TAR

# The IPSI BgD Transactions Advanced Research

**Trans-disciplinary Issues in General Science and Engineering**

A publication of

**IPSI Bgd Internet Research Society**
**New York, Frankfurt, Tokyo, Belgrade**
**January 2010 Volume 6 Number 1 (ISSN 1820-4511)**

**Table of Contents:**



This is a special issue dedicated to: Bilateral Scientific Cooperation of Slovenia and Serbia, Project #BI-RS/08-09-007, in years 2008 and 2009 [E-speranto: vmesna platforma in razpis in shranjevanje večjezikovnih dokumentov s poudarkom na slovenščini, serbščini in ruščini]

# The IPSI BgD Internet Research Society

The Internet Research Society is an association of people with professional interest in the field of the Internet. All members will receive this TRANSACTIONS upon payment of the annual Society membership fee of €100 plus an annual subscription fee of €1000 (air mail printed matters delivery).

***Member copies of Transactions are for personal use only***

**IPSI BGD TRANSACTIONS ON ADVANCED RESEARCH**

www.internetjournals.net

# Interview – Mastery of Multimedia

**Furht, Borko**

1. **What is the essence of the contribution for which you received your highest recognition?**

In the mid 90s, my graduate students and I invented and developed two crucial technologies that enabled video and audio broadcasting over the Internet. These two technologies include a new video compression technique based on the 3-dimensional Discrete Cosine Transform, which outperforms MPEG compression schemes in terms of the compression ratio and real-time transmission, and IP Simulcast, which is an innovative Internet broadcast protocol, which provides inexpensive, efficient, and reliable audio and video broadcasting. The technologies have been reported in a number of publications, including several IEEE articles.

2. **What are the impacts of these contributions?**

These contributions represent the pioneering work in the field of peer-to-peer networks, which are today commonly used for video and audio broadcasting. Much of research work worldwide was based on these contributions.

3. **What are the applications of your contribution that may change everyday life?**

The applications of these technologies are tremendous, and include various applications of Internet-based broadcasting of video and audio, which today allow crisp audio and full-screen, high-quality video on-demand. When the technologies were developed, the start-up company Pipe Dream in West Palm Beach was founded in 1997. Several of my students worked in the company, and I served on the Board of Directors. The company has developed the first software product for video and audio broadcasting over the Internet. Later, the company was acquired by another company.

4. **Can you tell us, what are the issues that we have to teach our kids, so they become creative when they finish studies?**

It is very difficult to teach "creativity". Therefore, I will begin with the skills that students, who will become engineers and scientists, need to possess in order to be successful in today's world. In my time, we needed to be good mathematicians and understand technical issues. However, in today's global world, a successful engineer and scientist should also have other skills, including communication and writing skills, understand various cultures, and also possess interdisciplinary skills (for example bioinformatics or business management).

The role of a good teacher is to identify promising students and challenge them with projects and problems, which will develop their creativity.

5. **What are the major things to keep in mind, when you form a team for a scientific experiment, or similar?**

I will present my views, which relates to my experience when initiating a new research (rather than a scientific experiment) and establishing a research team.

First, the research begins with selecting the right topic of research. It is very important to understand very well what has been already done, so you don't "reinvent the wheel!" This phase consists of studying the literature including patents.

In today's world, it is almost impossible to conduct research, which is not funded. So, the second phase consists of finding sources for your research. In my case, these are either

Manuscript received April, 4th, 2009.
Dr. Borko Furht is with the Florida Atlantic University, Department of Computer Science and Engineering (e-mail: borko@cse.fau.edu)

government agencies, such as NSF, NASA, Department of Defense, and NIH, or private high-tech corporations. This phase consists of writing and submitting research proposals, contacting government agencies, and talking to industry representatives.

Once when the research funding is assured, the scientific team is created depending on the available finds. For large, interdisciplinary projects we include researchers and graduate students with various expertises from different departments, even colleges or universities. For small projects, a typical team consists of a few graduate students.

**6. What are the people to avoid, when trying to generate a breakthrough achievement?**

First, try to avoid publishing it, before protecting it. Protect it first, by applying for the patent. Patent takes lots of time to get approved; however, you can always apply for provisional patent, which can be obtained in a very short time and is valid for a year.

Second, make sure that you are the first in generating "this breakthrough achievement". I have seen many cases where researchers "invented" something that was invented earlier by simply not doing a good homework in researching the literature.

**7. What is your opinion about the impact of math?**

In my time, math was a very important topic for engineering students. It still is, if it is taught correctly. European students are much better in math than American students, which is a big weakness in American schools. On the other hand, American students are more pragmatic and practical.

The recipe is to find a good balance between theory (in this case math) and applications (specifically in terms of computer applications).

**8. When targeting a major breakthrough, how sensitive one has to be about the direct interests of tax-payers?**

In our case, this is very well defined at the university level. Every patent invented by a researcher at the university belongs to the university (it is the state university, so it basically belongs to tax payers). If the patent invention generates some revenues through licenses, than university gets 70% and the individual inventor gets 30%. This sounds a fair game.

**9. What is the driving force that motivates a person like you to continue to create and generate results after he receives such a big prize?**

Most researchers are not motivated with monitory awards; it is typically curiosity which motivates and drives true researchers to continue their work until they can produce. In our College, we have a famous researcher Dr. Bill Glenn, who is 82 years old and has 128 patents. He is still very active and he recently submitted another patent. In USA there is a relatively new law (about 15 years old) that university professors do not ever need to retire unless they select to do so. This gives them the opportunity to continue their creative work all their lives!

**10. For a small nation like Serbia, what is your advice, which road to take, when it comes to science?**

Leading scientific forums in Serbia, such as Academy of Science and Ministry of Science, should develop a long-term vision which scientific areas should be developed in Serbia and provide funding to these areas. It would be very wrong for a small nation like Serbia to try to do research in many different areas; this would be wasting of resources. In my field of computing and engineering, I see a great chance for Serbian researchers to conduct applied research and development in the areas of software development and Internet engineering. I would be delighted to see a number of start-up companies in these areas, which today can easily grow and become another Google, or YouTube, or Facebook. For this kind of businesses, with great ideas you don't need to be in USA; you can be in Serbia and be very successful worldwide!

### 11. Major impression about Serbian scientists that you met abroad?

Being a Serbian, who left Serbia in 1981 and went to become a professor in USA, I met a large number of Serbians before and after. I also had several PhD students from Serbia who graduated under my supervision. I am currently supervising three PhD students, two from Serbia and one from Montenegro.

I am proud to say that my best students were always from Serbia. They were all superb; they successfully completed their work toward their PhD degree in a record time, and they are all very successful today. For example, Dr. Daniel Socek, who graduated several years ago in the area of multimedia encryption, has a very successful consulting company in USA in this field; his main client is Real Networks Company. Dr. Dubravko Culibrk, who also graduated several years ago with the PhD degree, is the only Serbian student who decided to return to Serbia; he is currently Docent at the University of Novi Sad.

Besides my students, I met many Serbian scientists in USA and elsewhere. They are all very successful in their fields of expertise. My very close personal and professional friend is Dr. Leon Alkalai, who finished electrical engineering at the University of Belgrade and later received his PhD from UCLA in Los Angeles. He is one of the leading researchers and a manager of Robotic Lunar Exploration at NASA JPL Center in Pasadena, California. He recently became a member of the International Academy of Astronautics.

There are many other examples of successful Serbian scientists and researchers.

### 12. Major impressions from your visits to Serbia?

I visit Serbia every year, and I see some improvements from year to year. Belgrade looks nicer every time. As usually, restaurants are full, people are smart, tall and beautiful (young people!). Some of them are happy, some of them not!

In terms of research and science, my impression is that there is no vision, no collaboration, no team and interdisciplinary work, no focus on "niche" research. I can only see some individual success stories. Good researchers/scientists are just trying to survive doing mostly consulting and not true research. No available funds for grand research projects that can motivate researchers in a small nation like Serbia

I don't see that with this kind of research politics and research strategy Serbia will ever become visible on the world research scene!

### 13. Major impression about the University of Belgrade?

It is a great university, great people, and great students! However, there is a lack of well-developed research strategy and funding for research. Professors (potential researchers) are mostly focused on teaching and consulting and less on true applied research. The university definitely needs significant increase of resources to become a world-accepted research institution.

Dr. Furht has over twenty five years of academic and industry experience in the field computer science and engineering. He is Professor and Chairman of the Department of Computer Science at Florida Atlantic University (FAU) in Boca Raton, Florida. He has been a PI or Co-PI of a number of large projects funded by various government agencies such as NSF, ONR, NASA, and from industrial corporations, including IBM, Motorola, Hewlett Packard, Xerox, and others. He is an author of more than 250 technical and scientific publications, including 28 books.

# More Precise Fairness Bounds of Deficit Round Robin Scheduler

Kos, Anton; Tomažič, Sašo

***Abstract*— Fairness is a much desired property of scheduling algorithms. In a fair scheduler each data flow gets its fair share of the available bandwidth, and this share is not affected by the presence of other flows and their possible misbehaviour. Deficit Round Robin (DRR) scheduler is easy to implement, it provides fairness for flows with variable packet lengths, it allows bandwidth reservation, and it has *O*(1) complexity. The fairness of schedulers has frequently been evaluated using one of the two fairness measures: the absolute fairness measure (AFM) and the relative fairness measure (RFM). In this paper we carry out a thorough fairness analysis of DRR and we derive more precise bounds of both RFM and AFM. We show that our new fairness bounds are mathematically more accurate and that they give tighter approximation of the worst-case fairness behaviour of DRR scheduler, than the bounds derived earlier.**

**Index Terms—DRR, Deficit Round Robin, Fairness, Fairness Measure, Packet Networks, Scheduling**

## 1. Introduction

In packet networks with statistical multiplexing (like internet) overload causes congestion that is solved either by delaying or by dropping excess packets. There are different solutions that try to solve a challenge of assuring high resource utilization and high application performance at the same time. Essentially they can be grouped into two categories: end-system based solutions and router based solutions. In this paper we analyse router based solutions.

Many multimedia applications rely on the ability of the network to provide some sort of quality of service guarantees. The term Quality of Service (QoS) can generally be defined as a set of network mechanisms that satisfy the varied quality of service levels required by applications, while at the same time maximizing bandwidth utilization. Applications rely on traffic scheduling algorithms in switches and routers to guarantee performance bounds and meet the agreed QoS. There are several measures that are to be considered when choosing a scheduling algorithm. The most important are: fairness, latency and complexity.

