12th IMEKO TC1 & TC7 Joint Symposium on
Man Science & Measurement
September, 3 – 5, 2008, Annecy, France

A METHOD OF DISTRIBUTED TEACHING SYSTEM REALIZATION BASED ON FUZZY MEASUREMENT OF USER'S FEATURES

V.N. Yakimov, G.N. Dyakonov, E.G. Udaltsova, V.I .Batishchev, A.V. Mashkov

Samara State Technical University, Samara, Russian Federation,
E-mail: yvnr@hotmail.com, hernd@yandex.ru, caged@yandex.ru, mav@it.samgtu.ru

Abstract: The purpose of the paper is to apply the main ideas and methods of soft (fuzzy) measurements to the problem of network learning. Teaching system is considered to be really distributed. As secondary effect of its content (in some point of view it may be called primary) we suppose a method to increase an effectiveness of the WWW.

Keywords: adaptive learning system, capacity, user's features profile, Semantic Web.

PREAMBLE

Education is the most important part of human life. Our skillfulness, our ability to support high social status, our professional success and many other aspects of our being depend upon our education.

That is why every civilization takes care about effective ways, forms, methods of teaching both in common and special (technical, musical, religious a. o.) education.

As it is commonly known, the best teaching is “eye – to – eye” method. Its sense is that each lesson consists of colloquial (more or less with dialogue using) lection and a kind of practical exercise and/or test.

When both of teacher and pupil are human beings it is easy to implement the scheme above. One may think about traditional English educational system with tutor and one or more (but not too many) pupils or students. We may construct a model of this form of teaching.

The teacher (tutor) describes a portion of a new didactical material. He uses different levels of explication, different temps of narration in view of a student’s understanding. To detect it a lot of skills may be used, but all of them are based upon the feedback signals both of lingual and non-lingual nature. The last may be: face mimics, shoulder rising, hand gestures.

What about lingual feedback signals they are no more than questions like “Is it clear?” and questions of current control.

But the modern tendencies challenge the new educational technologies, as such as automated teaching systems, or ATS for shot. The history of these systems begins from the first “Yes – No” quizzes used devices with punched cards. When computers were introduced to create the learning systems the new approach was inspired by teachers and scientists.

Further development of information technologies leads the ATS – society to idea of creation two types of the automated learning systems: adaptive and distributed ones.

Primordial adaptive learning systems provide determination of user’s features only on the installation or at last on the initial dialog stage.

What about distributed automated teaching systems they are in use a long time, but their renaissance starts with WWW access becomes popular. But up to now they do not use all Internet and

Measurement is increasingly seen as a key technology for management and control by the wide class of systems. Now engineers and technologists are skillful enough to measure almost complete set of physical parameters they deal with.

But a number of non-physical parameters exist and some of them ought to be measured or estimated. The possibility of measurement of the non-measurable parameters is based upon a set of different scales. They are nominal, ordinal, interval and ratio scales [1,2]. When we deal with human mental features we rather should say about measurement like about classification. Some variances may be observed when the explorer deals with clusterization. It occurs when a priory classification system not established, and its generation runs “in real-time mode”. This case is rather common in a lot of situations.

Among various applications of the human features measurement WWW applications are considered. First is an application, which makes possible an adaptive, flexible distributed teaching system creating. It corresponds with the new tendency: Semantic Web has to connect applications, not people. It means that intellectual work of the WWW increases. But it is impossible when user is somewhat unknown for Network.

A possible solution may be considered when a user is introduced to the Semantic Web or to similar system (e.g., to the adaptive automated teaching system, which becomes to be intellectual on and has a feature to be comfortable to deal with for pupil.

And the second profit we can get is that WWW would send to the users more useful (or more pertinent) information. It leads to increasing of WWW efficacy, reduction of useless downloads, and, at last, increasing of effective network capacity.

1. THE METHODOLOGY OF EXPLORATOIN

The used method consists of metaphoric approach to the problem of distributed teaching system in the terms of information storage and retrieval system.

From this point of view two indices, among other, are the most useful to describe information retrieval system. They are recall ratio and precision ratio. The first is determined as output volume (number of URLs) divided by the number of complete list of relevant URLs across the WWW.

