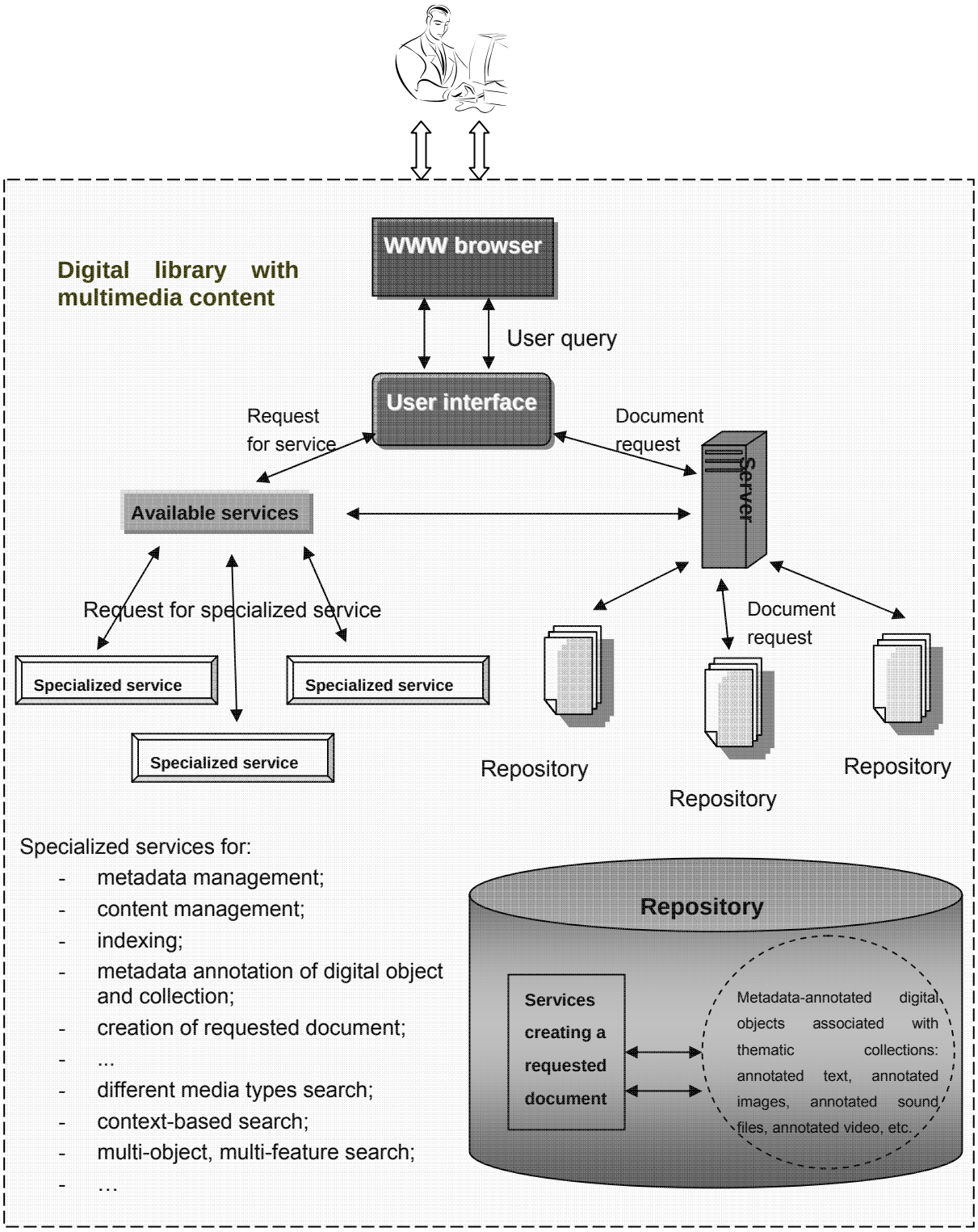


Figure 1: Hypermedia digital library

The mechanisms for protecting the displayed images and textual data from being copied, reproduced and later distributed are conformed to the collections and artifacts owners' requirements – i.e. the presented digital counterparts of Bulgarian icons are digitally protected with the credit line method, as well as by disallowing the copying and downloading of the images.

The project relies on the idea that the unity of the text information and the high quality of the digital images will represented the virtues of the Bulgarian icon in their entirety and will contribute to the its preservation, wider exhibition, and future potential restoration. The demonstrator that is developed is a tool for exploration of the techniques, styles, colours, and forms, as well as for the tracking and comparing of specimens and periods of the iconography and historical development of that art. [Pavlov at al. '05]

System realization of the demonstrator

The objects of the system, their attributes and the relationships between them are shown in the entity-relationship diagram on fig. 2.

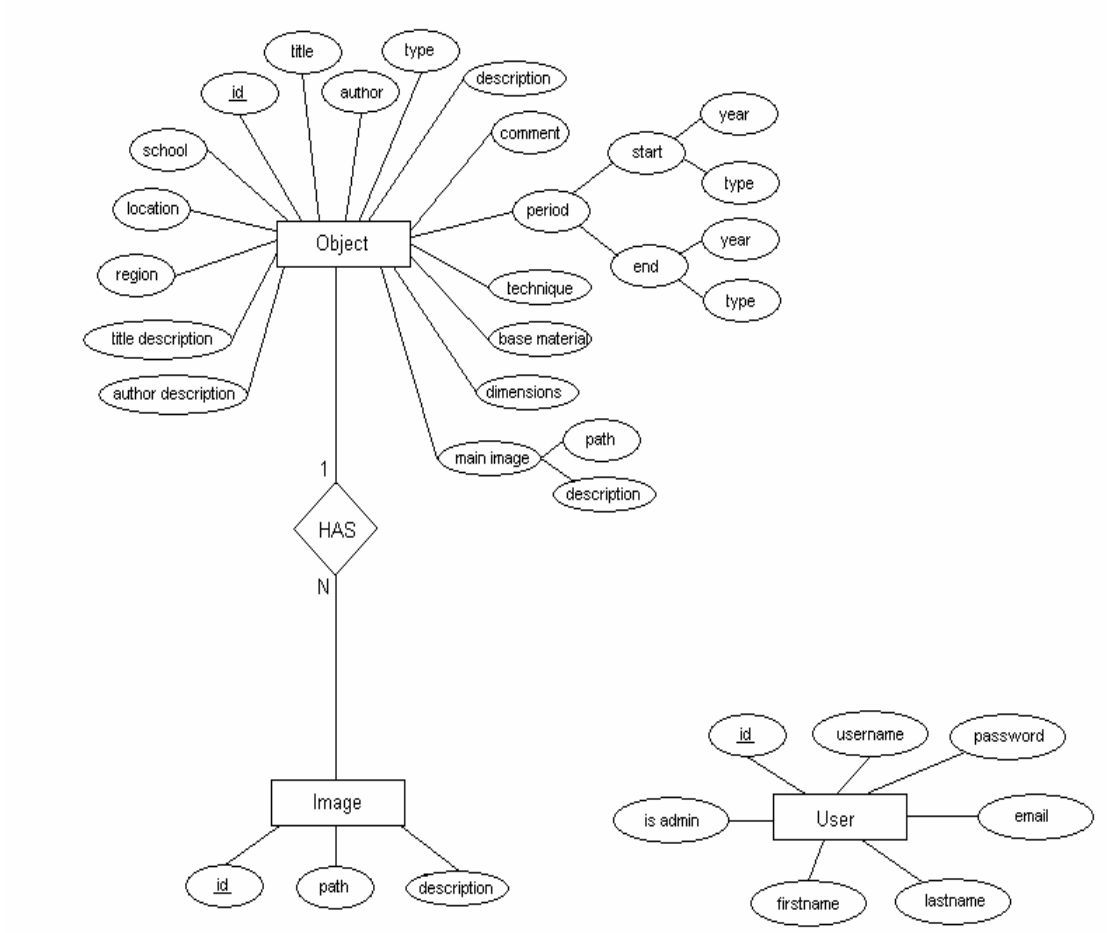


Figure 2: Entity Relationship diagram

Users are those who access the system and use it in various manners. Each user has its unique identifier (id), username, password, email address, first name, last name, and a flag which indicates whether the respective user is the administrator of the digital library. The administrator is the one who has the right to add, remove and edit objects, while the rest of the users are permitted to view them only. In order to gain access to the digital library, the users are required to register for it by specifying a desired username, which has not been already registered, a password, first and last names, and an email address. Upon initial opening of the library's website, the user is provided with a login box so that he can enter into the system, as long as he has not already done it. After a successful username and password fulfillment, she will have the opportunity to browse the separate sections of the virtual library, until her web browser is closed or until "Exit" is explicitly clicked.

The Object represents an entity of the digital library, which contains data describing a single element. An object has several fundamental characteristics – its title, author, type and main image. Apart from them, it can possess a description, comment, period, technique, base material, dimensions, school, location, region, title description, author description and main image description. The period is comprised of a start and end period and their

respective types – exact year or beginning, middle, or end of a century. Additionally, each object can be assigned multiple images, each of which has its own description.

After applying standard algorithms of deriving a relational model (which can be directly implemented into a relational database) from an entity-relationship model, three different tables emerged – Users, Objects and Images. The HAS relation between objects and images, which is a "one-to-many" relationship is realized by introducing an extra field *objected* into the Images table, which is a foreign key pointing to the *id* of the objects the images are attached to.

The source code of the demonstrator is separated into several scripts, written in PHP and using MySQL for storing the library data, which are executed via web. There are various library files, containing functions used throughout the code, grouped by their purpose, a global CSS file containing the HTML styles defining the user interface, a global JS file containing various JavaScript functions used for enhancing the user interface, and static HTML pages.

Conclusion

This article aims to present digital libraries with multimedia content as a modern technological solution for innovative presentation of the variety of Bulgarian Icon art from different artists, historical periods, and schools. For the past few years the need for a wide accessibility and popularisation of the Bulgarian icons is bigger aiming to disseminate our national cultural and historical heritage. Therefore, it is necessary that the icon's idiosyncratic art and exceptional values be made available in the global information space, so that they become accessible to both professional researchers and the wide audience. The presented project lays the foundations of the registration, documentation, and the exploration of a practically unlimited number of Bulgarian icons. The tools of the virtual encyclopaedia gives the users the opportunity to compare the icons in their historic context, so that some yet undiscovered treasured of the East-Christian iconography be manifested.

Bibliography

[i2010: Digital Libraries, 2005] "i2010: Digital Libraries", COM, 465 final, Brussels, 30/09/2005

[Krastev '05] Krastev D. (2005), Central Library of Bulgarian Academy of Sciences – present and future, The libraries of Bulgarian Academy of Sciences, Reference book, Bulgaria

[Kiernan at al.'03] Kiernan K., Kekhtyar A. (2003), EPT: Edition Production Technology for Multimedia Contents in Digital Libraries, Presented on Workshop on Multimedia Contents in Digital Libraries, USA

[Paneva at al. '05] Paneva D., Pavlova-Draganova L., Draganov L. (2005), Digital Libraries for Presentation and Preservation of East-Christian Heritage, "Generic Issues of Knowledge Technologies", Proceeding of HUBUSKA Second Open Workshop, Budapest, Hungary, 75-83pp.

[Pavlov, Paneva '05] Pavlov R., Paneva D. (2005), Towards a Creative Exploitation of Digitised Knowledge in eLearning Systems, Presented on Second CHIRON Workshop "Innovative Technologies and Solutions for Ubiquitous Learning", Paris, France

[Pavlov at al. '05] Pavlov R., Paneva D., Pavlova-Draganova L., Draganov L. (2005), Digital libraries with multimedia content and applications in Bulgarian cultural heritage (Analytical study), State Agency for Information Technologies and Communications, Sofia, Bulgaria

Authors' Information

Lilia Pavlova-Draganova – Laboratory of Telematics, BAS, Sofia, Bulgaria, 8, Acad. G. Bonchev str.; e-mail: lilia@cc.bas.bg

Vladimir Georgiev – Institute of Mathematics and Informatics, BAS, Sofia, Bulgaria, 8, Acad. G. Bonchev str.; e-mail: vlado80@gmail.com

Lubomil Draganov – Institute of Mathematics and Informatics, BAS, Sofia, Bulgaria, 8, Acad. G. Bonchev str.; e-mail: lubo@cc.bas.bg

SOFTWARE FOR EXPERT INFORMATION PROCESSING DURING CARRYING OUT THE EXAMINATIONS

Grigory Gnatienko

Abstract: *The system of the automated carrying out of examinations is described. Opportunities of system, its architecture and structure of communications between functional subsystems are resulted. Features of the subsystems making system are described. Types of questions which can be used at carrying out of examination are listed. In the near future the working version of system will be ready to input in commercial operation.*

Keywords*: Examination, decision-making, ranging of the answers, the automated carrying out of interrogation, types of questions, correlation of answers, registration of participants.*

Introduction

All human life can be counted by a line of examinations. Some of them are complex, others - are simpler. Some examinations are rather easily formalized; the some people in general cannot be modeled adequately.