The paper is organised as follows. In section 2 we briefly describe why fairness is an important property of each scheduler. In section 3 we give the advantages of our fairness measures over the existing ones. We continue with the explanation of the basics of Deficit Round Robin scheduler in section 4 and the definition of fairness measures in section 5. Sections 6 and 7 are the core of this paper where we derive the new relative and absolute fairness measures for the Deficit Round Robin scheduler. We conclude with the comparison of fairness measures of some well known schedulers in section 8.

## 2. Importance of Fairness

Packets belonging to different flows often share links in their transmission paths. Fairness is a very desirable property in the allocation of bandwidth on such links. In multiuser/multiapplication environments the protection guaranteed by fair scheduling improves the isolation between flows. Isolation offers more predictable performance of the system to users applications. Fair allocation of bandwidth ensures that the performance of one flow is not affected when another, possibly misbehaving flow, tries to send packets faster than its reserved rate. In addition, strategies and algorithms for fair management of network traffic can serve as a critical component of QoS mechanisms to achieve certain guaranteed services such as delay bounds and minimum bandwidths.

Since a scheduling algorithm should always provide the best possible QoS it has to be fair, it has a bounded maximum delay limit, low computational cost (complexity), easy implementation, and high efficiency.

## 3. Our Fairness Measures

Fairness bounds of DRR have been given or derived in many papers that discuss Latency-Rate or Round Robin like schedulers. Let us mention just the most important papers in this topic written by Shreedhar and Varghese [3] and Stiliadis and

Manuscript received May 4, 2009.
Kos, Anton is with the Faculty of Electrical Engineering, University of Ljubljana, Slovenia (e-mail: anton.kos@fe.uni-lj.si).
Tomažič, Sašo is with the Faculty of Electrical Engineering, University of Ljubljana, Slovenia (e-mail: saso.tomazic@fe.uni-lj.si).

Varma [7]. Their fairness bounds seem very similar to our bounds. Unfortunately they are inaccurate as they have made a slight mistake in presumption about the maximal size of a deficit counter and they have not taken into account the discrete nature of packet size. The detailed explanation of these differences is given in [1] and [2].

In sections 6 and 7 we give the simplified derivation of our fairness bounds. The exhaustive derivation with proofs can be found in [1] and [2]. Further comments about the differences in fairness bounds are given within the section 6.

The most important result of our paper is the fact that our fairness bounds for DRR are better than the ones calculated in [3] and [7].

## 4. Deficit Round Robin

Scheduling algorithms can be broadly classified into two categories: sorted priority schedulers and frame-based schedulers. Shreedhar and Varghese proposed one of the most popular frame-based scheduling algorithms in 1996 [3]. The main characteristic of all Deficit Round Robin (DRR) like scheduling algorithms is their ability to provide guaranteed service rates for each flow (queue). DRR services flows in a strict round-robin order. It has low complexity and it is easy to implement. Its latency is comparable to other frame-based schedulers. Below is the list of variables used in our analysis:

| | |
|---|---|
| $R$ | transmission rate of an output link, |
| $N$ | the total number of active flows, |
| $r_i$ | the reserved rate of flow $i$, |
| $w_i$ | weight assigned to each flow $i$, |
| $Q_i$ | quantum assigned to flow $i$, |
| $DC_i$ | Deficit Counter of flow $i$, |
| $F$ | frame size, |
| $L_{max}$ | the maximum possible packet size. |

A better fairness property of DRR comparing to Round Robin (RR) is achieved by maintaining a variable called *Deficit Counter* for each of the flows. It is denoted with $DC_i$, for flow $i$, and it stores the value of the deficit that a flow has accumulated in its active period (a flow is active during any period of time when packets belonging to that flow are continuously queued in the system). DRR assigns a *quantum* of service $Q_i$ to each flow $i$. When a flow is on its turn for service, its $DC_i$ is incremented by the quantum of that flow. A flow is served only if the packet at the head of the flow $i$ is smaller or equal than the current value of $DC_i$. Otherwise the scheduler begins servicing the next flow in the round robin sequence. Flow $i$ is being serviced until packets at the head of the flow remain smaller that the current value of $DC_i$. When a packet is removed from the flow $i$, the $DC_i$ is decremented for the length of the packet. A detailed operation of DRR algorithm can be found in [3].

Because all flows share the same output link, a necessary constraint is that the sum of all reserved rates must be less or equal to the transmission rate of the output link:

$$\sum_i r_i \leq R \tag{1}$$

Let $r_{min}$ be the smallest of $r_i$: $r_{min} = \min_{\forall i} r_i$. Each flow $i$ is assigned a weight that is given by:

$$w_i = \frac{r_i}{r_{min}}. \tag{2}$$

Note that $\forall i \in 1, 2, \cdots, N$ holds $w_i \geq 1$. Each flow $i$ is assigned a quantum of $Q_i$ bits, that is a whole positive value, i.e. $Q_i \in \mathcal{N}$. This quantum is actually the amount of service that the flow should receive during each round robin service opportunity. Let us define with $Q_{min}$ the minimum of all the quanta. Then the quantum for each flow $i$ is expressed as:

$$Q_i = w_i Q_{min}. \tag{3}$$

## 5. Fairness

When there is contention for resources, it is important for resources to be allocated fairly. Among scheduling algorithms significant discrepancies may exist in service provided to different flows over the short term. For example, two scheduling algorithms may have the same delay guarantees but can have very different fairness behaviours.

There is no commonly accepted method for estimating the fairness of a scheduling algorithm. In general, we would like the system to always serve flows proportional to their reserved rate and distribute the unused bandwidth left behind by idle flows proportionally among active ones. In addition, flows should not be penalized for excess bandwidth they received while other flows were idle.

Based on this intuitive definition of fairness the Generalized Processor Sharing (GPS) [5] is identified as an ideal resource sharing discipline. GPS is defined with respect to a fluid-model, where packets are considered to be infinitely divisible. The share of bandwidth reserved by flow $i$ is represented by a real number $r_i$. Let us assume that there is $N$ active flows in the system during the time interval ($t_1$, $t_2$) and let us with $SENT_i(t_1, t_2)$ denote the amount of service received by flow $i$ during the same time interval. According to [7] the next inequality holds:

$$SENT_i(t_1, t_2) \geq \frac{r_i}{\sum_{j=1}^{N} r_j} R(t_2 - t_1) \tag{4}$$

where $R$ is the link rate. The minimum service that a flow can receive in any interval of time is:

$$\frac{r_i}{\sum_{j=1}^{V} r_j} R(t_2 - t_1). \tag{5}$$

where $V$ is the maximum number of active flows in the server at the same time. It is also true that $N \leq V$ and

$$\sum_{j=1}^{V} r_j \leq R \tag{6}$$

Thus, at each instant GPS serves each active flow with a maximum rate equal to its reserved rate; in addition, the excess bandwidth available from flows not using their reserved rate is distributed among all others backlogged ones in proportion to their reservations. This results in perfect isolation among all flows and the ideal fairness. However, GPS scheduler is not implementable since in real packet-switched networks data are forwarded in units of packets and not in infinitesimal quantities.

The fairness of scheduling algorithm used in communication networks has frequently been evaluated using two fairness measures. One of them is known as Absolute Fairness Measure (AFM), and it is based on the maximum difference between the service received by a flow under the discipline being measured and the service it would receive under the ideal GPS policy. The AFM of a scheduler is frequently hard to obtain. Therefore an alternative measure, known as the Relative Fairness Measure (RFM) is used. It is based on the maximum difference between the services received by two different flows under the scheduling algorithm being measured. In this paper we will use both of them to give a better perspective of the fairness of DRR.

To give more appropriate definitions of these measures we define some figures of merit. Since arriving packets of different flows are stored in different queues, let us denote with: $SENT_i(m)$ the number of bytes sent out for flow $i$ in round $m$, and by $SENT_i(p, m)$ the number of bytes sent out for flow $i$ starting from round $p$ to round $m$. According to [9] we have definitions:

*Definition 1: Absolute fairness bound of flow $i$ over time interval $(t_1, t_2)$, is denoted by:*

$$AF_i(t_1, t_2) = \left| \frac{SENT_i(t_1, t_2)}{w_i} - \frac{SENT_{gps_i}(t_1, t_2)}{w_{gps_i}} \right| \tag{7}$$

*where $SENT_{gps_i}$ denotes the total service received by flow $i$ under GPS scheduling policy and $w_{gps_i}$ is the weight assigned to flow $i$ in GPS scheduling policy.*

*Definition 2: The absolute fairness, $AF$, of an algorithm over some interval of time $(t_1, t_2)$ and absolute fairness measure, $AFM$, of an algorithm are given with:*

$$AF(t_1, t_2) = \max_{\forall i} AF_i(t_1, t_2) \tag{8}$$

$$AFM = \max_{\forall (t_1, t_2)} AF(t_1, t_2) \tag{9}$$

Relative fairness measure can be meaningfully defined only with respect to flows that are all backlogged during the time interval under consideration. Following previous work we will assume that all flows under consideration are active at all instants of time during that interval.

*Definition 3: The relative fairness with respect to a pair of flows $(i, j)$ over time interval $(t_1, t_2)$, denoted with $RF_{(i,j)}(t_1, t_2)$, is:*

$$RF_{(i,j)}(t_1, t_2) = \left| \frac{SENT_i(t_1, t_2)}{w_i} - \frac{SENT_j(t_1, t_2)}{w_j} \right| \tag{10}$$

*Definition 4: The relative fairness with respect to flow $i$ over some interval of time $(t_1, t_2)$, denoted with $RF_i(t_1, t_2)$, is:*

$$RF_i(t_1, t_2) = \max_{\forall j} RF_{(i,j)}(t_1, t_2) \tag{11}$$

*Definition 5: The relative fairness, $RF$, of an algorithm over some interval of time $(t_1, t_2)$ and relative fairness measure, $RFM$, of an algorithm are:*

$$RF(t_1, t_2) = \max_{\forall i} RF_i(t_1, t_2) \tag{12}$$

$$RFM = \max_{\forall (t_1, t_2)} RF(t_1, t_2) \tag{13}$$

*Definition 6: A scheduling algorithm is* **fair** *if $AFM$ and $RFM$ are both bounded by some small constant.*

## 6. Relative Fairness Measure

We can now proceed with the analysis of a DRR algorithm. For the purpose of the analysis we need lemmas that are proved in [1] and [2].