The second is the number of relevant URLs in the output divided by output volume (number of URLs). The recall ratio determines a fullness of the WWW using. But the precision ratio may be used as index of efficacy of the system.

The matter is that exist a chain of information waste. The reason for user's appeal to WWW is his/her information need. First of all user tries to express it in the terms of natural language (Russian, French, English a. o.). But the mental intention and its lingual expression are not equivalent each time.

The next element of the chain is information request. It is an attempt to compress the lingual representation of user's need to acceptable size. The information request is a kind of restricted formulation of information need in pseudo-natural language.

At last, the retrieval machine as a part of WWW tries to compress the request (or expand it to execute generic and specific search) and elaborates in such manner a retrieval instruction to interoperate with its inner databases.

There are some differences in terms, used in traditional WWW approach and traditional information storage and retrieval systems, but equivalency in concepts takes place. This is illustrated in the Table 1.

Table 1. Terms correspondence between two traditional approach: WWW and ISRS.

Terms of Internet and WWW practice	Terms of information storage and retrieval system practice
—	Information needs
All URLs set	Documents (books, journals, magazines, newspapers, drawings a. o.)
Set of descriptors to be included in database dictionary	Secondary documents, or information retrieval apparatus
Request for search	Information request
Request to databases for search in internal language of the searching machine	Information prescription
GET command to send a list of URLs.	Output criterion.

Every step of this descending ladder from human intellect down to formal prescription for computer is connected to losses of semantics and expressibility of the requested field. The chain "information needs – information request – retrieval instruction" leads us farther from pertinence to simple technical coincidence of two retrieval attributes.

Scheme (see Fig. 1) demonstrates the main functional features of information storage and retrieval system. It illustrates a case when a user with some information needs makes an inquiry to the system.

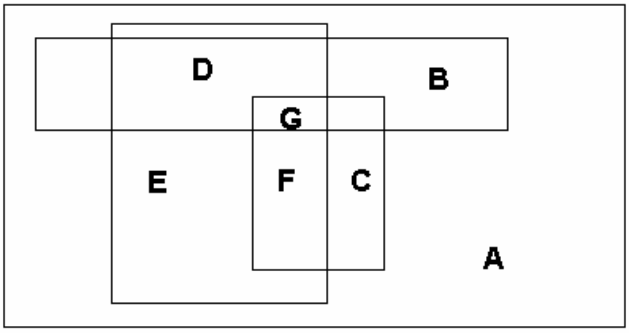

Fig. 1. Example of ISR - system functioning.

The legend to Fig. 1 is represented in the Table 2.

Table 2. The main functional features of information storage and retrieval system.

Area sign	Area content
A	the Knowledge Universum
B	user's Interests Area
C	pertinent sources found
D	relevant Documents
E	common documentary storage
F	a part of documents found, which are not pertinent or suitable for user
G	relevant and pertinent documents, found by request and useful for the user

2. RESULTS

Some basic set of learning ability features is constituted. It may be considered as a generalized vector with components (the most suitable form of information representation, ability to customize new knowledge, the level of the knowledge adoption). The first component describes those forms of information representation, which are convenient to the definite user ant may be listed as:

1. Formal theory. This form of information representation is extremely compact. But high level of intelligence and an ability for logical (and/or other sort of) thinking are necessary to apply it as a form of didactics.

2. Pictures with formal content (e.g., graphics, diagrams, plots a. o.). This feature correlates with the last one, but it is more comfortable for picture-minded high-level intelligence user.
3. Pictures with non-formal content. This form is not rigorous; it may be understandable for lower intelligence user.
4. Narration of high-level content. We set this form in the fourth place, but it may be useful for rather advanced user, who is informed in the interest area.
5. Popular narration (may be, with pictures). A form, almost of comics level, but it is useful for primary learning of some topic.

One note ought to be considered. We can say about three levels of verbal thinking.

Initial, or lower verbal level, when there are neither formal nor mathematical models to be discussed.

Higher verbal level, when a person's knowledge is complete enough to consider these models.

The highest verbal level, when a person is skillful to discuss problem under interest as like these models seem absolutely obvious. The problem, not models is discussed by users.