In ability to live of the person there are spheres which should be automated with the purpose of prevention of possible subjectivity in estimations, reinsurances from abusings, disposals of nominal carrying out of examinations, etc., in particular, it is possible to rank carrying out of examinations As the specified categories on examination of " Rules of traffic » which are easily enough formalized and give ample opportunities for abusings; regular examinations by representatives of separate trades of "Safety precautions regulations" which frequently have formal character, are routine, but are necessary in view of the current legislation; carrying out of examinations for specialities which it is possible to find room in narrow frameworks of possible variants of examination cards; and a plenty of other real situations of preparation and carrying out of examinations.

Lack of the majority of the examinations spent today in different spheres of human ability to live, the impossibility of unification of requirements to their results is. It speaks that, first, the given procedure is carried out as a rule by different teachers and, second, in view of subjective factors as even in activity of one examiner significant deviations in criteria of an estimation of knowledge of students both aside overestimate of requirements are possible, and aside unjustified overestimate of estimations. It, certainly, entails an opportunity both "diligent" mistakes, and abusings at carrying out of examinations, let alone the order of results of examinations which thus become beforehand known and procedure of examinations gets only formal value, etc.

With the purpose of the automated decision of the specified problems and preparation of the user for carrying out of the specified class of examinations the automated system "Іспит" is developed. The system functions in MS Windows environment and is written in programming language Delphi. The program complex is intended for operation in the organizations and establishments where activity is connected to situations of carrying out of examinations which can be substantially formalized.

Problem Definition

At carrying out of many types of examinations the collective of experts which is engaged in organizational - methodical maintenance of carrying out of examination, as a rule is too much time and resources is compelled to give works on gathering and processing of the information. In particular, it is possible to name the following kinds of activity:

- Registration of teachers and participants of examinations;
- Distribution of teachers and participants of examinations in places of their carrying out;
- Enciphering works before check by their teachers with the purpose of increase of objectivity of check;
- Distribution of works for check and preservations of the information on the one who from teachers checked what works;
- Gathering the information on check and exhibiting of estimations, its processing, decoding of works, formation of various accounting documents and other.

Also the significant amount of other questions, which decision demands a lot of time: check of a correctness of the entered information on participants of examination and teachers, a choice of the works which have received a maximum quantity of points, for more diligent check and other important questions.

It is necessary to develop the software which fully automates process of preparation, carrying out and checks of results of a significant amount of types of examinations, and also provides conducting databases (DB) about participants of examinations and formation of accounting documents.

Opportunities of System «Іспит»

The automated system "Іспит", developed by the author, has ample opportunities:

- Allows carrying out preparation of the user for passing examinations in the automated mode, i.e. carries out training functions.
- Reduces time of carrying out of examination in a case when it is the essential factor, and procedure of examination - rather formal process to which should subject a significant amount of people.
- Practically does not change the circuit of examinations stipulated by traditional approaches.
- In case of need plays only a consultative role and the final decision on estimation is accepted by the teacher.
- Provides full anonymity of the student in case of need, i.e. excludes an opportunity for the teacher to identify answers with their author before summarizing.
- Supports unification of requirements to students within the limits of one examination as in mutual relations between people, in particular interaction the teacher - student, necessarily there are value judgment which sometimes bring undesirable factors of inadequacy in process of an estimation of results of examination.
- Excludes an opportunity of intervention of factors of bias in a course, check of results and process of calculation of estimations by results of examination.
- Provides preparation of the future participants of examination for its carrying out: testing in a mode of training.

Architecture of System «Іспит»

Architecturally the system "Іспит" will consist of the several interconnected subsystems.

SS enciphering and option of work of system it is intended for:

- Preparations of structure of system for concrete group of participants (creation of a folder and a file which identify the participant, creation of structure of folders according to each round, copying of additional files, etc.);
- Enciphering examination works with the purpose of increase of objectivity at their check;
- Options of the following parameters of system:
  - A disk and a working folder in which the database (DB) is created, works of participants of examination and files of reports;
  - Quantity of computers for check of works;
  - Maximal time of check of work of the participant for the current test.

SS data input provides:

- Input, accumulation and preservation of the information on participants of examination and teachers;
- Input and preservation of the information on "workplaces" (the name of the organization, number of a cabinet or an audience, quantity of computers, data on periphery);
- Input and preservation of the information on distribution of participants of examination and teachers in places of carrying out of examinations;
- Importation of the data from files and a DB, not descending in system "Іспит";
- Transformation of the data on examinations which are used in system "Іспит", in formats which are used by other systems.

SS checks and the analysis it is intended for:

- Formalizations of procedure of exhibiting of estimations that allows to raise objectivity of process of carrying out of examination;
- Automation of work analytics: calculation of statistics of examination, i.e. distribution of results of examination on quality as sometimes examination will consist in ranging students, instead of in definition of

absolute points (estimations always are subjective to some extent) which adequately correspond to the knowledge shown at examination;
- Selection of criteria of an estimation of knowledge by results of examinations;
- Achievements of desirable distribution of results of examination on group due to a corresponding variation of criteria of an estimation if the task of examination will consist in classification of group to some attributes;
- Input, editing and revision of results of check;
- Realizations of the mechanism of interaction with the checking program, automatic record of results of check in a DB;
- Checks of the data on completeness and consistency;
- Formations of the integrated data by results of processing the primary information;
- The analysis of changes of parameters (results of examinations) separate participants or groups of participants depending on the name of examination or during some period of time.

Service SS allows the user:
- To see while in service systems "Іспит" not only the coded data, but also the full decoded information which corresponds to these codes if this information is not confidential;
- To conduct system magazine of changes in a DB;
- To renew the data in case of emergency stops of system on fault of the user;
- To operate archive of the data.

SS preparations of reports it is intended for:
- Promulgations of the full information on a course and results of examination concerning each student in case of need its illumination;
- Creations of accounting documents: lists of participants, distribution of participants and teachers on "workplaces", the report of an encryption, the report of results of check of works;
- Generation of a regular information and reports;
- Seales of results of examination and all necessary accompanying documentation as, despite of routinism and formality of process of examination, its results frequently have great value in human life and should be reliably are documentary.

SS archive carries out the following functions:
- Copying works of participants in a DB with creation of corresponding structure of folders for each round of examination and with use of codes of participants;
- Reservation of the information from a DB for its operative use in an operating time with system.

Structure of System «Іспит»

The basic components of system "Іспит" and functional communications between them are resulted on fig. 1.

Practically from each top in which the user can have misunderstanding or ambiguity, the opportunity of the reference to help SS is stipulated; after reception of the instruction or an explanation about the further actions management from help SS is transferred in a point, the reference to it whence has been initiated.

From each top where it is possible without loss of logic of dialogue or uncontrolled change of the information, the opportunity of "recoil" on a step is stipulated back. Thus, in case of need the user has an opportunity to fulfil some steps of dialogue some times and to come back to the beginning of a dialogue session, to fulfil, for example, the branch of the dialogue which is distinct from chosen it earlier.

The system "Іспит" contains SS recordings and documenting of the information which carries out unconditional record of a course of each examination in corresponding files. This important element of a dialogue session is stipulated with the purpose of avoidance of abusings during carrying out of examinations.

The structure of system "Іспит" is resulted on fig. 2.

Figures (fig. 2) designate numbers SS which names are resulted below. A design **«→#<figure>»** designates numbers of tops to which management is transferred in case of activization of the corresponding instruction. If the figure after a position is absent, it means, that the corresponding position is a sheet of a tree of dialogue.

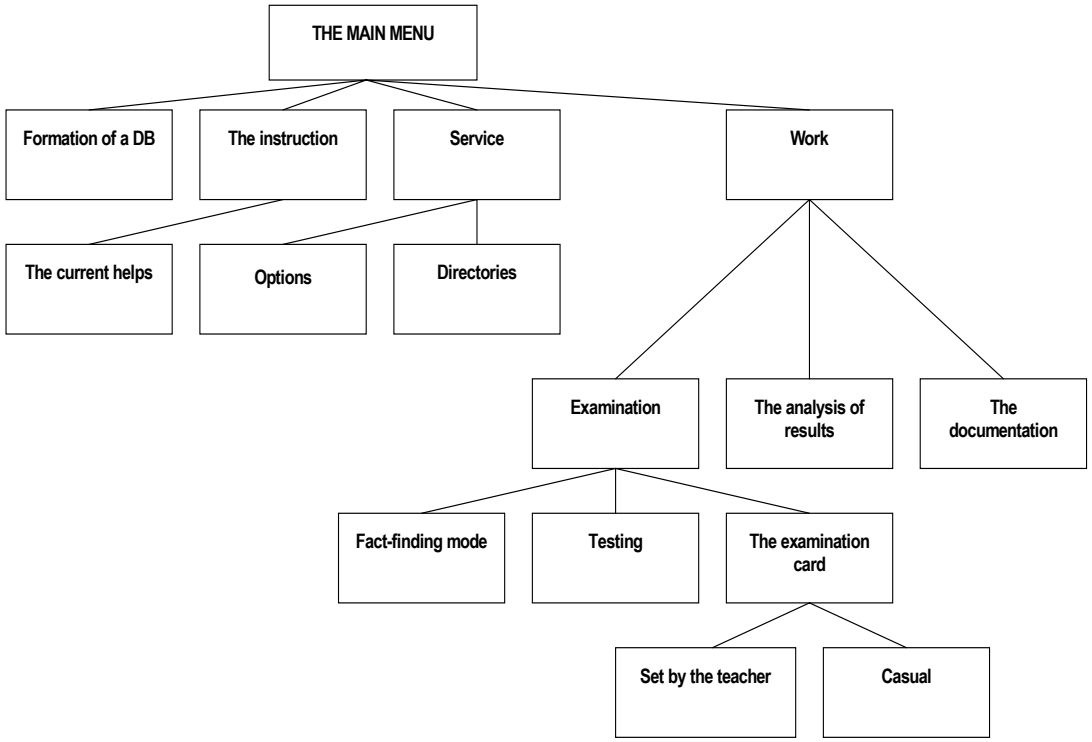


Fig. 1. The basic components of system «Іспит»

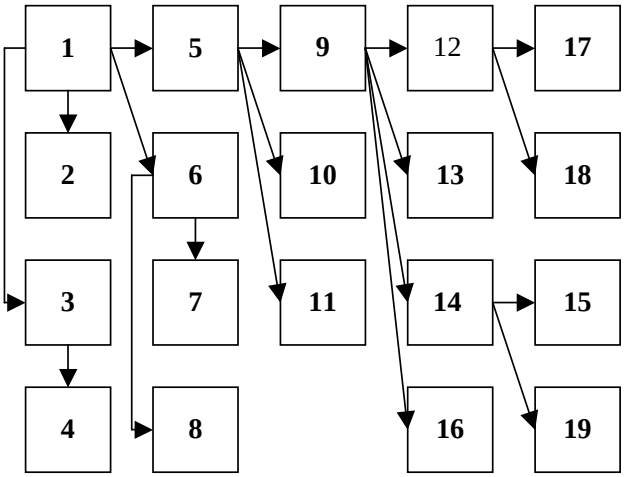


Fig. 2. The structure of system «Іспит»

#1. The main menu of system containing the following positions:

- Work → #5;
- Service → #6;
- Instruction → #3;
- Archiving → #2.

The top *1 contains also the software of advertising prompt, option of date, a soundtrack of a dialogue session and other auxiliary functions.

#2. SS "Archiving" will consist of group of programs which carry out functions of archiving, compressions, recyclings (controlled destruction) corresponding data if these actions are stipulated by the manager of system.

Archiving is carried out on a command of the user or the manager of a DB, and archiving - is unconditional, each time after the next start of system "Іспит".

#3. SS "Instruction" will consist of a DB containing all instructive materials on system, and the program managing this DB. The DB of instructions contains the full data on system as a whole and features of its functioning. SS it is intended for a selective conclusion to the screen or to a seal of the information necessary for the user in a present situation, and also for tracking a point from which the input in SS is carried out, and returnings in a point of interruption. SS "Instruction" will consist of four sections:

- The instruction on work with system ();
- The instruction on option of system;
- The instruction on archiving results;
- Current helps → #4.