*Lemma 1: For all integers $i$, the following statement holds:*

$$0 \leq DC_i \leq L_{max} - 1 \tag{14}$$

*where $L_{max}$ denotes the size of the maximum packet that can arrive.*

*Lemma 2: Let us consider any execution of the DRR scheme and any interval $(t_1, t_2)$ of that execution, such that flow $i$ is constantly backlogged during given time interval, and let $m$ be the number of round robin service opportunities received by flow $i$ during the time interval in progress. Then the following inequality holds:*

$$mQ_i - L_{max} + 1 \leq SENT_i(1, m) \leq mQ_i + L_{max} - 1$$

*Lemma 3: A basic invariant of DRR algorithm is that during arbitrary interval in which two queues, $i$ and $j$, are backlogged, between any two round robin service opportunities given to queue $i$, queue $j$ must have had a round-robin opportunity.*

*Theorem 1: For an arbitrary interval in time and in any execution of DRR service discipline the following inequality holds:*

$$\begin{aligned} RFM &= \left| \frac{SENT_i(m)}{w_i} - \frac{SENT_j(m')}{w_j} \right| \quad (15) \\ &\leq \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} + Q_{min} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned}$$

*where $m$ is a number of round robin opportunities given to flow $i$ in the interval $(t_1, t_2)$ and $m'$ is a number of round robin opportunities given to flow $j$ in the same interval.*

*Proof 1:* Let us consider an arbitrary time interval $(t_1, t_2)$ in any execution of DRR algorithm and arbitrary two flows $i$ and $j$ that are backlogged during this interval. According to the Lemma 3 we have:

$$|m - m'| \leq 1 \quad (16)$$

Now according to Lemma 2 we have:

$$\frac{SENT_i(1,m)}{w_i} \leq \frac{mQ_i}{w_i} + \frac{L_{max}}{w_i} - \frac{1}{w_i} \quad (17)$$

$$\frac{SENT_j(1,m')}{w_j} \geq \frac{m'Q_j}{w_j} - \frac{L_{max}}{w_j} + \frac{1}{w_j} \quad (18)$$

We have used only the one side inequalities stated in Lemma 2, for queue $i$ its right hand side and for queue $j$ its left hand side. Since from (3) we have:

$$\frac{SENT_i(1,m)}{w_i} \leq mQ_{min} + \frac{L_{max}}{w_i} - \frac{1}{w_i} \quad (19)$$

$$\frac{SENT_j(1,m')}{w_j} \geq m'Q_{min} - \frac{L_{max}}{w_j} + \frac{1}{w_j} \quad (20)$$

Then it follows

$$\begin{aligned} & \left| \frac{SENT_i(1,m)}{w_i} - \frac{SENT_j(1,m')}{w_j} \right| \\ \leq & \left| mQ_{min} + \frac{L_{max}}{w_i} - \frac{1}{w_i} - m'Q_{min} + \frac{L_{max}}{w_j} - \frac{1}{w_j} \right| \\ \leq & |Q_{min}(m - m')| + \left| \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \right| \\ = & Q_{min}|m - m'| + \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \quad (21) \end{aligned}$$

In order to justify equality in the last line of expression (21), we have to prove that the expression $\frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j}$ is always greater or equal to zero. Reorganizing the expression to:

$$\frac{L_{max} - 1}{w_i} + \frac{L_{max} - 1}{w_j}$$

and since $L_{max} \geq 1$ and $w_i$, $w_j \geq 1$ we can easily see that our statement holds. Now we get:

$$\begin{aligned} & \left| \frac{SENT_i(1,m)}{w_i} - \frac{SENT_j(1,m')}{w_j} \right| \quad (22) \\ \leq & Q_{min}|m - m'| + \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned}$$

Using expression (16) and inequality (22) we can conclude:

$$\begin{aligned} & \left| \frac{SENT_i(1,m)}{w_i} - \frac{SENT_j(1,m')}{w_j} \right| \quad (23) \\ \leq & Q_{min} + \frac{L_{max}}{w_i} + \frac{L_{max}}{w_j} - \frac{1}{w_i} - \frac{1}{w_j} \end{aligned}$$

■

*Comment 1:* Our objective is to find the maximum value of the expression on the right hand side of inequality (23) and along with that the upper bound of RFM. Since our interest is only to find the local maximum, we can proceed by analysing every variable separately. $L_{max}$ is defined by system properties, so we can treat it as a constant.

$$Q_{min} + \frac{L_{max} - 1}{w_i} + \frac{L_{max} - 1}{w_j} \quad (24)$$

So the expression (24) has its maximum when both $w_i$ and $w_j$ take their minimal possible value that is 1.

We can conclude that for a given $Q_{min}$, the upper bound of RFM occurs when at least two flows have their quanta equal to $Q_{min}$. Legitimacy of this statement lies in the definition 5 because we are only interested in the worst possible case between any two flows.

| Scheduling algorithm | Relative fairness bound |
|---|---|
| GPS | $0$ |
| FIFO | $\infty$ |
| PGPS = WFQ | $\max_{i,j}\left(\max\left\{C_j+\frac{L_{max}}{w_i}+\frac{L_j}{w_j}, C_i+\frac{L_{max}}{w_j}+\frac{L_i}{w_i}\right\}\right)$, $C_i=\min\left\{(N-1)\frac{L_{max}}{w_i}, \max_{1\le n\le N}\frac{L_n}{w_n}\right\}$ |
| Virtual Clock Fair Queuing | $\infty$ |
| SCFQ | $\frac{L_i}{w_i}+\frac{L_j}{w_j}$ |
| Worst Case Weighted Fair Queuing | $\frac{L_i}{w_i}+\frac{L_j}{w_j}$ |
| Frame Based Fair Queuing | $\max\left\{\frac{2F-Q_i}{R}+\frac{L_i}{w_i}, \frac{2F-Q_i}{R}+\frac{L_j}{w_j}, \frac{L_{max}}{w_i}+\frac{L_j}{w_j}, \frac{L_{max}}{w_j}+\frac{L_i}{w_i}\right\}$ |
| Packet Based Round Robin | $\infty$ |
| DRR - old | $\frac{L_{max}}{w_i}+\frac{L_{max}}{w_j}+Q_{min}$ |
| Surplus Round Robin | $\frac{L_{max}}{w_i}+\frac{L_{max}}{w_j}+Q_{min}$ |
| Elastic Round Robin | $\frac{L_i}{w_i}+\frac{L_j}{w_j}+L_{max}$ |
| DRR - new | $\frac{L_{max}}{w_i}+\frac{L_{max}}{w_j}+Q_{min}-\frac{1}{w_i}-\frac{1}{w_j}$ |

TABLE 1

**Relative fairness bounds for some well known scheduling algorithms**

The RFM bound expression (23) may seem very similar to the RFM bound calculated in [3] and [7]. But it is not. In addition to the fact, that in [3] RFM has not been correctly derived, it should be emphasized that our RFM bound is smaller by some margin than the one calculated in [3] or in any other work to our knowledge. We can conclude that our RFM result is always better than the one calculated in [3] and [7] for the original DRR algorithm.

## 7. Absolute Fairness Measure

Absolute fairness measure is intuitively closer to us because it shows the difference between the fairness of scheduler being analysed and the fairness of ideal GPS scheduler. We use the relationship between the RFM and AFM derived in [9].

*Lemma 4: The relationship between the relative and absolute fairness measures, under any work-conserving scheduling policy, is described with the inequality:*

$$AFM \le \left(1-\frac{Q_{min}}{F}\right) RFM \qquad (25)$$

*where $F$ denotes frame size, i.e. $F=\sum_i Q_i$.*

Using Lemma 4, that was proved in [9], we have that for an arbitrary interval in time and in any execution of DRR service discipline we have:

$$AFM \le \left(1-\frac{Q_{min}}{F}\right)\left[\frac{L_{max}}{w_i}+\frac{L_{max}}{w_j} + Q_{min}-\frac{1}{w_i}-\frac{1}{w_j}\right] \qquad (26)$$

When $Q_{min}$ is large the absolute fairness can be low even if the corresponding relative fairness measure is high.

We have already discussed relative fairness bound and its worst and best case, what is left to be analysed is the factor $(1-\frac{Q_{min}}{F})$. Its minimum value is being reached when $\frac{Q_{min}}{F}$ reaches its maximum value, i.e. when $Q_{min}$ reaches its maximum value, since frame size $F$ is a constant. We have previously concluded that $Q_{min}$ takes its maximum value when all the quanta are equal. In that case AFM is the smallest and the furthest from RFM.

## 8. Comparison of DRR with Some Other Schedulers

Let us compare fairness bounds of DRR and some of the schedulers from Table 1 that lists fairness bounds for some of the most known scheduling algorithms. Some of them belong to the class of sorted-priority schedulers and others to the class of frame-based schedulers. Thus, this table shows descriptive differences in these algorithms looking through their fairness measure properties.

We observe that the bounding function of PGPS takes higher values than the one just derived for DRR scheduler. The difference between PGPS and DRR is only in factors $C_i$ and $Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$. If we look closer into the definition of $C_i$,

$$C_i = \min\left\{(N-1)\frac{L_{max}}{w_i}, \max_{1\leq n\leq N}\frac{L_n}{w_n}\right\}, \qquad (27)$$

we see that since both factors: $(N - 1)\frac{L_{max}}{w_i}, \max_{1\leq n\leq N}\frac{L_n}{w_n}$ are bigger than $Q_{min}$, their minimum is bigger too. It follows that the discussed minimum is also bigger than $Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$. This leads us to conclusion that DRR algorithm has better fairness property than the PGPS scheduler.

When we compare expressions for SCQF and DRR given in Table 1, we see that they differ in factor $Q_{min} - \frac{1}{w_i} - \frac{1}{w_j}$. If this factor is bigger than zero, the fairness of DRR algorithm is worse, since in original DRR algorithm the minimum value of quanta assigned to flows should be bigger than $L_{max} - 1$ for DRR to have an $O(1)$ complexity. So SCFQ scheduler has better fairness properties than DRR scheduler.