The second component may characterize some personal abilities of a user from the point of view of information adaptation:

1. Quick-minded pupil or advanced user. They are usually correlated with formally – minded persons, but it is not obligatory correspondence.
2. Middle- (or common-) minded pupil/user. We can suppose a lot of people belong to this category. When adaptive distributed teaching system is under construction we may deal with it.
3. Slow-minded pupil or naïve user. As alternative case we may consider this category as beginner user or junior user category.

At last, the third component describes a level of user's qualification in the field of interests. We suggest the next ladder of levels:

1. Ability to repeat a topic. It means that pupil may repeat the main terms, tell about basic concepts of the topic under consideration.
2. Ability to answer of straight questions about the topic content.
3. Ability to solve a standard task.
4. Ability to solve a non-standard task.
5. Ability to formulate a non-standard task.
6. Ability to formulate any new, non-standard, creative task and solve it. This is the highest level of human intelligence.

This set is used as a system of generalized coordinates, and the determined in such way space is supposed to represent the set of users' didactic feature as points in the space above.

It is obvious, that any metrics in such terms has to be fuzzy one. So, it is impossible to involve an etalon of any feature above. We can say about only approximately defined significations of them.

Because of impossibility of multi-dimensional space graphical representation we deal with any palliative form of this.

An obfuscate character of the human's psychic cannot be described in any more or less rigorous or precise manner. We have to deal with approximate estimation.

Because of complexity of the problem of human modeling our attempt ought to be seen as primordial. Results are by Fig. 2 illustrated.

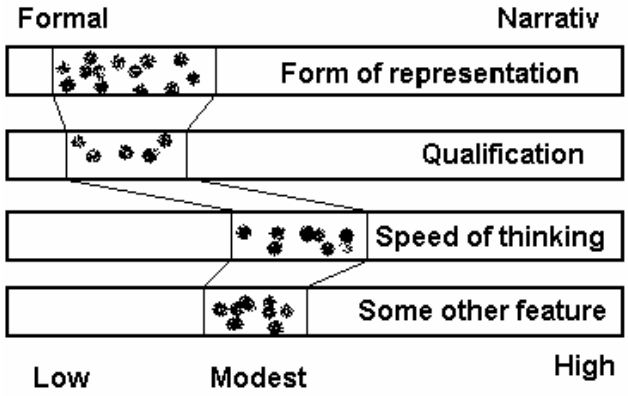

Fig. 2. Example of fuzzy multi-dimensional space and a user's model like learned person.

This picture illustrates a sort of evolved multi-dimensional space. Each dimension is displayed as rectangle area. Spotted sub areas illustrate a set of features of any concrete user.

These coordinates do not seem orthogonal. It means that a correlation may exist between each pair of those coordinates. But the construction of really orthogonal space to characterize a human facility is rather difficult task.

A number of scales may be applied to measure human being features. Among them the nominal and ordinal ones seem to play main role in the matter of user's features characterization.

A method to increase equivalent capacity of the WWW is offered [3]. It is supported with two means, rather new but not complicated in realization. All we need is, firstly, a support of users' profiles repository. This mean is practically in use, because different profiles are stored across WWW. Now they used for users' account, or for eTrading, but a matter is that we suggest supporting the users' profiles during system lifetime (or, pitifully, user's).

The matter is that they ought to be added with information about user as a student of a virtual school. Then the system may be configured in the manner of individual tutor. On the user's request for any didactic resource the most appropriate one may be send to the user. The appropriation here means as formal aspect, as contextual one. In such way the pupil may find himself in more comfortable condition.

The second is three or four obligatory pairs of tags to mark up WWW documents. We mean those of common set of WWW documents, which may be a part of o distributed teaching system. Data is generally contended in HTML files, and it does not cost a lot of effort to add a pair of tags

or include corresponding information into META section of HEAD.

But using of XML seems more perspective way. Anyway a tag pair like <INFORMATION_AQUISITION_FORM> - </ INFORMATION_AQUISITION_FORM > or <INFORMATION_ADAPTATION_LEVEL> - </ INFORMATION_ADAPTATION_LEVEL > may be useful.