#4. The program, allowing to carry out selective reading a DB "Instruction" depending on a point of dialogue, whence it is caused by the user. There is also a mode of the program, allowing to change and supplement the contents of a DB "Instruction".

#5. SS "Work" contains the following positions:

- SS "Examination" → #9;
- SS «The Analysis of results » → #10;
- SS "Documentation" → #11.

#6. Service SS allows to leave on the following operating modes:

- SS "Directories" → #7;
- SS "Option" → #8;
- Check of a correctness of a DB;
- Copying;
- Repeated indexation;
- The calculator.

The program of check of a correctness of a DB in case of its activization checks all possible connections between the DB accessible to it. In case of infringement of connections or other discrepancies there is a record in the report of dialogue. Is the preventive program which the manager of a DB uses as a rule.

Copying - a mode, allowing to speed up filling some fields in a DB or completely to copy a DB in case of their identity or when it is easier and more expedient to user to correct a DB, than to create new. Allows to copy also a DB on devices of external memory with the purpose of their preservation or operation by other program.

Repeated indexation of a DB - auxiliary function which the manager of a DB uses, and sometimes and the non-privileged user of system with the preventive purpose.

The calculator - the built - in function for the current calculations.

#7. SS "Directories" access to a DB which are directories for the user allows to provide. The reason of this action can be desire to receive the information, to create a new DB, to change the information in existing DB, to destroy some DB, etc. Clearly, that performance of these actions depends on a priority of the user, its level of access. There is an opportunity of access to the following directories:

- The information on examinations;
- Names of examinations;
- Teachers;
- Groups;
- Students;
- Circuits of examinations;
- Structures of the screen;
- The contents of questions.

«The information on examinations» will consist of references to examinations which already surrendered and to which there can be a necessity of the repeated reference. To this information concrete values from directories of "The Name of examinations», "Teachers", "Groups", «Circuits of examinations», «The Contents of questions» belong also.

Directories of "the Name of examinations», "Teachers", "Groups", «Circuits of examinations», «The Contents of questions», "Students" contain the listed data which are available in corresponding DB of system "Іспит".

#8. SS "Option" is intended for a choice from the DB listed in SS "Directories", concrete values for option for the current session of dialogue. SS "Option" contains the following positions:

- The information on examination;
- The name of examination;
- The teacher;
- Group;
- Students;
- The circuit of examination;
- Structure of the screen;
- Questions;
- Formation of examination cards.

#9. SS "Examination" allows the user to set one of such types of examination:

- Fluent interrogation → #16;
- Testing → #13;
- Examination cards → #12;
- Casual choice → #14.

#10. SS «The Analysis of results» it is intended for calculation of the following sizes:

- Correlations of answers of users in various metricss with use, in particular, the mathematical device described in works [Волошин, 1997], [Гнатієнко, 1997], [Гнатієнко, 2000];
- Classification of results by a technique described in article [Гнатієнко, 2001];
- Calculation of an estimation of the user for its answers, etc.

#11. SS "Documentation" is intended for preparation of the data on a course of examination, record in corresponding DB and seales in case of need. The majority of functions of this SS is carried out certainly, irrespective of the teacher, students and other users.

#12. "Examination cards" - the mode specifying type of examination at which on the screen there are numbers of examination cards and is simulated a choice by the user of the ticket "dropped out" to it. This mode has the following positions:

- «Set by the teacher» → #17;
- «The Casual ticket» → #18.

#13. "Testing" - a mode at which the circuit of examination such as sociological, psychological or other testing (questioning) is realized, i.e. is offered to answer the user all set of questions which contains in a corresponding DB and appears on the screen.

#14. « The Casual choice » - a mode during which a quantity of questions gets out of a corresponding DB depending on a choice of the user:

- The limited quantity of questions → #15;
- Any questions → #19.

#15. «The Limited quantity of questions» - is offered to answer the user the set quantity of questions from a corresponding DB.

#16. «Fluent interrogation» - a mode during which in the casual image from any concrete DB the significant amount of questions gets out. This mode differs from previous the casual order of a conclusion of questions.

#17. «Set by the teacher» - a mode at which the teacher in the beginning of examination forms examination cards on themes, to own representations about complexity of questions and to the other criteria. The user is offered to answer not the casual ticket, and on one of generated by the teacher.

#18. «The Casual ticket» - a mode at which examination cards are formed not by the teacher, and casual image. Thus the requirement of presence in the ticket of questions on various themes of examination is kept only.

#19. «Any questions» - the mode allowing the user to answer casually initiated questions from an any sphere of activity (or a concrete DB) without restriction of quantity of questions before compulsory end of a dialogue session.

Types of Questions

For expansion of opportunities and spheres of application of system "Іспит" various types of questions have been developed and realized as the corresponding software. We shall result their brief description.

T#1. <Variants_of_answers> <A_degree_of_importance_1> <A_degree_of_importance_2> ... <Number_of_a_"corresponding"_question>.

The type of a question admitting answers on such format, belongs to a class of examinations such as « the Rule of traffic » and a little more difficult. To each variant of the answer the degree of importance and a degree of reliability of the answer is put in conformity. After a choice the user of a variant of the answer calculates the integrated estimation as product: <Degree_of_importance>* <A_degree_of_correctness>.

T#2. Values: <fixed>, <bottom>, <top>.

Questions of this type are characteristic for examinations such as "Safety precautions regulation". In some cases any from the values above mentioned, act as variants of answers.

T#3. An insert of symbols. It is used in a case when examination will consist in testing the student for knowledge of some rules or laws. Necessity for use of such type of questions arises in case of examination, for example, grammatic rules during modelling a dictation. The student should enter or not enter in the specified fields the missed either unreasonably changed letters or other symbols, words, word-combinations and so forth. To Test texts there correspond variants of right answers with which the answers received from the student are verified. The metrics for definition of a mark estimation of answers of the student is entered.

T#4. Test questions. In this case variants of answers are verified with test. The teacher sets (chooses from the menu) criterion (or criteria) correctness of answers to the test as a whole. In concrete cases it can be some formulas or the whole algorithm of calculation of the integrated estimation of the test. If the integrated estimation is calculated by the unique formula this formula is interpreted by the special block of system "Іспит".

T#5. Complicated questions. The opportunity to generate the answer as the offer is given the user. Correctness of the answer is checked by comparison of this offer to patterns.

T#6. Ranging of answers. The user should enter variants of answers by way of reduction of their correctness. The distance from the set ranking answers up to correct is calculated in various metricss. The estimation of the student is generated under the formulas chosen the teacher.

Clearly, that the resulted types of questions do not cover all variety of situations of carrying out of examinations and decision-making by their results. In this connection the system "Іспит" is open for addition with other functions which model situations of carrying out of the examinations, admitting formalization.

Conclusion

The demonstration prototype of system "Іспит" which passes check in the several organizations is developed. The technology of the automated preparation for carrying out of examinations is constantly improved. In the near future the working version of system will be ready to input in commercial operation.

Bibliography

[Гнатієнко, 1997] Гнатієнко Г.М., Єпик Н.Б. Про визначення міри схожості вподобань експертів // Вісн. Київ. ун-ту. Фіз.-мат. науки. Київ, 1997, №3. С. 159-165.

[Волошин, 1997] Волошин О.Ф., Гнатієнко Г.М. Проблеми узгодженості експертної інформації в задачах прийняття рішень // Збірник наукових праць Міжнародної конференції “Знание-Диалог-Решение” KDS-97, Ялта, 1997. Т. 2. С. 332–335.

[Гнатієнко, 2000] Гнатієнко Г.М. Деякі математичні аспекти соціальної експертизи//Соціальна експертиза в Україні: методологія, методика, досвід впровадження / За ред. Ю.І.Саєнка.-К.: Ін-т соціології НАНУ, 2000. 194 с.

[Гнатієнко, 2001] Гнатієнко Г.М. Визначення міри схожості експертних розподілів об’єктів за належністю до кластерів // Вісн. Київ. ун-ту. Фіз.-мат. науки. Київ, 2001. № 3. С. 220–223.

Author’s Information

Grigoriy M. Gnatienko – T. Shevchenko Kiev National University, Faculty of Cybernetics, Dr.Ph., Kiev, Ukraine; e-mail: G.Gnatienko@veres.com.ua

LECTURER'S WEB-SITE AND ITS ROLE IN DISTANCE LEARNING

Evgeny Zabudsky

***Abstract**: Structure of University lecturer's web-site is suggested. A need for higher education system hyperspace is demonstrated.*

***Keywords**: Internet, web-site, information, electric machines.*

***ACM Classification Keywords**: K.3.1 Computer Uses in Education: Distance learning*

Introduction

Complete information availability for students and entrants is one of the aspects of distance learning [1]. For the range of considered subject, information is a combination of the details required for the specialization selection and educational learning, viz. state standard of specialization, educational plan, working program, multimedia books and school books, audiovisual course books, virtual lab equipment, etc.

An access to information is provided by means of Internet and computer technologies. Interactive TV will also contribute into distance learning development in the forthcoming future.

Information is located on the Web-sites of Russian Ministry of Education, University, Faculty, Department and Professor (Lecturer). It's required to develop typical structure of the listed Web-sites. This helps universities to actively participate in practical distance education, whereas students are involved into educational process.

Suggested structure (model) of lecturer's Web-site realized by the author is given below (http://zei.narod.ru).

Lecturer's Web-site: Typical Structure

1. Brief details of place of employment, position, academic degrees and scientific titles (Curriculum Vitae), contact information.

1.1. Scientific specialization and performed science projects.

1.1.1. Scientific specialization.

1.1.2. Performed science projects. Titles of main scientific research works, customers, etc are listed. Summary with photos is given for each work is given in html-form.

1.2. Lecture courses and publications.

1.2.1. Lecture courses. Lecture courses titles are listed. Working program is given for each course. It's necessary for distance learning to represent tutorial and methodological suite of the discipline on the Web-site.

1.2.2. Publications. The following lists are given as separate sheets (html-pages): monographs and course books for universities; sets of registered computer programs; certificates of recognition and patens; scientific articles; reports on science and technical conferences; reports on science and methodological conferences, methodological notes and manuals. Each list includes only major publications. Full text is given for resulting scientific papers and reports (html- or pdf-files).

To increase contacts with foreign colleagues its worth translating the Web-site into (http://zei.narod.ru/engl/index.html).

Publication of course texts, labs and equipment descriptions is a further development of lecturer's Web-site [4]. All texts and descriptions should be developed with multimedia features: hypertext, sound, animation, etc. Use of different colors and fonts is also an important methodological tool. Multimedia capabilities allow transforming manuals and course books to new quality level that is much deeper than traditional ones. An illustrating example follows. Course books on Electrical Machine (represented by author) could be written on the basis of Field Theory [3, 5] rather than Circuit Theory [2] that is a current forced choice. Magnetic field is a 'working body' of electrical machines and transformers. For its understanding well developed abstract thinking technique is required. Most listeners don't possess such a technique (at least on the beginning stage of education). Multimedia features allow

visual representing of magnetic field creation in the given devices as well as filed existence through spatial-temporal continuum [1, 6].

Thus, no matter how far listeners are located from the university, they have an access to almost all materials required for successful studying of any discipline via Internet.

It's evident that professors', departments', faculties' or universities' Web-sites as well as Russian Ministry's of Education and some other industrial Ministries' Web-sites should be integrated into unique *Higher Education System Hyperspace*.

It was already done a lot to realize this. Particularly Web-site http://informika.ru contains hyper-textual lists of Russian universities and other useful information. However it's still necessary to do much more.

Department's Web-site should contain the following initial information: history of department; contact details; characteristics of specialization that is offered by the department (it's practicably to put together hyper-textual links to Web-pages of companies and organizations that finally employ department's graduates; State Standard of specialization; educational plan of specialization; hyper-textual list of staff members with their positions, etc).

Faculty Web-site's information could be as follows: history of faculty; hyper-textual list of Faculty Scientific Board members; contact details; hyper-textual lists of specializations with their State Standards; hyper-textual list of departments involved into educational process, etc.

University Web-site (http://www.msau.ru/) should contain the following basic information: university history; university top-staffers and hyper-textual list of chancellor's sub-organizations; hyper-textual list of University Scientific Board members; contact details; hyper-textual lists of faculties, departments, etc.

It will be useful for entrants to include URL-addresses of Russian Universities into periodically published Entrants Hand-Books.