In the class of frame-based schedulers DRR has slightly better fairness bound than the rest of them. The difference is in two subtracting factors, $\frac{1}{w_i}$ and $\frac{1}{w_j}$, which contribute to its better fairness bound.

## 9. Conclusion

In this paper we have derived a new and improved fairness bounds for DRR scheduling algorithm that contribute to more exact analysis of DRR. We have shown that our fairness bounds are tighter than other previously derived bounds for DRR scheduler. More detailed DRR fairness analysis can be found in [1] and [2].

## References


[1] Anton Kos, *Zagotavljanje različnih stopenj kakovosti storitve v omrežjih s paketnim prenosom podatkov*, Doctoral Thesis, Faculty of Electrical Engineering, University of Ljubljana, February 2006

[2] Anton Kos, Jelena Miletić: *Sub-Critical Deficit Round Robin*, Technical Report, Faculty of Electrical Engineering, University of Ljubljana, July 2005

[3] M. Shreedhar, George Varghese: *Efficient Fair Queuing Using Deficit Round Robin*, IEEE/ACM Transactions on Networking, Volume 4, Issue 3, June 1996

[4] J. M. Arco, D. Meziat, B. Alarcos: *Deficit Round Robin Alternated - A New Scheduling Algorithm*, Conference on Protocols For Multimedia Systems - PROMS2000, Cracow, Poland, October 22-25, 2000

[5] Abhay K. Parekh, Robert G. Gallager: *A Generalized Processor Sharing Approach to Flow Control in Integrated Services Networks: The Single-Node Case*, IEEE/ACM Transactions on Networking, Volume 1, Issue 3, June 1993

[6] Chuck Semeria: *Supporting Differentiated Service Classes - Queue Scheduling Disciplines*, White Paper, Juniper Networks, Inc., December 2001

[7] D. Stiliadis, A.Varma: *Latency-Rate Servers - A General Model for Analysis of Traffic Scheduling Algorithms*, IEEE/ACM Transactions on Networking, Volume 6, Issue 5, October 1998

[8] Nihal Pekergin: *Stochastic Bounds on Delays of Fair Queueing Algorithms*, Proceedings of the Conference on Computer Communications (IEEE Infocom), New York, March 1999

[9] Yunkai Zhou, Harish Sethu: *On the Relationship Between Absolute and Relative Fairness Bounds*, IEEE Communications Letters, Vol. 6, No. 1, January 2002

[10] Jian Zhang, Timo Hamalainen, Jyrki Joutsensalo: *Stochastic Analysis of Upper Delay Bound of GPS-Based Packetized Fair Queuing Algorithms*, Dept. of Mathematical Information Technology, University of Jyvaskyla, Finland

[11] S. Golestani: *A self-clocked fair queueing scheme for broadband applications*, Proc. IEEE INFOCOM, Toronto, ON, Canada, June 1994

# Interpretation of Multilingual Documents in e-speranto Using the Client-Server Architecture Model

Jakus, Grega; Varga, Emil; and Tomažič, Sašo

**Abstract —** ***The paper introduces an overview of a multilingual translation system based on e-speranto.***

***E-speranto is a computer language intended for recording multilingual documents on the Web. It can also serve as an intermediate language, or interlingua, in multilingual translation. Its advantage over similar approaches is the compatibility with HTML and the intelligibility of documents both to computers as well as to people.***

***The development of the e-speranto-based translation system consists of four stages. They include the development of language, tools for writing documents directly in e-speranto, interpreters of e-speranto and translators from a natural language to e-speranto.***

***A proof-of-concept implementation of the translation system uses the components from the first three stages of the development. These components are organised using the client-server architecture model.***

***By using a Web browser, the user can request the interpretation of a multilingual document. On the server side, the requests are passed to the interpreter of e-speranto. The result of the interpretation is a HTML document that is returned to the client. Language-specific data, such as language rules, word and phrase dictionaries, are contained in a database. The database can be located on the client or on the server.***

***The interpretation of a document in e-speranto is realized in three steps that include semantic generation, lexical transformation and structural transformation.***

***The interpreters implement a modular architecture. The modules are divided into three layers of abstraction. A level of abstraction refers to the degree of abstractness of language structures that enter a certain module as data and on which transformations are performed. The procedures in the modules in the first layer perform language-independent operations that are common to rule-based machine translation. The procedures in the second layer are typical of a group of languages, while the ones in the third layer are language-specific.***

Manuscript received April 28, 2009.

G. Jakus and S. Tomažič are with the Faculty of Electrical Engineering, University of Ljubljana, Slovenia. E. Varga is with the School of Electrical Engineering, University of Belgrade, Serbia (e-mails: grega.jakus@fe.uni-lj.si, emil.varga@gmail.com saso.tomazic@fe.uni-lj.si).

**Index Terms —** ***client, e-speranto, interpretation, multilingual translation, server***

## 1. Introduction

THE multilingualism is gaining importance in today's society when the World is in the process of globalization. The barriers and divides among different nations and cultures are blurring. A typical example of such a process in Europe is the uniting of nations and their economies in the European Union. While different markets are merging into single one, the language divides remain.

One of the demanding problems related to the multilingualism on the Web is the number of units needed for translation among different languages. The problem is burning since there are practically no political, geographical or any other kind of obstacles for the interaction among individuals. One can come across any language while using the Internet. In order to develop translators for all known 6900 languages that are spoken in the World today, about 47,610,000 of translators should be made.

One of the possible solutions to this problem is the use of an intermediate language or interlingua. Interlingua is an abstract presentation of the content that is independent from any natural language [1]. The record in interlingua must contain the whole information required for generating text in a natural language. Thus, the entire meaning we want to express in a natural language must be captured in interlingua. The advantage of using the interlingua is a two-phase course of translation between two natural languages. During the process, the modules that perform the conversion from a natural language to the interlingua (translators) are independent of those that perform the opposite conversion (interpreters). Moreover, the interpreters and translators of different languages are also mutually independent. The effect of this independence is the reduction of the number of units that would be needed in case of direct mapping among the individual languages. The cost of the latter approach is as high as $n(n-1)$, where $n$ denotes the number of languages among which we want to translate. By using the interlingua approach, the cost of the interpreter development reduces to *2n*, since only a translator and an interpreter for each language must be made.

## 2. Related Work

Numerous attempts of creating an interlingua-based multilingual system were conducted in the past. Some more notable implementations are presented in the next paragraphs.

DLT (*Distributed Language Translation*) [2] was a project of the development of a multilingual system in the 1980s that used an adapted version of Esperanto as an interlingua. The document written in Esperanto would be carried over the network and interpreted in a chosen language by the target computer. Although DLT presented a novel and interesting approach to machine translation, the results were not promising in practice.

The KANT system [3] is based upon controlled English (a language with a limited scope of vocabulary) and was created with the intention of translating technical documentation. It produces very accurate sentences, but due to the limited field of use it is not directly applicable for general multilingual translation.

UNL (*Universal Networking Language*) [4] is a computer language for recording and exchanging information and it is basically intended for communication on the Web. It supports 15 languages, which makes it currently one of the largest multilingual systems intended for use on the Web. Its main deficiency is the limited power of expressiveness [5] and somewhat lower degree of intelligibility of texts written in this language. The latter already proved as a disadvantage during the development of Internet standards in the past.

Generation of a text from an interlingua can be carried out in different manners [1]. Most often it is based on language rules (rule-based) that define the conversion from a source presentation to the target one. Another widely used approach is based upon the semantic and pragmatic knowledge of a certain field (knowledge-based). The statistical and example-based approach are not generally used in generation of texts from an interlingua because they both require a bilingual corpus which is, however, hard to create when one of the languages is an interlingua.

The majority of the approaches of the rule-based natural language generation has a modular design with two basic steps of conversion. They represent the lexical transformation (conversion of the lexical units) and structural transformation between the language structures in both languages [1]. When generating text from an interlingua, the *semantic generation* is also present. During this step, the deep syntactic structure of the content is generated according to the semantics in the interlingua representation.

An example of a well-established framework that uses similar design is ARIANE [6]. ARIANE is a flexible framework for the development of machine translation systems between language pairs which can also be an interlingua and a natural language. The system separates the algorithms from the linguistic content as the parameterization of algorithms. Several interpreters of the interlingua are based on ARIANE, for example in French [7], [8].

## 3. e-speranto

E-speranto [9], [10] is a formal computer language for recording multilingual texts. Its main goal is to overcome language divides in the Internet. A document in e-speranto is interpreted in a chosen language, when the user requests the document to be displayed.

The basic syntax of e-speranto is based on XML (*eXtensible Markup Language*) which is an important technology on the Web today. XML is compatible with HTML and therefore e-speranto can be incorporated into Web pages. Syntactic and grammatical rules are taken from Esperanto, but are expressed explicitly by means of metadata.

The main advantage of e-speranto over similar approaches is especially the intelligibility of documents both to computers as well as to people and compatibility with existing Web technologies, such as XML and HTML.

The development of the e-speranto-based translation system consists of four stages. They include the development of language, tools for writing documents directly in e-speranto, interpreters of e-speranto in a natural language and translators from a natural language to e-speranto. When the interpreters of e-speranto are developed, the displaying of Web pages in e-speranto will be possible. When the translators from natural languages to e-speranto are available, also multilingual translation between different pairs of languages using e-speranto as an interlingua will be feasible.

## 4. Interpretation of e-speranto in Natural Languages

The interpretation of a document in e-speranto is based on tree transformations. A tree consists of symbols that denote concepts, concept attributes and the relations among concepts. It contains enough information to perform the interpretation on the semantic, morpho-lexical and syntactic level. A tree is represented in the syntax of Mathematica programming language in which the core of the prototype interpreter INES (the **IN**terpreter of **E**-**S**peranto) was made.