Let us consider a sub network for some special functions, as learning or differential informing. It is obviously tied with a person features and abilities. A scenario of such system may be considered.

At the primary log-on a user executes a registration procedure. He is asked to answer the questions about his field of interest, the level of training, the most convenient form of information representation a. o.

The system executes his/her registration and forms corresponding primary profile.

After every logging-on the user is invited to input his/her request.

The user (as skillfully as he is possible) conveys his information need in the form of information request.

The system transforms the information request to a retrieval instruction, taking into consideration rules and restrictions of its information retrieval language.

System finds out the user' profile from one of the profile repository in WWW sub-network.

System executes the search in its intrinsic databases and forms an output.

Data from user's personal profile and from information request is used as relevancy criterion.

The system sends output data to the user, who analyses it and sends the results of analyses back to the system as feedback message.

On receiving of feedback message system executes necessary correction (adaptation) of the user's personal profile.

1. When a user fulfills his initial registration procedure, he may be asked some questions to form his primary profile. The questions seem to be like these:
2. What is your educational level?
3. What is your profession? What is your special interest area in the profession?
4. What problem are you working on?
5. What is your job position?
6. What kind of information representation is the most suitable for you?

When some current didactic dialog starts, or when it terminates, adaptive teaching system may renew the user's current profile.

Tests and measurements may be standardized instruments, which are used to evaluate the level of, individual features, or speed of information processing. They may be useful when individual profile is created or corrected.

The existence of universal and specified ontologies ought to be provided [4]. The universally developed ontology may be used as standard when quantity of semantic information is measured for any text [5]. As a matter of fact, information quantity ought to be measured not at all in irrelative terms. The picture below illustrates a situation. One can see that both of source and receiver have their own thesauri.

There are a lot of different approaches to perform a teacher — pupil interaction. One of them (may be the most popular) is communicational approach. But an important notice ought to be considered. It is linked with intellectual features of human being.

The classical C.E Shannon's model of communication system does not include anything than source, receiver and a channel with noise addition. It does not consider any ability of receiver terminal agents to be informed about "something very fresh and cool and the newest stuff" from the point of source's view.

The fact is that we ought to consider a kind of information storage, where at last discourse data may be stored. A similar model may be seen in the Fig. 3.

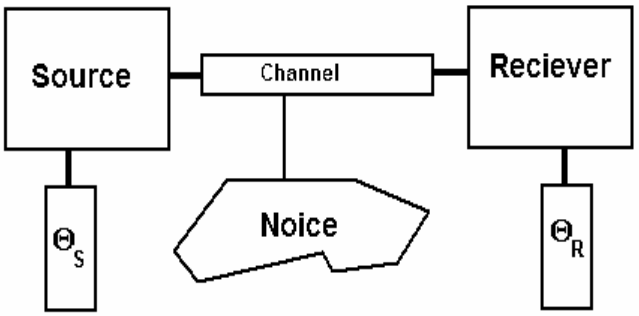

Fig. 3. Yu. Shreyder's extension of Shannon model of communicative process.

Here we designated the source thesaurus as Θ_S and receiver thesaurus as Θ_R.

We may consider thesaurus as some quantity of a priori information. Any communicative process may be performed and considered as effective only if:

1. the thesauri have one or more common areas;
2. no one of them is a proper part of another.

Any portion of distributed knowledge system may be considered as a part of NFL-continuum, where NLF is abbreviation for "Natural and Formal Languages (documents)" [6]. It means we can use any form of teaching information representative form. This way seems to be a way to solve a risk of "traditional schooling … anachronism" [7].

To be used as didactic text, it has to be structured in some special manner:

Rather small portion of new didactic information is included into such text.

This portion ought to be extruded from some part of ontology-correct knowledge model.

This portion ought to be correlated with previous ones in terms of input metainformation interface [3].

This portion includes the necessary data to form the output interface.

This portion contains markups to characterize some extralingual tasks and features of this piece of didactic information.

Ontology representation may require a special formal language. It seems to be similar to OWL but with more

flexible and universal (at any subject area, at last) apparatus of paradigmatic relations [8].