Basic information provided by *Russian Ministry of Education* should be as follows: brief history of Russian educational system development; hyper-textual list of Ministry's and sub-divisions' top-staffers; hyper-textual lists of State and Private Universities in Russia; hyper-textual lists of specializations available for students in Russia, etc.

It's also necessary to create material and resource pre-conditions to solve the task. University lecturer should have an opportunity to use Internet's information collection working with a computer set up at his working space (university department). Moreover if lecturers' home PC's are global network connected that'd be even more effective. Actual experience with Internet opens unlimited net capabilities in terms of access to global information resources, contacts development among colleagues, relations management with manufacturers, organizations, etc.

Conclusion

Lecturers', departments' and faculties' Web-sites are less developed for the time being. Community initiative is of primary importance for higher education united hyper-system development. Internet development in Russia goes fast due to such an initiative though our country faces well-known difficulties and problems.

Bibliography

1. Забудский Е.И., Павлов М.В. A Geometric interpretation of the calculation of the Magnetic Field in electromechanical devices // 1-я Международная конференция по дистанционному образованию в России "Дистанционное обучение и новые технологии в образовании". Российская Академия Наук, Москва, 1994.
2. Забудский Е.И. Математическое моделирование управляемых электроэнергетических устройств: Учебное пособие для вузов. – Ульяновский государственный технический университет, Ульяновск,1998. http://zei.narod.ru/Up1.html .
3. Забудский Е.И. Анализ управляемых электроэнергетических устройств методом конечных элементов: Учебное пособие для вузов. – Московский государственный агроинженерный университет, Москва,1999. http://zei.narod.ru/Up2.html .
4. Забудский Е.И. Электрические машины. Ч. 1. Трансформаторы. Учебное пособие для вузов. - Московский государственный агроинженерный университет, Москва, 2002. http://zei.narod.ru/soderghanie.html .
5. Забудский Е.И. Совмещенные регулируемые электромагнитные реакторы. - Энергоатомиздат, Москва,:2003. http://zei.narod.ru/regreak.html .

6. . Забудский Е.И. Компьютерный фильм "Геометрическая интерпретация результатов расчета магнитного поля в электромеханических устройствах". // Каталог сертифицированных программных продуктов по электротехническим дисциплинам. Сертификат # 3/93 от 10.09.93 г. - Астраханский государственный технический университет, Научно-методический совет по электротехнике Минобразования РФ, Астрахань,1993. http://zei.narod.ru/filmr/film.html .

Authors' Information

Evgeny Zabudsky - Moscow State Agro-Engineering University, Professor Dept. of Electric Power Supply and Electrical Machines; Timiryasevskaya St., bl. 58, Moscow-127550, Russia; e-mail: zei@inbox.ru ; web-site http://zei.narod.ru .

E-COURSE OF THEORETICAL MECHANICS [5]

Elena Ponomaryova, Tatiana Nevenchannaya, Vladimir Pavlovsky

Abstract: *The concept, structure and contents of the Internet textbook on classical mechanics intended for Higher Technical Institutions are presented in this work. Aspects of program realization of textbook applications and the technology of elaborating the textbook in the "Hecadem" Internet-teaching environment are given too.*

Keywords*: distant education, electronic textbook, Internet, mechanism.*

Introduction

The achievements in the field of information technology help to work out and use computer learning systems (CLS) in the course of teaching process, aimed at performing various types of teaching activity. The main principles of new CLS are the following:

- interactiveness;
- dynamics;
- integrativeness (program compatibility with other products);
- Web-compatibility;
- work providing in the off-line mode and distance mode;
- updating, replenishing with new material by subject teachers;
- modeling and simulation of examined processes and phenomena;
- visualization of educational information;
- illustration by sound aids;
- various services;
- control of knowledge with error diagnostics and feedback.

CLS is a more advanced means as it not only intensify non-textual parts (graphics, animation, video, sound) and various services, as well as, wide opportunities of Internet network.

Such systems we will name hereinafter as distance teaching systems (DTS). Now we consider the most common opportunities of DTS (Fig.1).

The first type (semantic elements): helps to study theoretical material of a subject, consolidate the know- ledge of the material covered by solving problems and performing calculations, carrying out computer experiments and

[5] The work is done with grant №**06-07-89195** support investigations.

exercising intermediate and final knowledge control. These options bring closer DTS and traditional teaching process.

The second type (multimedia options): includes illustration of teaching material with the help of video and sound, animation and graphics as it helps to improve presentation of teaching material and arouse interest for learning.

The third type (interface, services): it includes options of creation of multilevel menu system, multiple window interface, system of hypertext jumps, text look-up. It helps to avoid laborious calculations and it accelerates the process of solving tasks as well as it helps to motivate learners for study and to liven up the work.

The fourth type (Web options): it helps to provide interactive link of a learner and a teacher through Internet (from primitive E-mail to the formation of off-line or on-line conferences), it provides searching of scientific and educational information in Internet and it helps to use references, etc.

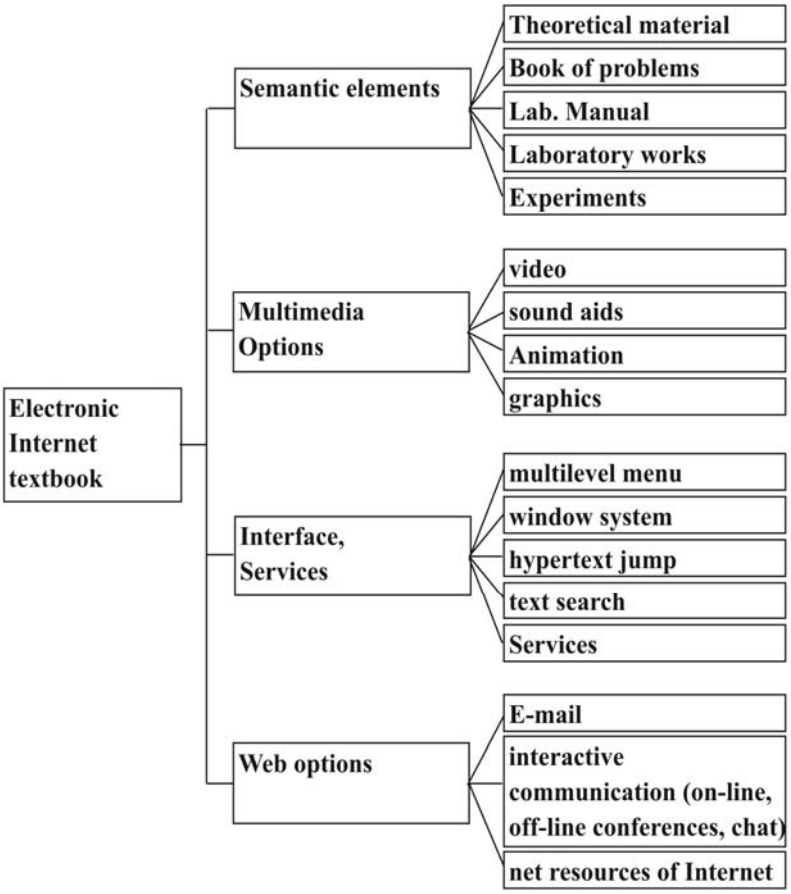


Fig 1. DTS possibilities

Astrakhan State Technical University together with M.V. Keldysh Institute of Applied Mathematics of Russian Academy of Sciences developed technologies of creation computer distance/local learning system and so DTS on theoretical mechanics was created for this purpose [Pavlovsky, 2003], [Nevenchannaya T.O., 2003].

Computerization learning is of great independent and interdisciplinary importance in the field of theoretical mechanics as this subject is the basis for learning general engineering subjects and special courses as it teaches to highlight the modals for solving formulated problems, to lay down conditions, to solve problems, to determine the limits of their acceptability.

Developed teaching system includes the following parts:

- organization and information block;
- electronic Internet-textbook on theoretical mechanics (certificate on official registration of computer program №2004612620, Rospatent);
- manual aid “Theoretical mechanics on computer”.

Let us consider each part in detail.

1) Organization and information block.

Organization and information block includes: course curriculum with indication of subject list and date of their performing; schedule of laboratory works, seminars and classes; syllabuses; programs; teaching charts; problem exercises; manual aid on theoretical mechanics.

2) CLS on theoretical mechanics.

Let us illustrate the designed textbook (see [1-3] as well) and for its illustration we include the proposed technologies of creation DTS. They are based on the technology and host system of differentiable Internet training "Hecadem" in distance version [4]. Pattern of electronic Internet-textbook is illustrated in fig. 2.

Two versions of textbook were worked out; local and distance ones. These two versions of electronic textbooks were designed in shareable formats and they are Web standards.

Distance version of the textbook assumes the work by global information Internet and it supplies distance development of the textbook in Internet; creation of general array of knowledge elements on mechanics; formation of new teaching courses on mechanics type on the base of created array of knowledge.

Local version of the textbook is designed for work at autonomy device or in local network.

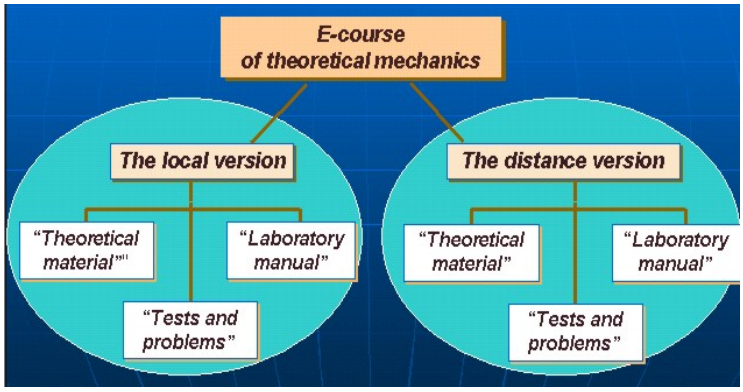


Fig. 2. Textbook pattern

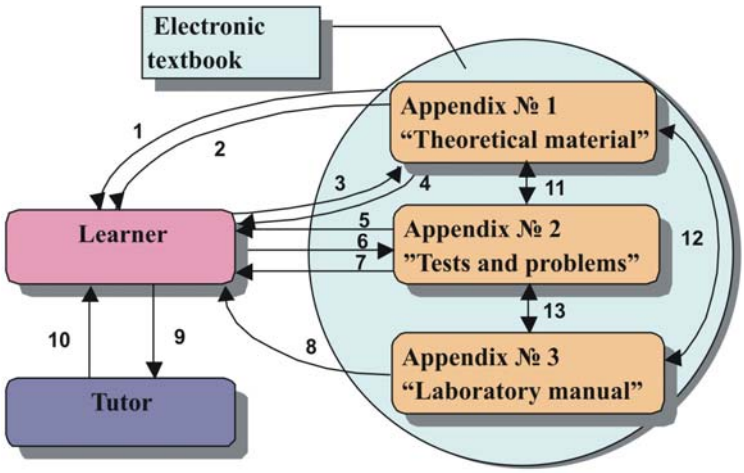


Fig. 3. Scheme of interconnections of a learner and electronic textbook

Each version of the textbook includes program appendix:

- "Theoretical material",
- "Tests and problems",
- "Laboratory manual" (fig. 3).

When working with appendix № 1 a learner is offered to study theoretical material (1), including the main definitions, theories, illustrations for worked standard problems. Each section has been studied, the learner answers the questions (self-evaluation) (2).

Fed answers (3) are transmitted back to appendix № 1, which returns the result to the learner after answers being analyzed (4).

When applying to appendix №2 a learner is offered a set of tests and problems (5), and he/she enters the result of a solved problem (6), then it is carried to appendix №2. The answers are analyzed by the appendix and then the result is reported (7).

The interaction of a learner and appendix № 3 assumes the participation of a tutor and it is reduced to the following steps: the appendix offers a learner design basis and calculated and graphical work (8). The work being done; the learner passes it to the tutor (9). The tutor evaluates the correctness of work, then the learner is informed the result (10).

Let us consider these appendixes.

Appendix "Theoretical material" contains the theory of studied subject with the examples of solved standard problems. The appendix is performed in the form of two realizations: realization № 1, or "Course of lectures" where theoretical material is presented in a logical order as in a typical textbook; realization № 2, or "reference-précis", structured according to semantic elements. Hyper textual references help to work on this reference material.

Theoretical material of the textbook is presented in the following parts: “Kinematics”, “Statics”, “Point dynamics”, “Dynamics of a system”, “Analytical mechanics”, “Calculation of mechanisms”.