Because e-speranto is a computer language, we can draw some parallels with the interpreters and translators of computer programming

```
<sentence original="E-speranto is a design of a computer language." feelings="declarative" organization="simple">
  <subject detail="personal_name" number="singular">
    <word>E-speranto</word>
  </subject>
  <predicate detail_predicate="main" mood="indicative" voice="active" tense="present" person="third">
    <word>be</word>
    <subordinate>
      <predicate detail_predicate="predicate_noun" number="singular">
        <word>design</word>
        <subordinate>
          <object detail_object="of_genitive" number="singular">
            <word>language</word>
            <subordinate>
              <attribute>
                <word>computer</word>
              </attribute>
            </subordinate>
          </object>
        </subordinate>
      </predicate>
    </subordinate>
  </predicate>
</sentence>
```

**Figure 1: Record of a sentence in e-speranto.** The basic building element in e-speranto is a clause. A clause is a semantic unit that corresponds to a sentence in a natural language. Clauses are composed of sentence elements introduced by XML tags. The grammatical characteristics are expressed explicitly by means of XML attributes. The concepts representing the essence of e-speranto are marked in English for the sake of better intelligibility.

languages regarding the interpretation. Namely, the process of interpretation is similarly divided into several stages. We distinguish:

- lexical and syntactical analysis of the source code,
- generation of the intermediate code,
- code optimization, and
- compilation phase.

The lexical and syntactic analyses are performed every time during the composition of a document in e-speranto and are provided by the development environment. For this purpose the development environment based on the Eclipse platform was developed. The built-in XML editor performs the verification of the document conformity with the e-speranto grammar, syntax and vocabulary.

The other phases are realized in the INES interpreter. The conversion of the e-speranto document into the expressions of the Mathematica language is analogous to the generation of the intermediate code. The phase of code optimization in the classical translators corresponds to the adaptation of the intermediate representation structure to the form that is used in the process of interpretation in INES (compare Figures 1 and 2). The phase of optimization is important since it enables, to a certain degree, the independence of the tree structures in INES from the changing grammar and syntax of e-speranto, as the latter is still being developed.

The compilation phase is the main step of the interpretation in a selected natural language. In INES this phase is realized in three steps.

In the first step, the so-called *semantic generation* is carried out. For each e-speranto concept a syntactic structure is generated that represents the same meaning in a natural language as the corresponding e-speranto concept. This is in accordance with the *principle of compositionality* which states that the meaning of a phrase is a function of its constituents and their position in a phrase. In the interpretation of e-speranto, this phase is important for the generation of common phrases and idioms.

In the second step (the lexical transformation), lexical units are formed on the basis of lexico-morphological attributes of concepts in e-speranto and linking of these concepts to the lexical units in a natural language (e.g. e-speranto record {*'lingua', tense: plural*} is interpreted as '*languages*' in English).

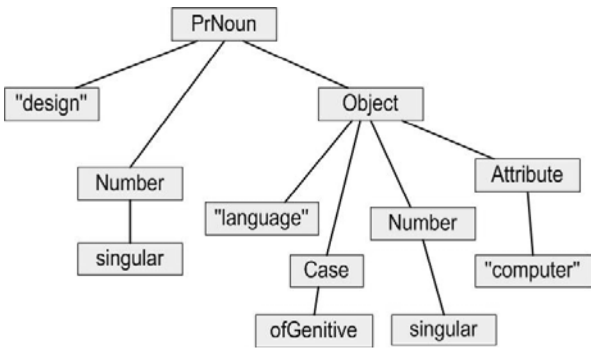


**Figure 2: Simplified representation of the noun phrase "*design of a computer language*"** from the sentence in the Figure 1 in the form of the tree structure of the symbolic programming language Mathematica.

In the third step, the structural transformation is carried out. The tree that comprises lexical units is rearranged according to the purpose of the message (e.g. the mood of the sentence), the syntax of the target language, etc. This step also includes the processing of punctuation and conjunctions. Finally, the tree is transformed to a sequence of lexical units representing a sentence in a natural language.

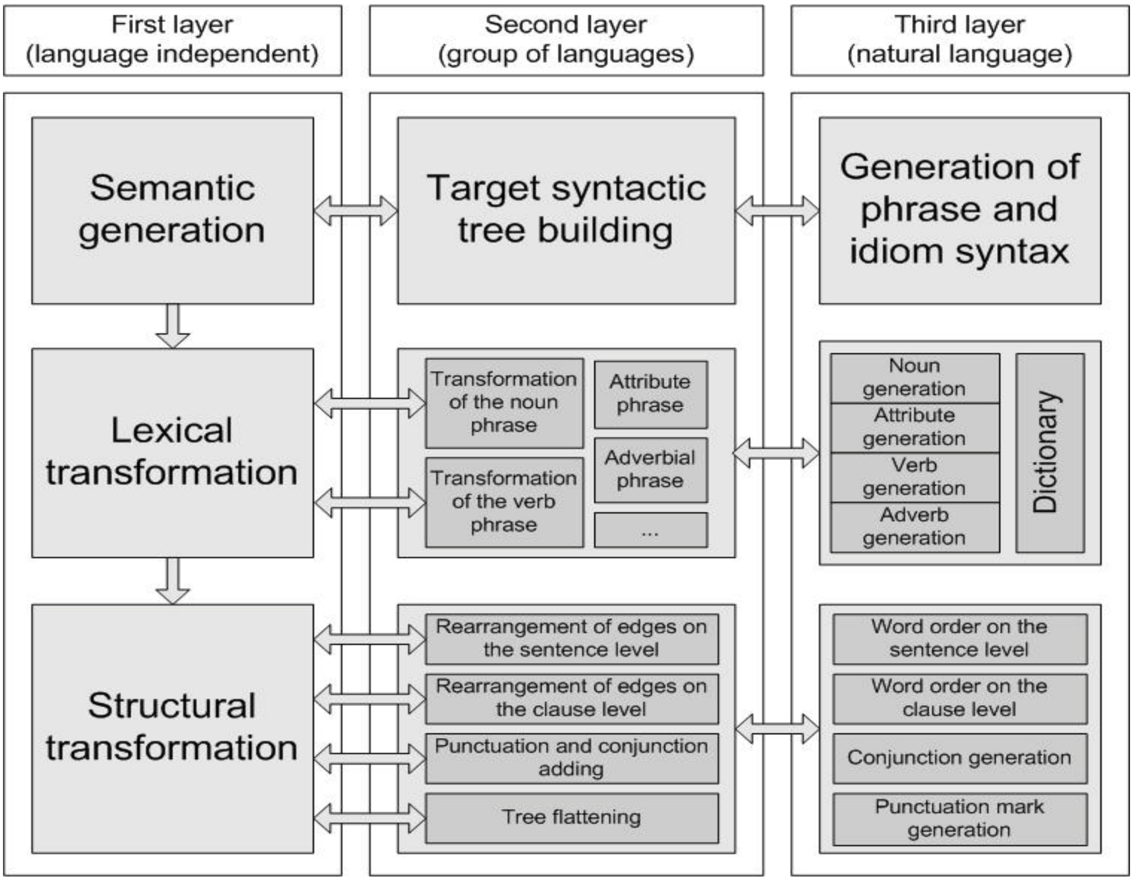


**Figure 3: Scheme presenting operation of INES.** Every phase of interpretation is divided into three layers. The procedures in the modules in the first layer perform language-independent operations that are common to rule-based machine translation. The procedures in the second layer are typical of a group of languages, while the ones in the third layer are language-specific.

## 5. ARCHITECTURE OF THE INTERPRETERS

The interpretation is carried out with the modules that are arranged into three levels of abstraction [11]. A level of abstraction refers to the degree of abstractness of language structures that enter a certain module as data and on which transformations are performed. The architecture of INES with some distinctive procedures on individual layers is shown in Figure 3.

The first layer comprises modules that dictate the course of interpretation and are independent of the target language. The layer is only aware of the fact that a sentence in a natural language contains the elements that express an action or activity (i.e. the predicate) and the holder of this action or activity (i.e. the subject). This layer also contains the algorithms for movements in the tree structure. Among the various possible methods the recursive *depth-first search* algorithm is implemented in INES. The individual subtrees are identified according to their type; their transformation is then performed by the lower layers.

The modules in the second layer are closer to the language families. These modules in general perform transformations of particular subtrees in accordance with their type. The type of a subtree is determined by the syntactic and/or semantic role the root element is performing according to the parent element in the tree representation. In general, a subtree of a certain type corresponds to a particular clause or its part in a sentence of a natural language. Figure 2 shows a subtree that corresponds to a predicate noun in a natural language.

The procedures with language-specific rules can be found in the third layer. These procedures map the parts of a tree structure to the elements of a natural language in a way that is specific to the language of interpretation. An example of such a transformation is the replacement of the e-speranto concepts and their attributes with the words of the target language or the rearrangement of the tree edges in accordance with the word order in which particular clauses appear in the target language. The access to word and phrase dictionaries is also implemented in this layer.

The characteristics that are common to a certain group of languages can be introduced on an abstractly higher level (e.g. on the second layer for a particular group of languages) than the actual characteristics of individual languages in this group (the third layer). The use of this concept reduces the cost of development for enabling some of the modules to be reused when developing the interpreters of related languages.

## 6. ARCHITECTURE OF THE SYSTEM PROTOTYPE

The implementation of the translation system prototype (Figure 4) is roughly divided into two parts: the client side and the server side. The server side is only present during the development of the e-speranto language and the interpreters. When the language is standardized, the interpreters will migrate from the server to the client. The functionality of the interpreters will be implemented in browsers through the plug-ins.

The performance of the system prototype can be tested on the project's Web site [12].

### 6.1. Client side

On the client side, the documents in e-speranto are composed. For this purpose an integrated development environment (IDE) based on the Eclipse platform [13] was developed. The IDE provides the user with the content assistance, document validation, access to the dictionaries and their effective use during the composition of documents.

By using a Web browser, the user requests the interpretation of a multilingual document that is located on the server.

### 6.2. Server side

On the server side, Apache Tomcat, a Java HTTP Web server, handles the requests for the interpretation of documents in e-speranto. The requests are passed to the INES. The interface between the Internet and the core of the interpreter is written in Java programming language.