In another word we deal not only with information level of knowledge structure, but with metainformation one. We must know not only about a thing. We must cognize, that we do know about it, and we must systemize our individual knowledge system according to that metainformation level of some universal system. It seems we may tell about metacognition yet [9].

When these conditions are taken in the system, it may be used like adaptive one.

3. CONCLUSION

A multi-purpose tool is suggested. Its essence is generation and maintaining through entire World-Wide Web a special type of user's registration document. Such permanent storage makes possible to monitor any individual user's request historization, to explicate different profiles of that user [10].

Due to restriction of non-pertinent information (or increase a ratio of pertinent information) decrease of non-effective dialog takes place.

A scheme of basic components of semantic-oriented retrieval machine is illustrated in the Fig. 4.

Worth to note that development of key-word search technique and using of formal request languages do not eliminate neither high ratio of information noise nor of output.

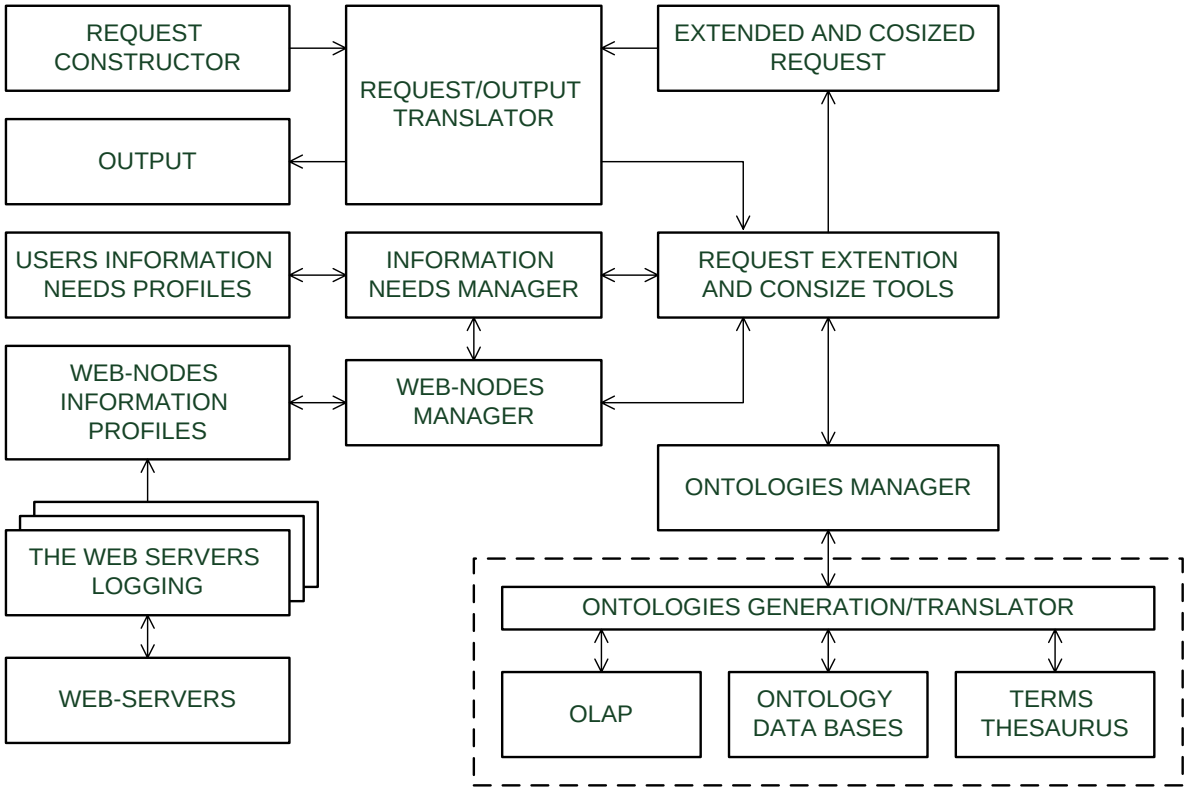

Fig. 4. Basic components of semantic – oriented searching machine and their interconnection.