The first five sections are included into traditional course on theoretical mechanics. Section “Calculation of mechanisms” is also put into the textbook. The textbook is an original authors’ teaching aid, the main concepts of mechanics being presented here as well as intersubject approach to the calculation of mechanisms and it helps to use knowledge from different fields for solving the problems [Nevenchannaya T.O., 2003].

Concepts of composing kinematical design on machine engineering drawing are presented here (fig. 4).

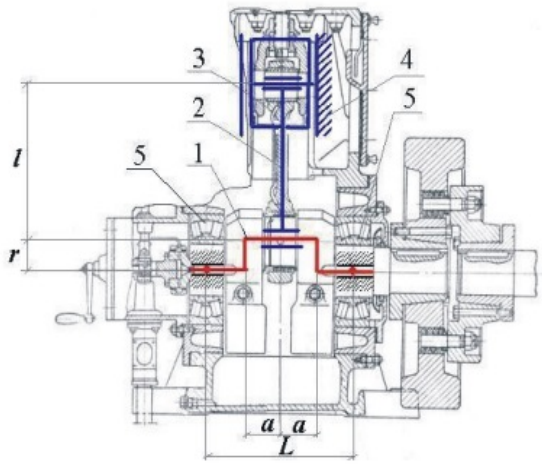


Fig. 4. Machine engineering drawing

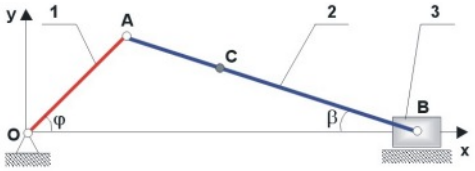


Fig. 5. Kinematical design

Link OA – crank - executes rotary movements.

Link AB – coupler - connecting rod-moves plane parallel.

Link B – slider - reciprocating along the horizontal guide.

Point O with coordinates (0; 0) is stationary.

Execution of crank-slider mechanism in the form of a rod an in the form of a circular links are examined in the kinematical circuit design (Fig. 6).

Fig. 6

Presentation of theoretical material is illustrated by a number of Web pictures, including animated illustrations 2D-, 3D-models of mechanical objects, mechanisms in particular (Fig. 7). Animated models of mate mechanisms, ellipsograph, scotch-yoke mechanism, and link gears with swinging and rotary link, planetary gears with inner and outer catching and many others are worked out and used now.

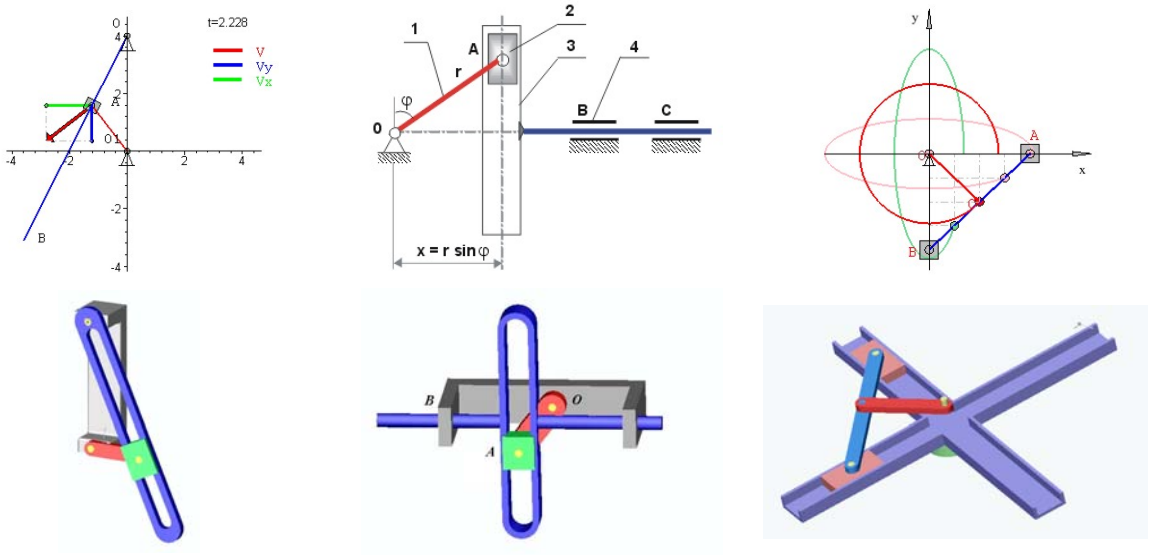
The models illustrate total cycle of mechanism movement, they are parametric and complicated.

Development of each model is rather labour-consuming process. Let us discuss the technology of producing 2D-, 3D-animated models of mechanisms in more detail.

A) Creation of computer animated models in mathematic packing MAPLE.

Options of MAPLE in the field of development animated graphics help to create 2D-, 3D-animated models of real mechanisms and to deduce graphical plots of parameters to time.

Sequence of operations for production of animated model in the environment of mathematical packing MAPLE is presented in Fig. 8.



a) Coulisse mechanism b) Scotch-yoke mechanism c) Ellipsograph

Fig. 7. Models of mechanisms (illustrations)

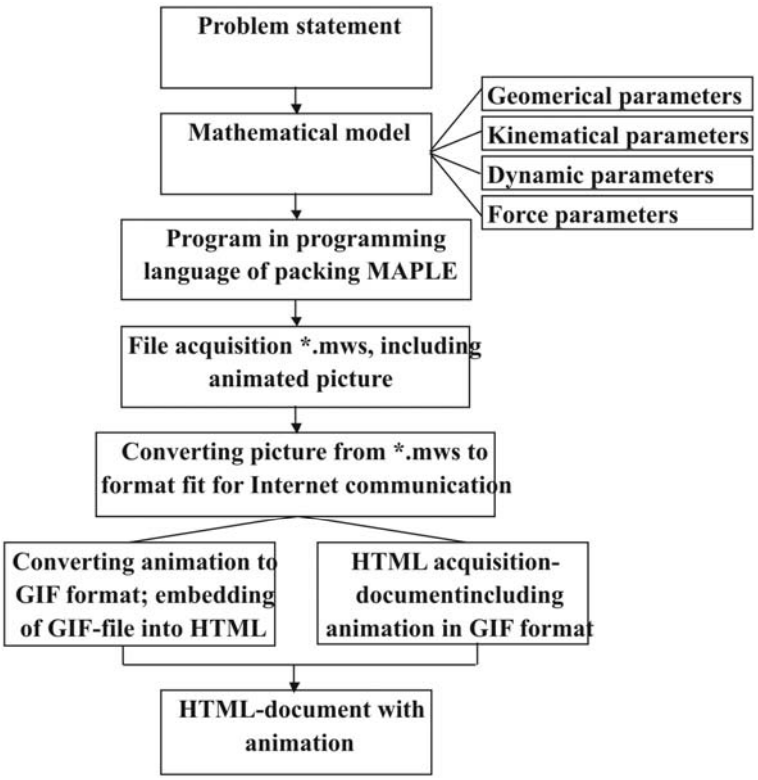


Fig. 8. Production of animated models in packing MAPLE

1). Problem statement includes the description of input and output parameters.

2). Composition of mathematical model of mechanism: description of geometrical, kinematical, force and dynamic parameters is used depending on handling a problem:

- geometrical parameters: link dimensions, coordinates of datum points, link slopes with respect to axis;
- kinematical parameters: velocities, accelerations of datum points, angular velocities and acceleration of all links;

- force parameters: loads (single and distributed forces, moments), applied to parts of a mechanism;
- dynamic parameters: mass, force and moments of inertia.

3) A new program is being prepared on the base of a constructed mathematical model in the built-in programming language MAPLE.

Coordinates x_A, y_A, x_B, y_B of points *A*, *B* of a mechanism are defined as time functions (in the interval from *0* to *T*, where *T* - is a time period of one revolution of *AB* crank), and links *OA* and *AB* are set as lines with coordinates $\{(0,0),(x_A, y_A)\}$ and $\{(x_A, y_A),(x_B, y)_B\}$ accordingly and slider *B* - as a rectangular.

2D-model of a crank mechanism (is included in file *.mws) is the result of executed program. Some frames illustrating the position of mechanism at any instant are shown in fig. 9.

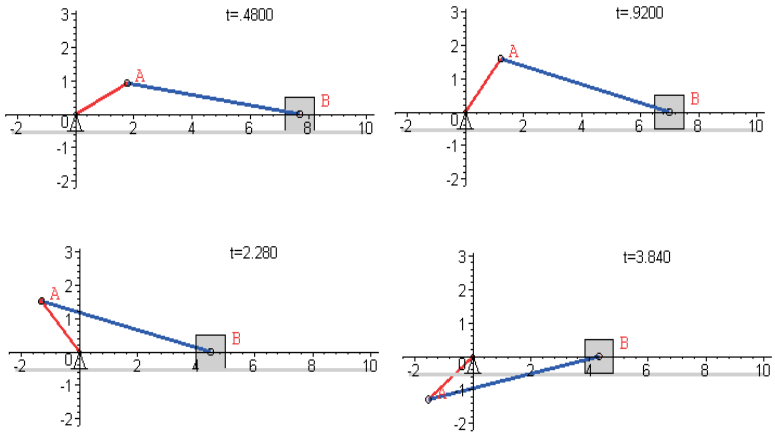
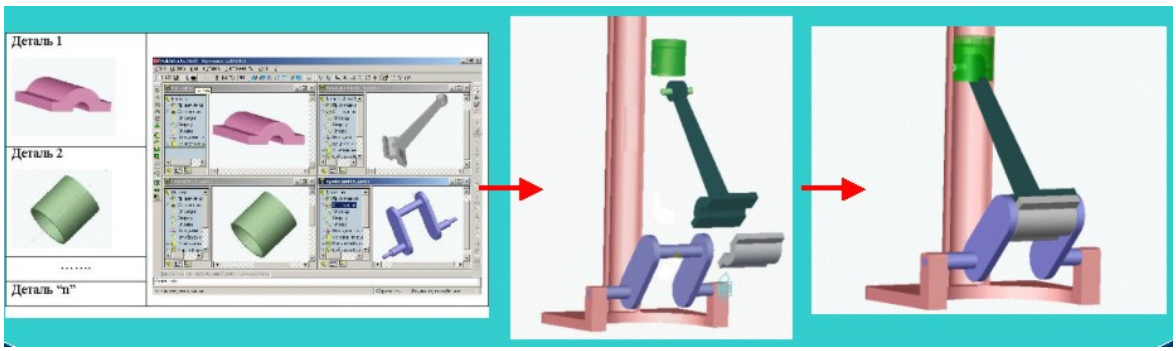


Fig. 9

4).Converting. Manufactured model is converted into WEB –format (CIF) and is built-in into HTML-page.

B). Animated system of designing SOLID WORK together with program set DYNAMIC DESIGNER MOTION (ADAMS) is used as alternative environment of computer modeling.

Let us consider the acquisition of 3D solid model as illustration of crank mechanism.



a) b) c)

Fig. 10. Design of three-D parts

The first stage-formation of 3D parts of a mechanism (Fig. 10, a) and it is necessary:

a) to draw a flat sketch, made with the help of simple primitives, entering an outline;

b) to “define” an outline (preset dimensions, correlations: parallelness, perpendicularness, concentricness, equality, etc.);

c) to design 3D parts (outline drafting by pressing out along directive, rotating along the axis, drawing jackets, holes, bevels, etc.)

The second stage-uniting parts into assembly (Fig. 10, b) with the help of applying so called “integration” on them (conditions, coincidences, parallelness, perpendicularness, concentricness, etc.)

Limiting their relative mobility.

3D solid-state parametrical model of a mechanism (Fig. 10, c) is received as a result, the model being passed to the partner systems such as: DYNAMIC DESIGNER MOTION (DDM).

The third stage-receiving an animated model in the packing DDM. In order to do it animated parameters are added to the geometrical model, worked out in SOLID WORKS: leading link is chosen, assigned a view of its movement, the number of frames of animation, time of its reproduction. It is possible to watch colored 3D of moving mechanism, design its single parts in space, and examine kinematics and dynamics on the monitor screen (Fig. 11).