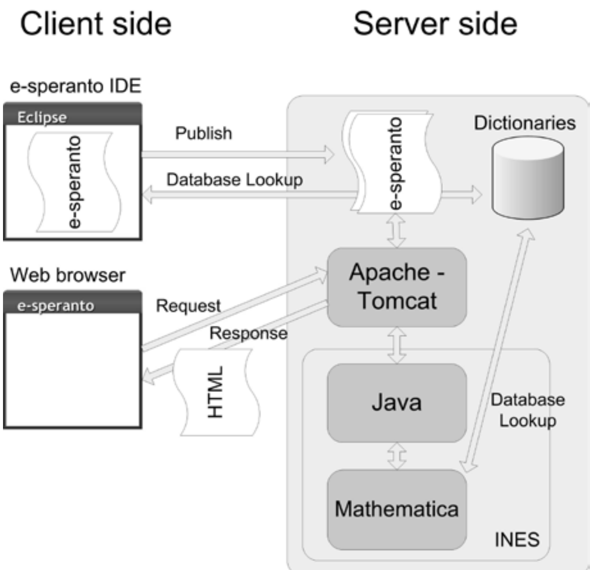


**Figure 4: Implementation of the multilingual translation system** based on e-speranto using the client-server architecture model.

The sentences in an e-speranto document are converted to symbolic trees. Beside the transformation of the data structures, the compilation of a "starter script" is carried out. The script comprises trees that need to be transformed, specifications of the modules that are needed to perform the interpretation, as also global variables and processing instructions. The module selection depends mostly on the language of interpretation.

### 6.3. Database

The database is primarily located on the server, but it can also be located on the client. Database serves as a source of language-specific content such as language rules, word and phrase dictionaries.

## 7. CONCLUSION

In this paper we introduced a brief overview of the multilingual machine translation system based on e-speranto. The most important features of the presented system are:

- the use of an interlingua that is compatible with HTML and it can be therefore used for recording multilingual content directly in Web pages;
- the use of an interlingua that is simple enough to be user-friendly, and is complex enough to have enough expressiveness and unambiguity;
- the support for composing documents in e-speranto by the tools developed on the Eclipse platform;
- the implementation of the architecture of the interpreters that allows the reuse of some modules and therefore it reduces the cost of the development of the interpreters into similar languages.

Our further research will be towards the following fields. We intend to define the format and the content of the dictionaries in detail. Moreover, we intend to upgrade the integrated development environment with features that would enable automatic publishing of the composed e-speranto documents on the Web server. We also want to perform a more thorough research in the optimal number of layers in the layered architecture of the interpreters and determine the content that needs to be placed into individual layers. In this context, we will pursue the aim of optimally high factor of module reuse when interpreting into similar languages.

### REFERENCES


[1] Hutchins, W., Somers, H., "An Introduction to Machine Translation", *Academic Press*, New York, 1992.

[2] Schubert, K., "The Architecture of DLT – interlingual or double-dialect", *Floris Publications*, New Directions in Machine Translation, Holland, 1988.

[3] Nyberg, E., Mitamura, T., "The KANT system: Fast, accurate, high-quality translation in practical domains", COLING, 1992.

[4] Uchida, H., et al., "Universal Networking Language: A gift for a millennium", *The United Nations University, Tokyo*, Japan, 1999.

[5] Bugoslavsky, I., "Some controversial Issues of UNL: Linguistic Aspects", Universal Network Language: Advances in Theory and Applications, Research on Computer Science, 2005.
[6] Boitet, C., "GETA's methodology and its current developments", PACLING'97, Meisei University, Ohme, Japan, 1997.
[7] Sérasset, G., Boitet, C., "On UNL as the future "html of the linguistic content" and reuse of existing NLP components in UNL-related applications with the example of a UNL-French deconverter", COLING, 2000.
[8] Blanc E., From the UNL hypergraph to GETA's multilevel tree, MT2000: machine translation and multilingual applications in the new millennium, University of Exeter, British Computer Society, 2000.
[9] Omerović S., Jakus G., Filimonova T., Tomažič S., "Zapis večjezičnih besedil v e-sperantu", *Elektrotehniški vestnik*, Vol. 74, No. 3, 2007, pp. 151-157.
[10] Tomažič S., "Multilingual Web with E-speranto", *IPSI Bgd Internet Research Society*, The IPSI BgD Transactions on Internet Research, Vol. 3, No. 2, July 2007, pp. 13-15.
[11] Jakus G., Omerović S., Filimonova T., Tomažič S., "Classification of the language group characteristics into a multi-layered architecture of the interlingua interpreter in multilingual translation", *Elektrotehniški vestnik*, Vol. 75, No. 5, 2008, pp. 285-292.
[12] e-speranto Web Page, http:// www.e-speranto.org
[13] Eclipse, http://www.eclipse.org

# A User Study of Auditory Versus Visual Interfaces for Desktop Computer

**Sodnik, Jaka and Tomažič, Sašo**

**Abstract—** ***This paper describes a user study on the interaction with auditory interfaces for desktop computers intended to be used by blind computer users. Two new auditory interfaces are proposed, and their effectiveness and efficiency are compared to a standard visual interface. All three interfaces represent a hierarchical menu structure of a simple word processing application. The two auditory interfaces consist of spatialized auditory cues implemented in two different spatial configurations. The AH auditory menu has a ring shaped horizontal configuration of the sound sources, whereas the AV auditory menu has a vertical configuration of the sound sources. The use of spatial sounds enables the simultaneous playback of multiple sounds (i.e. menu commands), which increase the information flow between the user and the application. A pair of headphones is used for the playback of the auditory interfaces, while the 17" computer screen is used for the visual menu. The interaction with the interface is preformed by the QUERTY computer keyboard. The task completion times, interaction efficiency, perceived workload and overall user satisfaction are evaluated. Sixteen test subjects who reported no visual deficiencies were asked to perform five different tasks with two auditory interfaces and a classic visual GUI. The main goal of the experiment was to compare all three interfaces and select the more efficient and better evaluated of the two auditory ones. The experiment proved that both auditory interfaces are effective to use in the interactions with desktop computers; however, they are not faster than the visual interface. The AV auditory interface proved to be more efficient and better evaluated by the users and will therefore be used in further experiments with visually challenged computer users.***

**Index Terms—** ***Human-computer interaction, spatial audio, desktop computer, blind users.***

## 1. Introduction

Nowadays there are over a billion computer users worldwide and the number is still increasing rapidly. Visually impaired people constitute a significant number of computer users which use the computers especially for various word processing applications and web browsing. The majority of existing computer applications for blind computer users is based solely on simple voice synthesizers which read the content of the documents word by word [1][2]. They are sometimes combined with tactile keyboards. We believe all those systems suffer from slowness and low information flow due to the narrow human auditory channel.

We developed a special auditory interface for controlling the desktop application which could also be used by blind users. In this paper we focus primarily on the hierarchical auditory menus for the interface that can be controlled via the computer keyboard. The interface is based on multiple spatial sounds played simultaneously, which enables high information flow between the user and the computer and therefore improves the user interaction with the computer.

The main goal of our user study was to evaluate the effectiveness of such an acoustic interface in the interaction with an application. Three different interfaces - two auditory and one visual - were compared, all of them based on the same hierarchical menu structure. The GUI in the visual interface resembles the interfaces of the typical word processing application such as MS Word or Open Office Writer. The menu items in the two acoustic interfaces are presented with spatial sounds coming from different pre-fixed positions. All three interfaces were evaluated with an extensive user study, resulting in sets of objective measurements and subjective user comments. The main objective of the experiment was to select one of the two proposed acoustic interfaces to be used in the future development of auditory software for blind computer users.

## 2. Related Work

Several researchers focused on spatial auditory interfaces used for browsing or manipulating data. Many different solutions were proposed, each of them with its specific benefits and limitations.

Schmandt described a virtual acoustic environment used for browsing collections of related audio files [3]. The user had the possibility of traveling up and down the virtual hallway by head motion, passing various rooms on the left

Manuscript received April 28, 2009. This work has been supported by the Slovenian Research Agency within the research project: "Assisting visually impaired people in using the computer with the aid of spatial sound".

J. Sodnik is with the Faculty of Electrical Engineering, University of Ljubljana, Slovenia (e-mail: jaka.sodnik@fe.uni-lj.si).

and right side and selecting and playing audio clips.

A surrounding three-dimensional audio interface for various user applications was proposed by Frauenberger and Noistering [4]. The binaural audio rendering implemented directional hearing and room acoustics via headphones. The users could move around freely using a joystick. The system was capable of producing a virtual audio reality that functioned as some sort of a user interface for people with a visual impairment.

A prototype of such an auditory interface was then evaluated with common interactions like tasks with menus, text input and dialogs [5]. The authors concluded that such an auditory interface could be used effectively for presenting various interface components and that there is no significant difference in the user performance between normal sighted and visually impaired users. They also pointed out the importance of the quality of spatial sounds. The real world applications such as MS Explorer also proved to be manageable by auditory display [6].

An example of an auditory interface for a communication device in a vehicle was proposed and evaluated by Sodnik et al. [8]. The interface was based on speech commands positioned in space and manipulated through a custom interaction device.

Mynatt proposed a general methodology for transforming graphical interfaces into non-visual interfaces [7]. Various salient components of graphical interfaces were transformed into auditory interfaces by using the so-called auditory icons. These types of interfaces were designed primarily to meet the needs of blind computer users.

## 3. AUDITORY INTERFACE DESIGN

In our study, we designed a simple word processing application that is be controlled by the proposed auditory interfaces. The hierarchical menu structure of the application is based on MS Word menu structure. The auditory interface enables all the operations needed for handling documents. The menu commands are represented with sound sources (recorded voice commands). At each level of the menu, all available options are presented to the user. After the user selects an option, a new level of the menu is loaded. The options at the basic or main level of the menu are File, Edit, Format, Tools and Help. If File is selected, the options on the next level are New, Open, Save, Print, Page Setup and Exit. When the user selects an option in the basic level and moves to the next level, a gentle music starts playing in the background. Different music is assigned to each branch of the menu (File, Edit, Format, etc.) in order to identify the selected branch. The sound sources at various levels of the menu are organized in two different ways, thus constituting two auditory interfaces.

### *A. AH interface*

In the AH interface, the individual sound sources are placed on a virtual horizontal ring that surrounds the user's head [8]. The spatial angle between two individual sources changes with the number of sources constituting the ring. For example, if there are six simultaneous sources, the angle between the sources is 360° / 6 = 60°, if there are four sources, the angle is 360° / 4 = 90°, etc. The selected source is always directly in front of the user.