It occurs because of losses of information when requests are translated from natural language into language of information search. The using of ontology allows more exact interpretation the semantics of request terms. It allows to add terms, which are connected with request terms as generic ones, synonymic, part-entire, associative or other relationships. It may minimize a possible output incompleteness.

The filtering and concretization of requests are used in the ISR systems.

Ontology Generator/Translator permits to map the conceptual representation of ontology onto goal languages of implementation. It provides an access to ontology from remote application.

The OLAP performs analyses to extract knowledge from network databases. Different cross-sections of data cube may be requested.

It leads to rise of effective capacity of WWW channels. The second profit is that pupil has in his service a kind of computerized tutor, which may be adapted to user's personal facilities and features.

A didactic portion may be formatted like that (an example of definition of limit).

<DIDACTIC>
<TOPIC_NAME>
Limit Definition
</TOPIC_NAME>
<NEEDED_TERMS>
Numeric series, delta – epsilon form, Open interval
</NEEDED_TERMS>
<NEW_TERMS>
Limit
</NEW_TERMS>
<PUPIL_FEATURES>
<LEVEL>
beginner
</LEVEL>
<PRSENTATION_FORM>
formal/narrative
</PRSENTATION_FORM>
<SKILL>
modest/quick_minded
</SKILL>
</PUPIL_FEATURES>
<CONTENT>
{***Here is one or more definitions of mathematical limit with formal, but not too formal, text***}
</CONTENT>
</DIDACTIC>

The signification of tags is obvious and may differ if some implementation facilities take place. Our system provides us with a lot of URLs as output of the search by "Mathematical Limit", but not all of them would be fitted to the current user. Due to <PUPIL_FEATURES> tag only adequate to his characteristics URLs have to form a real output. Besides if user's profile contains his stage and a number of current session and under point of view of needed terms user receives only needed portion of didactic information.

Such feature makes possible to obtain more adequate approach to the Semantic Web, which seems to be a thousand times more intelligent, then contemporary one.

REFERENCES

[1] R. Gottsdanker "Experimenting in Psychology", Prentice-Hall. Inc., Englewood Cliffs, New Jersey, 1982.

[2] C. James Goodwin "Research In Psychology: Methods and Design", 4th Edition, Wiley, 2004.

[3] E.G. Udaltsova "Development of the Metainformation Interface Concept for Synthesis of a Training Session in an Information Network", Proceedings of the 4th Scientific Conference "Management and information technologies", pp 150-152, Saint-Petersburg, October 2006.

[4] V.N. Yakimov, G.N. Dyakonov, E.G. Dyakonova, A.V. Mashkov "Soft Measurements in Structure N-Component NFL-Continuum", Proceedings of the IX International Conference on Soft Computing and Measurements, Vol. 2, pp 13-15, Saint-Petersburg, June 2006.

[5] Yu.A. Shreyder "About One Model of the Semantic Information Theory", Problems of Cybernetics, Vol. 13, Moscow, 1965.

[6] V.N. Yakimov, G.N. Dyakonov, A.V. Mashkov. "Verbal components of NFL-continnum as Structure and as System". – Proceedings of the IX International conference "Intelligent Systems and Computer Sciences. 23 - 27 October, 2006. Vol. 2, Part 2, pp. 323-326" Mechanical – Mathematical Faculty, Moscow State University. Moscow, 2006.

[7] T. Jörg, B. Davis, G. Nickmans. "Towards a New, Complexity Science of Learning and Education". Educational Research Review, 2 (2007) pp. 145 – 156.

[8] V.N. Yakimov, H.N. Dyakonov, A.V. Mashkov "Creation Ontology a Subject Domain on the Basis of Analysis NFL-Conlinuum". "Information Technologies", 3, 2006. Pp. 36-39.

[9] Anastasia Efclides. "Metacognition and Affect: What Can Metacognitive Experiences Tell Us About the Learning Process?" Educational Research Review, 1 (2006) pp. 3-14.

[10] V.I. Batishchev, G.N. Dyakonov, Ye.G. Dyakonova, V.Yu. Mishin. "Information technologies of cognitive process". "Open and Distant educacation", 4(20), 2005, pp. 49-55.