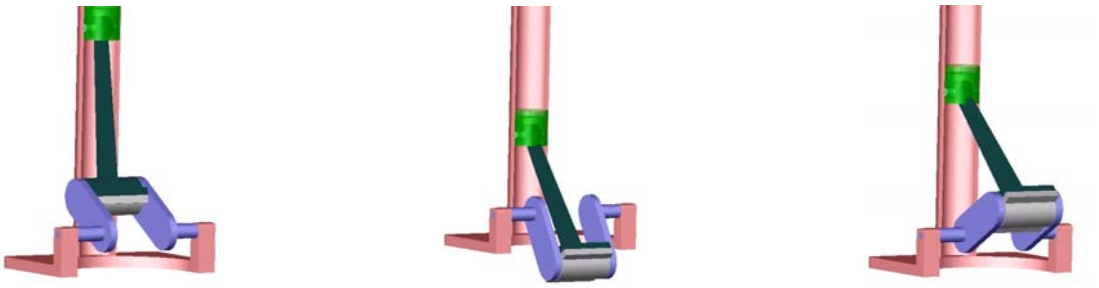


Fig. 11. Crank mechanism in rod running

Received animated model can be kept as a demo reel (in format *.AVI) or in format *.WRL (VRLM, Virtual Reality Modeling Language) which is a standard for passing 3D virtual scripts into Internet. It gives opportunity to embed reference for them into HTML-pages and publish them in Internet.

The fourth stage- display of graphical results. Options of SOLID WORKS and DDM programs help to display designs of kinematical (link velocities and acceleration), dynamic (forces, moments) and other necessary parameters from time to the screen.

Working with such DTS the user can not only study principles of analytical solving of mechanical problems but to become skillful at computer modeling.

Operation with appendix “Tests and problems” helps to:

- learn skills in solving problem;
- exercise self-control of knowledge (local version), intermediate and final control of knowledge (distance version);
- form test tasks (subsystem “Test Constructor”)
- carry out monitoring of learners’ standard of knowledge.

This subsystem registers the results of tests of each learner, the number of attempts and time spent on answers.

For assessing the proficiency we use such test assignments as: “yes-no”, “multivariate choice”, “list regulating”, “and input of one’s own answer”.

Appendix “Tests and problems” contains more than 300 test assignments in the sections of theoretical mechanics. At present the base of assessing proficiency is expanding.

Appendix “Laboratory manual in mechanics” contains a set of computer calculated-graphical works; each of them includes the name, aim, short theory, instructions in details for performing the tasks, condition of a problem, variants of assignments. File-samples are given to help a learner to solve standard problems on statics, kinematics and dynamics on the course of theoretical mechanics.

3) Teaching Aid.

Worked out DTS is used together with the teaching aid “Theoretical mechanics on computer”. This aid is a practical manual for the work with teaching system in local and distance modes. The manual also contains solution algorithms of standard problems on mechanics in the mathematical packing MAPLE environment,

description of comprehensive calculation of gears and it helps the learners study course of theoretical mechanics for technical universities independently and to solve engineer's problems on computers.

DTS on theoretical mechanics passed evaluation test in the teaching process at Astrakhan State Technical University in 2003 and it showed a good performance.

Problems of program realization were solved when developing DTS. They include the following:

- format choice with due account of their mutual compatibility and opportunity of transfer over Internet;
- choice of compliant software and realization of technological chain development of DTS;
- possibility of providing local and distance operation with appendixes;
- providing of textbook updating, adding new teaching material in particular. In the future it is presumed to develop an integral distance information environment on a mechanical type worked out on the base of DTS, an environment, where it is possible to form an integral array of knowledge elements (such as mechanics) and to realize dynamic access for reference data on various specialties.

Apart from reference and teaching theoretical material there will be provided model opportunities, control functions and administering.

Bibliography

[Pavlovsky, 2003] Pavlovsky V.E., Nevenchannaya T.O., Kurganskaya G.S., Ponomaryova E.V., "Concept, structure, program realization of electronic Internet-textbook: preprint IAM M.V. Keldysh RAS; №39, M., 2003, 28 p.

[Nevenchannaya T.O., 2003] Nevenchannaya T.O., Pavlovsky V.E. Ponomaryova E.V. "Integrated calculation of mechanisms in electronic textbook on theoretical mechanics»: Preprint M.N. Keldysh IAM RAM, №38, M., 203, 24 p.

[Matrosov, 2001] Matrosov A.V., Maple 6. "Solving of problems on higher mathematics and mechanics". B.P.b.:6. LXB. Petersburg, 2001, 528 p.

Authors' Information

Elena Ponomaryova - Astrakhan State Technical University, Associate Professor; P.O. Box: 414056, Astrakhan, Taticheva, 26, app. 53, Russia; e-mail: acmpax@rambler.ru

Tatiana Nevenchannaya - Astrakhan State Technical University, Professor; P.O. Box: 414025, Astrakhan, Taticheva, 16, Russia; e-mail: nevento@rambler.ru

Vadimir Pavlovsky - The Keldysh Institute for Applied Mathematics, Russia Acad. of Sci., Professor; P.O. Box: 125047, Moscow, Miusskaya, 4, Russia; e-mail: vlpavl@mail.ru

DESIGN STANDARDIZED WEB-COMPONENTS FOR E-LEARNING

Andrey Belonogov

***Abstract**: In this paper a flexible approach to design LMS with QTI Ready component based on the e-Learning standards AICC and IMS QTI is described. This system and component permits a dynamic learning and assessment process. QTI Ready component can provide these facilities to other real world virtual learning management system.*

***Keywords**: Web-based education, e-Learning technologies, metadata schemes, standardization and open standards, Servlet technology.*

***ACM Classification Keywords:** D.2 Software engineering – Standards: D.2.11 Software Architectures - Patterns (e.g., client/server, pipeline, blackboard); H.3.5 Online Information Services - Web-based services*

Introduction

Designers of online learning systems have an enormous variety of software tools and learning resources at their disposal. Unfortunately, the wide variety of software tools available from many different vendors is not able to share learning resources and interoperate with each other. Nowadays, Web-based education research efforts are focused into the standardization of learning metadata schemes, course structures and software interfaces to provide interoperability between applications and learning resources. This would allow both instructors to reuse learning resources and developers to reuse educational software.

The learning technology standardization process is taking the lead role in the research efforts into Web-based education. Standardization is needed for two main reasons: on the one hand, educational resources are defined, structured and presented using different formats; on the other hand, the functional modules that are embedded in a particular learning system cannot be reused by a different one in a straightforward way.

Standardized definitions for course structures are necessary to move courses from one system to another. In a similar way, learning management environments need to understand the course structure to schedule the next student activity. They propose a Web-based run time environment scheme based on the division between the learning contents and the Web-based management system that launches on the browser, delivers through the network and controls them. This allows different learning resources to be managed by heterogeneous management systems.

The objective of this project is to develop advanced program system that helps improve access to Europe's knowledge and educational resources via e-Learning technologies.

The Web application framework to build Learning Management System (LMS) “xDLS” for small educational organizations and components of this LMS are presented. It is demonstrated how frameworks services can be configured to create a complete system. This framework increases reusability and reduces the maintenance requirements of the LMS. It provides a way to develop the Web component “QTI Ready” which allows to edit and play QTI based quizzes and tests. This component already can be used in many of existing VLE (Virtual Learning Environment).

Methodology

The system xDLS combines the latest Internet and database technologies, incorporating a user-friendly and intuitive interface; xDLS raises the industry standard in Web-based testing. Flexible and powerful enterprise architecture of the system permits to effortlessly incorporate xDLS into new or existing training programs. The system is based on Java technology.

The package QTI Ready is the server side J2EE component for operation with tests in IMS QTI format. The package is realized on the Servlet technology with independent platform. Interaction of the package with Web-application is carried out extremely through the protocol atop HTTP protocol. XSL + XML technology is used for generating HTML pages. The relational database is the storage of resources and system information.

Technological Description

In this section two software products, developed by the author, supporting standard IMS Q*TI are described

xDLS Learning Management System

The system xDLS is Web-based enterprise-wide Learning Management System. xDLS software resides on a server, allowing to access it via a Web browser. It also allows teachers to make changes to quizzes readily – from any Web-accessible location – and to make these changes available to your students immediately.

The system xDLS can be used to

- evaluate students with quizzes,
- provide assessments that include text, html, image, flash,
- integrate Web resources.

Authors can create, modify, delete and rearrange questions in assessment with ease. The system resources are complied with open standards such as the IMS Project Question and Test Interoperability (QTI) Specification. The system provides several types of questions (multiple choice, multiple response, matching, numerical and string

fill-in-black). Assessment data can be exported and imported in zipped file according to IMS QTI specification. Image and HTML materials are supported.

It is proposed to organize questions in accordance with the Question and Test Interpretability's ASI (Assessment Section Item) model provided by IMS Global Learning Consortium, Inc. According to this model every assessment is divided into several sections and each section consists of several items. An Item is the smallest object needed for the assessment, which represents generally a question, defined as a combination of interrogatory, rendering, and feedback information. The IMS QTI specification will enable to exchange assessment items and results during the learning process.

The system xDLS combines the latest Internet and database technologies, incorporating a user-friendly and intuitive user interface. xDLS raises the industry standard in Web-based testing. Flexible and powerful enterprise architecture of xDLS permits to effortlessly incorporate xDLS into new systems or existing training programs.

The xDLS key features are listed below:

- Pure HTML based, no Java or Flash needed (unless your content does need it).
- "No frames" solution.
- Multilingual (currently available in Russian, English and German).
- Printview with optimized printing template (without menus or decoration) and optimized font.
- The course editor of xDLS allows create new courses within short time.
- Advanced enrollment system (self enrollment in groups, management of available seats).
- Advanced group management, associating groups to learning areas.
- Support for self organized learning/collaborative work via so-called buddy groups: everybody can create his own group and work together with his peers.
- Testing system with various testing types (Multiple Choice, Single Choice, Fill in blank), fully based on IMS QTI v1.2.
- Score based course system: score can be fed automatically (by online tests).
- Personalized task management.
- Questionary/Survey system, fully based on IMS QTI v1.2.
- Publishing and access management for learning content via the learning resource repository module.
- Easy backup files via ZIP compression.
- 100% Java based, only tomcat and a relational database required (tested with MySQL).
- Unlimited number of learner.
- Unlimited number of courses.

QTI Ready

Package QTI Ready is the server side J2EE component for operation with tests in IMS QTI format. QTI Ready is meant to do only one thing – to edit and play QTI tests. One other notable characteristic of QTI Ready is that it makes relatively few assumptions about a user's machine; playing and authoring tests happens on the server, and the result is rendered as HTML. So no plug-ins is needed, and no major restrictions on hardware or operating system beyond the ability to run a decent browser. It also simplifies authoring tasks by separating the roles of the layout designer and content author. In addition, it gives user the flexibility to view any content using any defined layout as well as to add annotations to the content.

The process of testing and editing is realized in Web-interface that can be easily customized. Export and import IMS QTI tests is realized through loading and uploading ZIP files.

Package QTI Ready consists of two components:

– QTI Ready Player (QRP) provides Web based test and survey delivery capabilities via HACP AICC protocol. Upload a QTI compliant questionnaire into a QTI Ready test server, launch it via AICC URL, take a test, and see your results recorded into your Learning Management System.

– QTI Ready Designer (QRD) provides Web based test and survey authoring capabilities and generates IMS QTI compliant tests and questionnaires with saving them on the server.

Package QTI Ready can be applied to extend the learning management system of remote training and to design test portal.

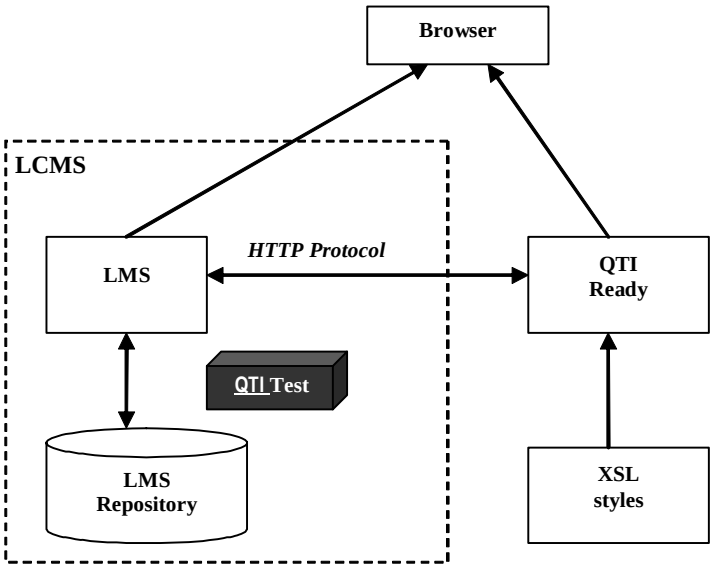


Figure 1. The QTI Ready Component Integration

Interaction of the package with Web-application is carried out extremely through the protocol atop HTTP protocol (fig. 1) that provides:

- possibility of integration with Web-application realized on any Web-technology: Java, Perl, PHP, ASP etc.; (If your server application is not based on Java technology it's enough for you to install Open Source Tomcat server);
- complete independence from system environment: operating systems and DBMS;
- allows centralized data storing.