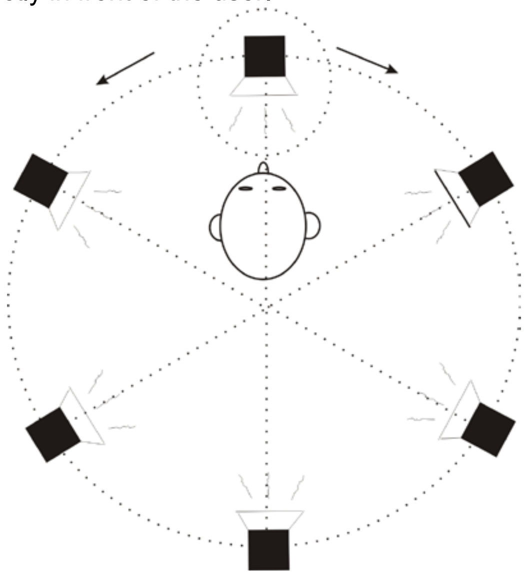

**Fig. 1. The basic principle of AH menu. The individual sound sources are placed on a virtual horizontal ring that surrounds the user's head.**

The ring can be turned in any direction (left or right) in order to select the desired menu option. The user is not located directly in the centre of the virtual ring, but slightly closer to the front. As the front item is the closest to the user, it is therefore also the loudest one.

### *B. AV interface*

In the AV interface, the individual sound sources of a specific level are aligned vertically in front of the user.

The vertical alignment resembles the typical menus used in graphical interfaces. The user can browse between the options by moving up and down. Again, the option or the command directly in front of the user is the selected one. Important features of the AV interface are the pegs at both ends of the menu. When the top or bottom of the menu is reached, the user has to go back in the opposite direction, while in the AH menu the horizontal circle can rotate indefinitely.

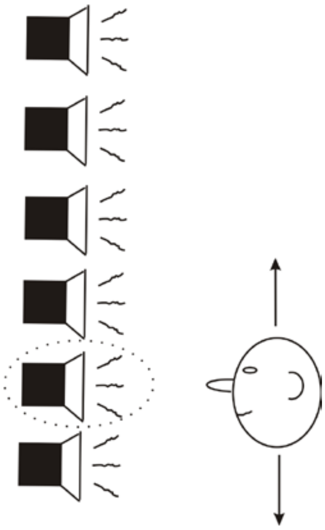

**Fig. 2. The basic principle of AV menu. The individual sound sources of a specific level are aligned vertically in front of the user.**

In the AV interface, the central pitch of the background music changes according to the current position of the user. If the user is located at the bottom of the current menu, the pitch is very low, and if he is located at the top of the menu, the pitch is high.

In the AV menu, three spatial sound sources are played simultaneously, representing the closest three menu options. The user can therefore always hear the selected, the previous and the following option.

### C. *Interaction*

In both interfaces, the interaction is preformed with the use of a QUERTY computer keyboard. In the AH interface, the "LEFT" and "RIGHT" keys are used for turning the circle in the desired direction, whereas in the AV interface, the "UP" and "DOWN" keys are used for moving the user up and down in the menu. The "Q" key is used for confirming the selected menu option and moving to the next level. The "W" key is used for canceling the option and moving one step back in the menu. The "SPACE" key is used to replay the current choice without moving the user up or down in the AV or turning the circle in AH interface.

The interaction in the visual interface is performed with a computer mouse. The menu is shown on a 17" computer screen.

### D. *Sound reproduction*

The sound commands were recorded in a professional recording studio by a female speaker. The sound clips were then sped up to double speed (similar to "spearcons" proposed by Walker [9]) in order to increase the navigation speed. The auditory interfaces were developed in Java programming language with OpenAL sound positioning library. The latter was used to create spatial sounds at various positions by using Creative X-Fi CMSS-3D technology on X-Fi Extreme Gamer sound card. Sennheiser HD 270 headphones with a very good attenuation of ambient noise (from -10 dB to -15 dB) were used for the playback of the auditory menus.

## 4. USER STUDY

A total of 16 test subjects participated in the experiment. They all reported normal sight and hearing. They were asked to perform 5 different tasks:

1. Change language of the document to US English.
2. Change font size to 16.
3. Sort table in the document by field 2.
4. Change font color to blue.
5. Go to Help and activate the product by phone call.

All test subjects performed the tasks with three different interfaces: a visual interface (V), the AH interface and the AV interface – 15 tasks per test subject. The auditory interfaces (AH and AV) were used without any visual support (in both cases the computer screen was blank).

A 5-minute break was assigned after each interface in order for the test subjects to answer some questions on the subjective evaluation of the interface and to complete the Questionnaire for User Interface Satisfaction (i.e. the QUIS test).

In order to eliminate the learning effects, two groups of 8 test subjects were formed. Each group performed the tasks with the interfaces in a different order:

1. group: V, AH, AV.
2. group: V, AV, AH.

Since the same menu structure was used with all three interfaces, the visual interface (V) was used in the experiment primarily to introduce the menu structure to the test subjects. That is why it was used as the first interface in both groups.

The test subjects were asked to complete the tasks as fast as possible. The duration times and the navigation performance were logged automatically.

## 5. RESULTS

In the five tasks preformed by 16 test subjects five variables were evaluated:

- task completion times
- interaction efficiency
- QUIS test
- overall ranking of the interfaces
- subjective comments on the experiment and interfaces

The main results with corresponding interpretations are summarized in the following subchapters. The analysis of variance (ANOVA) is used to establish the significant differences in the results.

## A. Task completion times

The task completion time was measured automatically for all three interfaces. The timer started right after the task was:

- read to the user by the voice synthesis system in the auditory interfaces or
- shown on the screen in the visual interfaces.

The timer stopped when the task was completed. The average task completions times in seconds are shown in Figure 3.

A significant difference between the three interfaces could be confirmed for tasks 1, 2 and 3, with the visual interface always being significantly faster than the two auditory interfaces:

$F_{Task1}(2, 21) = 11.367, p < 0.01$;
$F_{Task2}(2, 21) = 3.464, p = 0.05$;
$F_{Task3}(2, 21) = 6.523, p < 0.01$.

No significant difference could be noted for tasks 4 and 5.

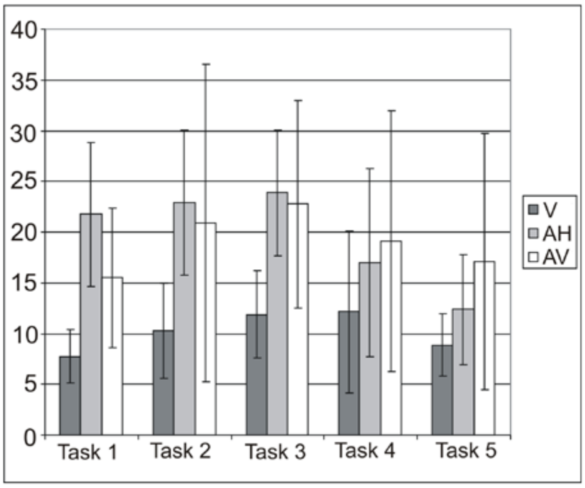


**Fig. 3. The average task completion times for all five tasks shows a significant difference between the three interfaces for tasks 1,2 and 3.**

The shortest task completions times were measured for the visual interface. The confidence intervals for the visual interface were small as well, which signifies small variations of the results. On the other hand, no significant difference could be found between the two auditory interfaces. The latter can be confirmed also by comparing the average completions times of all tasks:

TABLE I
THE AVERAGE TASK COMPLETION TIMES FOR THE THREE INTERFACES SHOW NO SIGNIFICANT DIFFERENCE BETWEEN THE TWO AUDITORY INTERFACES

| Interface | Time / s |
|---|---|
| **V** | 10.2 |
| **AH** | 19.6 |
| **AV** | 19.1 |

We expected the visual interface to be the fastest one, since the majority of test subjects were already skilled in using MS Word GUI for editing documents. We believe that the use of a computer mouse in the visual interface and the use of a keyboard in the auditory interfaces also contributed to the established difference.

## B. Interaction efficiency

The interaction efficiency was measured by counting the movements of the user in the menu (i.e. how many times the user pressed the “LEFT” and “RIGHT” or “UP” and “DOWN” keys) and the selections of various menu items (i.e. how many times the user pressed the selecting buttons). Both parameters were measured automatically by the application itself. The results were then compared with the optimal interaction procedure necessary to complete a specific task.

The interaction efficiency was measured only for the auditory interfaces. Figure 4 shows the average efficiency (the sum of movements and selections) per task:

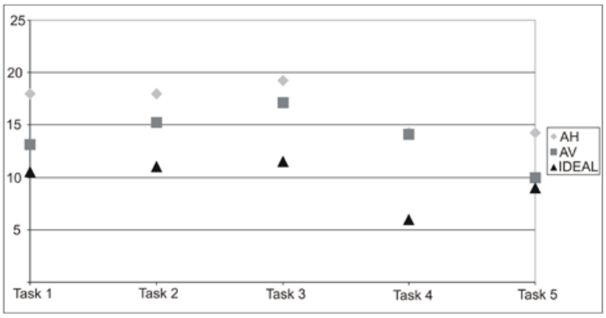


**Fig. 4. Menu navigation performance shows higher effectiveness of AV interface.**

The AV interface proved to be more effective and easier to use since less navigation was required for completing the tasks.

The interaction in both cases proved to be far from ideal. We believe the interaction could be improved significantly with continuous practice and a better knowledge of the menu structure.

## C. QUIS test

The QUIS test was designed to assess the users' subjective satisfaction with specific aspects of the human-computer interface. In our experiment, we intended to measure the reaction of the users to the software used in the experiment. We asked the users to rank the interface on a scale of 1 to 5 (1 being entirely false and 5 being entirely true) based on the following statements about each individual interface:

1. the interface was more wonderful than terrible (Q1)
2. the interface was more easy than difficult (Q2)

3. the interface was more satisfying than frustrating (Q3)
4. the interface was more adequate than inadequate (Q4)
5. the interface was more stimulating than dull (Q5)
6. the interface was more flexible than rigid (Q6)
7. it was easy to learn how to operate the system (Q7)
8. it was easy to explore new features by trial and error (Q8)
9. it was easy to remember names and use commands (Q9).

The significant difference between the interfaces can be established for the questions Q1, Q4, Q5, Q6 and Q7:

$F_{Q1}(2, 21) = 12.014$, $p < 0.01$;
$F_{Q4}(2, 21) = 7.000$, $p < 0.01$;
$F_{Q5}(2, 21) = 5.509$, $p < 0.05$;
$F_{Q6}(2, 21) = 5.619$, $p < 0.05$;
$F_{Q7}(2, 21) = 7.768$, $p < 0.01$.