The VLE can communicate with the QTI Ready authoring tool and player via the AICC HTTP based API (HACP). HACP is an older specification about sending commands from learning content to a VLE via encoding in URIs, which is considered to be pretty unsafe these days. The URIs can be fairly easily grabbed and tampered with on the way. On the plus side, it does not have the cross-domain scripting problems associated with AICC's newer sibling, SCORM.

For those security reasons, the HACP communication between a VLE and the QTI Ready player is limited to selecting and starting a particular test, and signalling when the test is done. The actual score calculation is done on the server side, from a Web form. This makes it less easy for the test results to be tampered with on a user's browser, but, since forms are even easier to grab than URIs, following QTI Ready's advice to use it over a secure HTTPs connection seems wise.

Package is realized on Servlet technology with independent platform and can be easily built in indirect web-application to accessible Servlet container.

Using of XSL+XML technology in generating HTML pages makes user interface customization very simple. There are several types of preformatted user's XSL styles for setting up user interface.

Plug-in supports several languages of user interface simultaneously. The language of the interface for curtain user can be set up while the program is running. Also, it is simple to add a new language.

QTI Ready Player – Web-plug-in, which can play tests stored in repository of Learning Management Systems. Tests are stored in external repository as ZIP archives. This archive contains XML file in IMS QTI format plus resource files used in the test. For Learning Management System format of the test is possible to consider as black box, all functionality of operation with these tests will be carried out by package QTI Ready. QTI Ready loads the test from repository once and returns results via AICC protocol.

There is no limitation for internal structure of tests' repository. It can be the relational database, in Blob fields of which ZIP files will be stored, or simple catalogue of the file system. Access to the tests by plug-in QTI Ready Player is carried out through HTTP protocol or the same way as to local files.

QTI Ready Designer (QRD) provides Web based authoring tool designed specifically to create surveys, tests, and assessments. This component generates IMS QTI compliant questionnaires.

QRD key features are listed below:

- Easy to integrate into any Web based learning management or survey administration system through HTML based API.
- Produces QTI compliant questionnaires.
- Packaged as J2EE Web Application.
- Easily customizable XSLT based user interface.

To set up QTI Ready Designer it will be necessary to store repository of the tests analogue to repository for QRP. Tests may be stored in the same form of zip archives. In addition to ZIP files it is required to store the set of attributes for each test. For editing test QRD loads it to the temporary local storage. Test will be uploaded to the repository of Web-application obviously when command about export of the test is given from QRD. Also in Web-application it's possible to organize the implicit load of test from temporary storage. If the user quits editing the test, without having made export of the test, then by the time of the next session the test will be given in that state, in which it was left.

Generation of HTML is executed on the server using XSL+XML technology. For setting up user interface it is recommended to customize user's styles where descriptions of general units of external interface is stored.

Conclusion

The system xDLS is used in more than 10 organizations in Russia (Moscow, Siberia and so on) and Ukraine. The package QTI Ready is integrated with xDLS. Package is tested on following Servlet containers: Tomcat 3.3, 4.0, 4.1,5.0, JRun, Orion Web Server.

The system easily was adapted for requirements of users. It allows to execute integration with existing systems or to create on its basis a new system in the shortest terms.

The current IMS QTI standard version does not provide a possibility to adaptive testing. Author would like to continue experiments with the all offered assessment types by IMS QTI and to investigate also more complex sequencing types to offer a more precise and faster marking process, as well as a timely feedback. This enables to provide the learner with a wide range of assessment possibilities, to permit a dynamic learning process based on an accurate evaluation of knowledge level.

Bibliography

IEEE PAPI. IEEE P1484.2.5/D8, 2002. Draft standard for learning technology. Public and private information (papi) for learners (papi learner). Available at: http://jtc1sc36.org/doc/36N0179.pdf

IMS Global Learning Consortium Inc. IMS Question and Test Interoperability ASI Information Model,2004.V1.2.Available at: http://www.imsglobal.org/question/qtiv1p2/imsqti_asi_infov1p2.html.

IMS Global Learning Consortium Inc. IMS Question and Test Interoperability Results reporting, 2004.V1.2. Available at: http://www.imsglobal.org/question/qtiv1p2/imsqti_res_bestv1p2.html .

IMS Global Learning Consortium Inc. IMS Question and Test Interoperability Specification, 2004,Version 1.2. Available at: http://www.imsglobal.org/question/index.cfm.

LTSC Learning Technology Standards Committee, Learning Object metadata standard, 2004. Available at: http://ltsc.ieee.org.

[Cheryl, 2001] G. Cheryl. XML: New Formula for e-Learning. 2001.

Authors' Information

Andrey Belonogov - Perm State University, Post graduate student of the Department of Computer Science; PSU, 15, Bukirev St., Perm, 614990, Russia; e-mail: abelon@perm.ru

I&R SYSTEMS ON THE INTERNET/INTRANET CITES AS THE TOOL FOR DISTANCE LEARNING

Andrii Donchenko

***Abstract**: This article considers the Internet/Intranet information systems as the tool for distance learning. Author considers the model of the 3-tier WEB based information system, the idea of the language for implementing and customized solution, which includes the original language and processor for fast prototyping and implementing small and middle sized Internet/Intranet information systems.*

***Keywords:** I&R Systems, client-server architecture, distance learning*

***ACM Classification Keywords:** I&R Systems*

Introduction

The distance learning takes an important place in the world. The useful tools for implementing such systems are small and middle Internet/Intranet based I&R systems.

This article considers the problems of designing, implementing and classifying these I&R systems. Author proposes the conceptual model of architecture of the system and describes the original tools those were created for prototyping, implementing and maintaining small and middle WEB based I&R systems.

Problems and Goals

Now there is an expansion of the tasks those are related to the designing, implementing, developing, supporting and maintaining Internet/Intranet based I&R systems. These tasks are related to the developing and supporting access to various databases using standard protocols for supporting the data transport layer (search engines, I&R systems, etc.). There are large amount of systems those solve the task of integrating I&R system with the database management systems and allow prototyping of the resulting documents. The sample of such systems is Oracle WWW. Unfortunately lot of these systems is oriented for one particular operation system. The sample is Microsoft Peer WEB server. Other systems require specific software on the server and client sides. The sample is Lotus Notes.

The important problems and aims of the designing, maintaining and supporting Internet/Intranet based small and middle I&R systems are following.

1. Choosing of the method of the representation of the information (the information architecture of the system)
2. The user interface of the system
3. The tools those allow prototyping and maintaining the user interface and whole system using the model of information representation, which has been selected.

The conceptual model and tools for prototyping, supporting and maintaining middle and small Internet/Intranet based I&R systems are described below.

Information Architecture

The information architecture includes two following key aspects.

1. The functional interaction with end user.
2. The methods of categorizing of information (the main method and the alternative methods

From functional prospective interaction user with I&R system can be considered as searching information; constructing queries and browsing the results.

Generally there are two main types of searching, which are the full-text search and the categorical search. The sequential browsing of the articles (insight, learning) is a degenerate search, which implies switching to the next document in the document's list by default.

According to the approach [Wurman Richard Soul, 1996] it is possible to consider five following methods of structuring information inside I&R system (LATCH-model) using following criteria.

1. L – Location (by the geographical location)
2. A – Alphabetical (by the lexicographical order)
3. T – Time (by the time of the events)
4. C- Category (by the category of the object)
5. H – Hierarchy (By the hierarchy of objects. For example moving from little to largest, depending on changing thickness, color, etc.)

Taking a decision as to the main and alternate methods of structuring of information is a first step in designing information architecture of the I&R system. These methods should take into account the features of human work during obtaining the creative solutions. They includes iterative cycle for generation of the hypothesis, verification, decomposing the problem for components, various methods for solving each particular task, spontaneous switching from one method of solving to another method. The history of the Internet I&R systems proves this thesis. Some those systems have been started like a specialized list of Internet resources. For example Yahoo has been started like a list of useful resources on the user's homepages. Other systems were started like a full text searching systems (Lycos, AltaVista). Now all providers of such services allow both searching methods as the complimentary methods.

User Interface of I&R System.

Now only basic set of interface elements is defined. This set includes only "atoms" (edit control, text box, and tree or list box) for constructing graphical user interface. The items for designing and implementing the high level paradigms of the interfaces like encyclopedia, textbook, I&R system only become to be formalized and standardized. From one hand these metaphors are based upon various paper editions those gave us such metaphors as the contents, the indexes, the tables of the illustrations so on. From other hand it is based upon new metaphors those were given us by the IT technologies. The best sample is an evolution of the hypertext systems. The programmer must fill the gap between these metaphors and basic elements by mean constructing various shields those allow experts in the particular areas to organize the information.

Choosing of the information architecture is an important problem. But beside of this problem there are needs of supporting I&R products, keeping these products in the actual state, the relations between developers of the system and the system users. For example it is reasonable to inform the user about known bugs, available patches and extensions of the system.

Lack of standards for information metaphors complicates joining information arrays and the interactions between different I&R systems. New interface elements came during solving particular tasks from the real life. The examples are the meta-tags in HTML page, the formats of the hypertext I&R systems like WinHelp so on. But generally the sphere of actions is still open in this area.

Conceptual Model of the I&R System

The analyzing of existing I&R systems gives the base for defining the conceptual model of the I&R system. The following statements describe the requirements for this conceptual model.

- The information inside system must be represented by the linear set of the information articles
- Each article must be included in one or more index sequences those are the links and represent the hierarchies.
- Each article id and the path in the hierarchy must represent the addressing method for particular article.
- The UI must be determined using frames.
- Navigation must be depend on the method of the article's addressing (the frame state or the article body)
- Full text search must allow the following

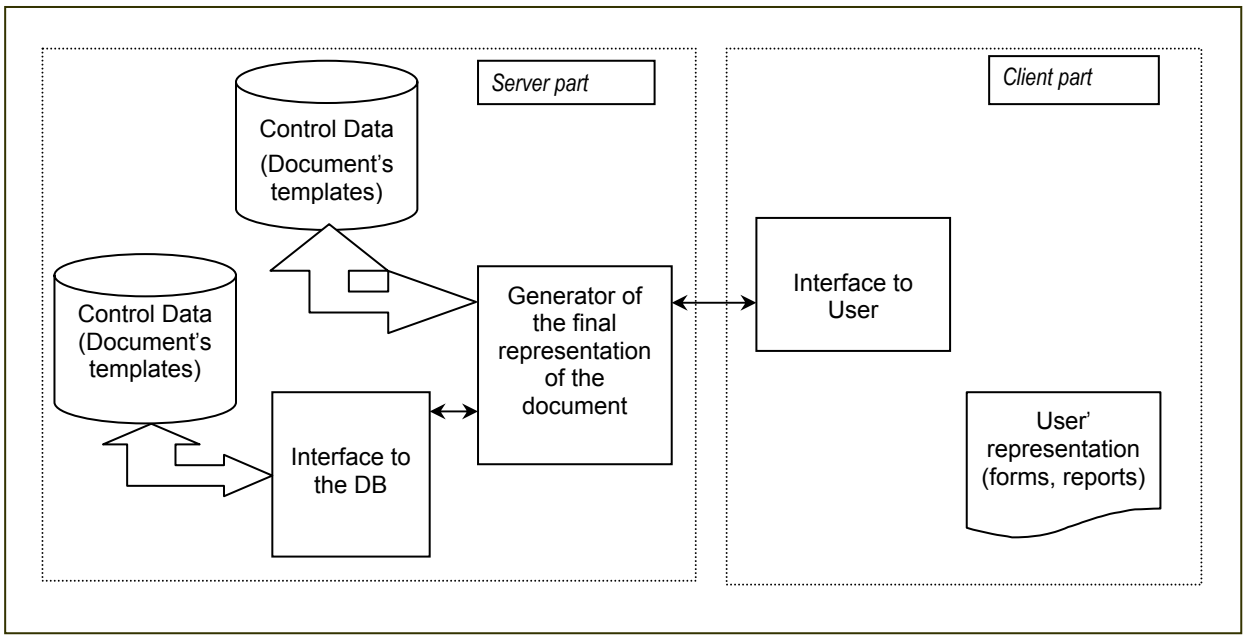
- determining the query conditions using the regular expressions;
- stating the limitation of the searching range
- browsing list of articles those were found.

- The possibilities for providing the feedback to the authors and to automatic notification about the system's sate changes must be allowed for the registered users.

Tools for Prototyping and Maintaining I&R Systems

The tools for prototyping, implementing and maintaining of the small and middle I&R Internet/Internet based systems have been developed taking into account the conceptual model of I&R system, which is described in the chapter above. The 3 tier client-server based solution has been used for the implementation of the software package, which allows prototyping and supporting of the hypertext documents using the tabular text data. The transport layer protocol is the HTTP protocol. The server part of the system also uses the CGI convention for the HTTP servers. The client side has no restrictions. Any standard WWW-client (Microsoft Internet Explorer, Netscape Navigator, Lynx so on), which supports HTML 2.0 standard for the hypertext document representations can be used.

The functional interactions between system components are represented on the picture 1.



Picture 1

This system has a 3-tier architecture. The first layer is an interface with the database. The second layer is a business rules of the data mussing. The third layer is and client application's interface.

On each step of the processing the generator of the final representations of the documents constructs the resulting document. This document is going to be delivered to the user using the standard WWW server and HTTP protocol. Then the client part allows browsing and printing of the resulting document. Usually this functionality is allowed by the standard WWW client.

System data structures.

All main data structures are represented by the text files.

The input data for the processor of the document templates are the following.

1. The document template.
2. The tabular data those are used during constructing the resulting document.

The program on the document template description language delivers the necessary for constructing resulting document information to the documents template processor. The tabular data can be also inverted and indexed for optimizing an access and allowing filtering of the corteges using the key.

The result of executing is a text document that corresponds to the requirements of [T.Berners-Lee, D. Connoly, 1995]. This document is constructed basing upon the tabular data and can be browsed using any standard WWW browser, which supports the HTML 2.0 or higher as an input language.

Work languages of the system

The system supports two internal programming languages for operating benefits. These languages are

1. The language of description of the document template
2. The language of the tabular data representation.

The language of description of the document template

The main goal of the document templates description language is allowing parameterization of the final document representation that is based upon the tabular data those are kept on the server.

The current state of the execution system is determined by the triplet <the current operator of the language; the current data table; the number of the record in the current data table>.

The macro command is the operator of the language of description of the document templates, which invokes changing of the parameters of the execution system (new table is opened; new part of the output text has been generated, etc.)

The macro values are the objects of the incoming program. These objects are connected to the particular input data of the particular current table.

The macro function is the operator of the language, which does not change the state of the execution system. It generates only the data those do not directly based on the data from the current table.

Using of the macro is using of the macro function, macro value or the macro function in the text of the program.

The language of description of the documents templates is implemented as the language of using of the macro. The program on this language is a document template. The document is represented by the superposition of the constant parts and using of the macros. The constant part is unchanged and does not depend on the input data. The macro represents the variable part. The representation of this part may be changed depending on the content of the input tables and input parameters of the execution system.

There are three following functional groups of the operators, macro variables and functions of the language.

1. The operators those conduct the process of the program executing and include the operators of branching and of cycling.
2. The operators those support the database layer interaction. These facilities allow to set the current table, set the current record, extract the particular domain value.
3. The auxiliary operators and macro functions those support the interaction with operation environment like access to the CGI interface parameters and to the system environment variables. Also these operators and functions include the data conversion functions.

The language of the tabular data representation.

This language allows the tabular representation of the source data. This representation has an internal format, which is specific for the particular system. Each table is represented by one file, which contains all corteges of the table. The attribute is the number of the domain in the cortege.

Conclusion

This article

- Introduces the conceptual model of information representation, which covers a big number of the encyclopedia-like systems those provide the standardized toolset of searching and browsing facilities. Embedded facilities for providing the feedback for the end user and announcing of the user simplify support of various life stages of the I&R system. The set of information metaphors has been considered. This set includes the basic high-level definitions of the I&R systems' interface items. These items allow description, fast prototyping and implementation of these systems.

- Considers the key features of the software package that has been developed and implemented. The package implement the I&R system's model, which has been discussed above. The distinguishing features of these tools are below.
    1. The package provides flexible and simple mechanism for support interaction between HTTP server and data, which have a tabular presentation.
    2. The package allows a cross-platform data transfer. The formats of templates of the text documents and the formats of the tabular data do not depend on the choice of the operation system. It is also possible to transfer the processor of the text document templates between various operation systems.
    3. The set of the resulting documents can be easily extended by mean ether changing the tabular data or extending the set of document templates.

Some intermediate project were already implemented using the conceptual model of the I&R system and the software package those have been discussed above. These projects were published as small and middle WWW sites in Internet and Intranet networks.

Bibliography

[T.Berners-Lee, D. Connoly, 1995] T.Berners-Lee MIT/W3C, D. Connoly. RFC 1866, Hypertext Markup Language – 2.0, November 1995, Network Working Group.

[Wurman Richard Soul, 1996]Wurman Richard Soul. Information Architects., © 1996, Graphis Press Corp., ISBN 3-85709-458-3.

[Leontyev A. N., 1981] Leyontyev A. N. The problems of Evolution of Mentality. 3-4 addition. M., 1981, Moscow State University Publisher.

Author's Information

Andrii Donchenko – Bonus Technology, Ph.D., Project Manager/Team Leader; 86D Bozhenko Str., Kyiv, Ukraine, e-mail: andriy.donchenko@gmail.com

USING ELEMENTS OF SEMANTIC PARSING IN E-LEARNING ENVIRONMENTS

Andrii Striuk

***Abstract**: Possibilities for using semantic parsing to estimate the correspondence of text materials to teaching aims, correspondence of test task to theoretical materials and other problems arising during the distance course designing and educational process itself in e-learning environments.*

***Keywords**: semantic parsing, e-learning environments, distance courses, teaching aims.*

***ACM Classification Keywords**: K.3.1 Computer and Education: Computer uses in education*

Introduction

A lot of researches are devoted to the possibilities of applying artificial intelligence tools in educational systems. The analysis shows that much attention is paid to semantic analysis. In particular there are many tools that use semantic analysis in automated assessment of student's complete answer to some question. Investigations conducted by the Modeling and Software Department of Kryvyi Rih Technical University show that possibilities of semantic analysis application in educational environments are considerably wider. In particular at the stage of distance courses design using semantic parsing it is possible to estimate the correspondence of text materials to teaching aims, correspondence of test tasks to theoretical materials used. During the process of training the elements of semantic parsing can be used for estimation of richness of messages content. The present article is devoted to the realization of these and other possibilities.

Possibilities for Semantic Parsing Application in E-learning Environments

The tools of key information extraction from the text and analysis of its semantic context lay in the foundation of the methods proposed by us. A lot of investigations (among them there are many works by Peter Turney [Turney, 1997; Turney, 2000]) have been devoted to key information extractions tools from the text, compiling of short summaries and annotations on the basis of this information. The elements of such investigations are widely implemented in searching web-systems, electronic libraries catalogue tolls, but almost no investigations have been conducted for implementations of similar tools in the distance learning systems. At the same time information exchange in the interactive modules of educational environments takes place using text information and its automatic analysis could considerably increase the efficiency of educational process.

Several tasks that can by solved using elements of semantic parsing have been distinguished in our investigations. Here are some of them.

Firstly, the distinguishing of the key information is in itself an important tool and can be widely used in the distance learning organization. The present method enables, for example, to automate the process of keywords distinguishing in educational texts and glossary building, the creating of concise context of each chapter and of the whole educational course. And the most important is that the method of key information distinguishing in the educational text is a basis for all other directions of our investigations and before considering the other possibilities of semantic parsing application we concentrated our efforts on the development and perfection of this very module.

Secondly, key information distinguishing in the educational text allows creating the analyzer of educational material correspondence to the set educational aims. This will allow the developer of a distance course, firstly, to pay attention to the aims to be achieved in the distance course. Strictly determined aims enable the correspondence of educational materials to the set aims to be traced automatically, to indicate to the completeness of one or another question disclosure and to conduct search for the necessary information in the proposed database even in automatic mode. A collection of text messages, files or World Wide Web resources can be such a database. Thus the semantic parsing becomes an indispensable assistant during designing and creating the distance course in the educational environment.

Thirdly, the conditions for correspondence analysis of test tasks to the aims of education and to the contents of theoretical materials are being created. There are some well-known facts when much attention has been paid to insignificant aspects considered in the theoretical material during composing the test tasks. This, of course, decreased the efficiency of testing and did not give the objective estimation of students' knowledge. On the basis of the key information distinguishing it is possible to determine the correspondence of test tasks to the main contents of educational material and to show the developer of knowledge control system the possibility of error. Insufficient number of questions concerning the subject matter, having been distinguished by semantic parsing as a key one, can become another problem while compiling test tasks. The module of test tasks analysis will be able to indicate the non-correspondence between theoretical material and test tasks, and, thus, to contribute to the increase of testing efficiency.

Fourthly, there is an opportunity of semantic parsing application to estimate the correspondence of text messages that pass through interactive modules, to the subject of the current lesson. The important feature of distance courses is the intensive application of communication tools, such as e-mail, forum, chat to organize the communication between students and a teacher. Of course, such communication is of educational character and is used to estimate the quality of learning. That's why certain grades for activity in communication have been provided in the students' assessment system. But for such communication not only availability of text material but also its richness of contents and correspondence of the lesson subjects, that is being conducted at the moment, or the course subject as a whole, are important. The teacher has to estimate the richness of messages content by himself, that's why he/she faces the necessity to work through a great number of text messages daily. This is especially important during the intensive work with a great number of students. In this case the semantic parser allows automate the process of text messages analysis and determining their correspondence to the given theme

or the course subject matter in general. This also creates the possibility to implement the automated knowledge control not only using testing but also using students' activity analysis in forum, e-mail or chat considering the richness of their messages content.

Semantic Parser Implementation and its Integration into E-learning Environments

As it has already been pointed, the procedure of key information distinguishing from the text messages, and in particular educational texts, underlies all the presented methods. Implementation of this procedure, proposed by us, is rather simple. First of all keywords are distinguished from the text on the basis of statistical data. Keywords determinations are performed on the basis of Zipf's laws. The implementation of them has become rather popular, for example, on modern search engines [Breslau, 1998]. Further, the context, where the keywords are used, is analyzed; a so called semantic cut for each keyword is prepared. A table from words that go with keywords is formed, semantic links between keywords and words-satellites are established. At this stage the analysis of synonymous and homonymous variations with application of corresponding dictionaries is provided. The generation of tables for each keyword is concluded with a cluster analysis, during which links between certain terms on the basis of their context are determined.

These tables are the foundation for realization of methods described above. Key notions and terms will be in the knots of the table. Their definitions can be obtained from the analysis of their context and, thus, a glossary for each educational chapter and for the whole course can be created. While determining the correspondence of educational materials to the course aims, the search of distinguished keywords is done in the formulations of aims, their context is analyzed, the semantic cut is built for each of them and is compared with the correspondent semantic cut. Coefficient of coincidence allows proving correspondence or non-correspondence of educational materials or separate chapters to the course aims. Similarly the correspondence of test tasks to theoretical materials is established and the analysis of content richness of text messages in chat, forum and e-mail is conducted.

After detailed elaboration and implementation of the semantic parsing methods for educational texts, our research was directed to the development of actions to integrate the semantic parser into the existing educational environments and developments of dialog modes between analyzer and user.

Conclusion

The conducted investigations enable us to build software using semantic parsing elements for effective solving of numerous problems arising during distance courses design and educational process organization itself in distance learning environments. We hope that the conducted investigations will be the basis of a more complex project, i.e. the elaboration of expert-system which acts as a methodologist while designing, filling with necessary material and implementation of distance courses.

Bibliography

[Turney, 1997] P.Turney. Extraction of Keyphrases from Text: Evaluation of Four Algorithms. National Research Council Canada, Institute for Information Technology, 1997.

[Turney, 2000] P.Turney. Learning Algorithms for Keyphrase Extraction. National Research Council Canada, Institute for Information Technology, 2000.

[Breslau, 1998] Lee Breslau, Pei Chao, Li Fan, G. Phillips, S. Shenker. On the implications of Zipf's law for Web caching. Proc. 3d Int. WWW Caching Workshop, Manchester, UK, June 1998.

Author's Information

Andrii Striuk – Kryvyi Rih Technical University; XXII Partz'yizd st., 11, room 219, Kryvyi Rih, 50027, Ukraine; e-mail: andrey_stryuk@mail.ru

TABLE OF CONTENTS OF VOLUME 1, NUMBER 3