Figure 5 shows the average QUIS scores of the individual interfaces:

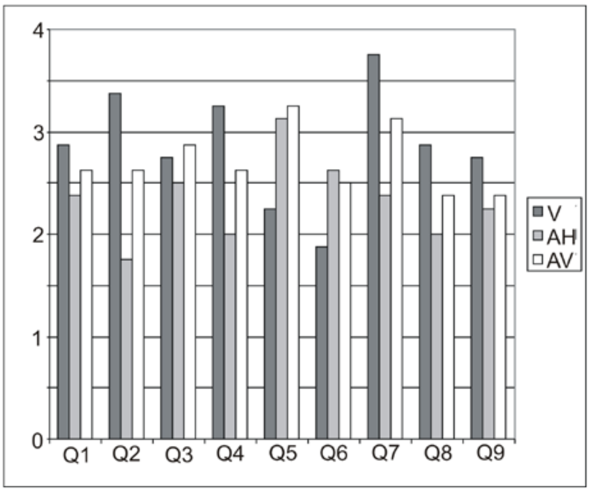


**Fig. 5. QUIS scores for the three interfaces show significant differences between the interfaces for questions Q1, Q4, Q5 Q6 and Q7.**

The visual interface got significantly higher scores in questions Q1, Q4 and Q7. The latter signifies that the test subjects found the visual interface much easier to use than the two auditory interfaces. They also described it as more adequate and stated it was easier to learn the commands and their positions in the menu.

The auditory interfaces scored significantly more points in questions Q5 and Q6. The users described the auditory interfaces as more stimulating and more flexible than the visual one.

No significant difference could be established in questions Q3, Q8 and Q9. It means that all three interfaces seemed to be equally satisfying and equally difficult when exploring new features by trial and error and also for remembering the names and use of commands.

We believe the users found the auditory interfaces more stimulating since a new and interesting prototype (the auditory interface) was compared to a well known and widely used GUI.

### *D. Overall ranking and subjective comments*

We also asked the users to rank the three interfaces with marks from 1 to 3 (1 being the worst interface and 3 being the best one). Table 2 shows the average marks for the interfaces:

TABLE 2
THE AVERAGE MARKS FOR THE THREE INTERFACES SHOW THE SELECTION OF THE AV INTERFACE AS THE BETTER OF THE TWO AUDITORY INTERFACES.

| Interface | Mark |
|---|---|
| **V** | 2.94 |
| **AH** | 1.25 |
| **AV** | 1.81 |

The users selected the AV interface as the better of the two auditory interfaces. We collected some subjective comments justifying their decision:

- the vertical arrangement of the sources in the AV interface seems more natural as it resembles the GUI menus (the vertical columns);
- the peg at both ends of the AV menu increases the overall orientation in the menu since the users always know when they reach the top or the bottom of the individual menu; in the AH menu, the virtual circle cand be rotated indefinitely without a peg at each end of the menu.

Other comments about the auditory interfaces:

- the background music which intends to identify the individual submenu is more disturbing than helpful;
- the spatial sound perception and understanding of the multiple simultaneous sources is better in the AH menu.

## 6. DISCUSSION

The main goal of this study was the evaluation of an acoustic interface as a substitute for the traditional visual interface (V) for a simple word processing application. The three main variables measured in the experiment were task completion time, navigation performance and overall user satisfaction. Our main goal was to compare the two auditory interfaces (AH and AV) and select one of them for our future research.

Both auditory interfaces proved to be significantly slower than the visual interface, which was an expected outcome. Visual perception is much faster than the auditory one. In addition, the resolution and the capacity of the visual channel are much better than those of the auditory channel. On the other hand, we could not confirm any significant difference in the task completion times between the two auditory interfaces.

The second parameter we evaluated in the study was the navigation performance of the users when solving tasks. By navigation performance we mean the overall orientation of the test subjects in the menu, which can be measured by following the movements of the user in the menu and the movements of the menu itself (turns, movements, clicks). The average navigation performances of the two auditory interfaces were measured and then compared to the ideal task completion navigation. According to the results, the AV interface enabled better navigation than the AH interface. We find the result somewhat surprising since we expected the AH interface to more efficient. It has already been established in [10] that the horizontally distributed spatial sounds can be localized much more efficiently and faster than the vertically distributed sound sources. It appears that the test subjects concentrated primarily on just one selected sound source and not so much on the two additional simultaneous sources.

The comparison of the subjective evaluations also shows a preference of the test subjects for the AV interface. Most of the users confirmed the AV interface to be more efficient and user friendly. The latter is confirmed also by a significantly higher overall rank.

Taking into consideration the results discussed above, the AV interfaces proved to be more efficient and preferred by the majority of the users.

## 7. CONCLUSION

With the designed interfaces we intend to improve the use of various desktop applications for the visually impaired users. We explored the use of spatial sound in order to increase the auditory information flow between the computer and the user. A simple interaction based on only two keys for controlling the position in the menu and two additional keys for confirming or canceling various options should effectively replace the mouse navigation in graphical interfaces.

The most important results of the evaluation show that both proposed auditory interfaces can be used effectively for manipulating documents in word processing applications, but one of them is slightly more efficient and also preferred by the users. The latter will therefore be included in our future research. We intend to develop a 3D auditory computer desktop to be used by blind computer users. The auditory desktop will support free space navigation and selection of various applications, similar to the movements of the mouse in the graphical interface. Spatial sound will play an important role for user orientation and accurate perception of the desktop structure.

## REFERENCES


[1] American Foundation for the Blind, Word Processing, http://www.afb.org/ProdBrowseTaskResults.asp?TaskID=425&SpecID=26

[2] TSBVI, Adaptive Technology, http://www.tsbvi.edu/technology/overview.htm

[3] C. Schmandt, "Audio Hallway: a Virtual Acoustic Environment for Browsing," ACM Symposium on User Interface Software and Technology, 1998, pp. 163-170.

[4] C. Frauenberger, M. Noistering, "3D Audio Interfaces for the Blind," Proc. of the Int. Conf. on Auditory Display, Boston, MA, USA, 2003, pp. 281-283.

[5] C. Frauenberger, V. Putz, R. Holdrich, "Spatial Auditory Displays: A Study on the Use of Virtual Audio Environment as Interfaces for User with Visual Disabilities," Proc. of the 7th Int. Conf. on Digital Audio Effects (DAFx'04), Naples, Italy, 2004, pp. 384-389.

[6] C. Frauenberger, V. Putz, R. Holdrich, T. Stockman, "Interaction Patterns for Auditory User Interfaces," ICAD Proc., 2005, pp. 154-160.

[7] E. D. Mynatt, "Transforming Graphical Interfaces into Auditory Interfaces," Doctoral dissertation, Georgia Institute of Technology, 1995.

[8] J. Sodnik, C. Dicke, S. Tomažič and M. Billinghurst, "A user study of auditory versus visual interfaces for use while driving," Int. Journal of Human-Comput. Stud., vol. 66, no. 5, 2008, pp. 318-332.

[9] B. N. Walker, A. Nance and J. Lindsay, "Spearcons: Speech-Based Earcons Improve Navigation Performance in Auditory Menus," Proc. of the Int. Conf. on Auditory Display (ICAD 2006), London, England, 2006, pp. 63-68.

[10] J. Sodnik, R. Sušnik, M. Štular, S. Tomažič, "Spatial sound resolution of an interpolated HRIR library," Applied Acoustics, vol. 66, no. 11. pp. 2005, pp. 1219-1234.G. O. Young, "Synthetic structure of industrial plastics (Book style with paper title and editor)," in *Plastics*, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.

# Efficient User Data Synchronization

Stančin, Sara; Đorđević, Nemanja; and Tomažič, Sašo


## Abstract

User data synchronization feature of computing user devices presents an essential role in the vision of pervasive networks. Typical user environment includes several devices - a work desktop computer, home laptop, and mobile device for example. Through each of these devices a user can access different data records. The aim of synchronization is to enable user access to a particular data record through a device other than the one on which the record was whether created whether last modified. Users can read, change, or delete a specific data record by accessing any of its available replications. Such independent access can embrace any type of file or web content in applications like email, file content distribution, and personal information manager.

The synchronization environment can include numerous devices with different processing, memory and connection capabilities. Further on, devices involved operate in two possible modes, the online and the offline. As long as a device is not connected to a particular network new data record modifications are not consistent with corresponding data records on other devices in that network. As soon as a device connects to a network, the synchronization process needs to determine and reconcile possible differences between documents replicated on multiple devices. Due to multi device involvement in data updating, different synchronization conflicts may occur. Synchronization protocols in force implement different, more or less complex mechanisms which dictate system behaviour if synchronization conflicts arise. The efficiency of these mechanisms varies.

A new paradigm for achieving efficient user data synchronization is presented. Documents are associated with globally unique identification marks while document replication validity is achieved using timestamp values. The proposed paradigm enables reliable synchronization of user data, while the synchronization process remains simple and efficient from the computational point of view. The synchronization procedure according to the new paradigm is efficient even in a dynamic environment which involves numerous devices.



### KEYWORDS

User data synchronization, Data reconciliation, Unique global document identification


## 1. INTRODUCTION

### *1.1. Fundamentals of user data synchronization*

Users access their data through different computing devices like mobile telephones, laptops, desktops, and handheld computers. Each single user device stores a specific user data subset. Depending on the used device, users can therefore access different subsets of data. Recently, more and more of these devices, belonging to the same user, communicate and collaborate among themselves in such information delivery. Access to user data that has previously been enabled through one particular user device is being enabled through other user devices as well. Users can read, change, or delete a specific data record by accessing any of its available replications. Such independent access can embrace any type of file or web content in applications like email, file content distribution, and personal information manager.

Multiple copies of data records that represent the same data on different devices must be the same in content. This leads us to the problem of user data record synchronization. The synchronization process needs to determine the differences between user data replicated on multiple devices and thereafter reconcile these differences as illustrated in Figure 1.1.

Manuscript received May, 11, 2009.
Sara Stančin is with the University of Lubljana, Slovenia
Sašo Tomažič is with the University of Ljubljana, Slovenia
Nemanja Đorđević is with the University of Belgrade, Serbia

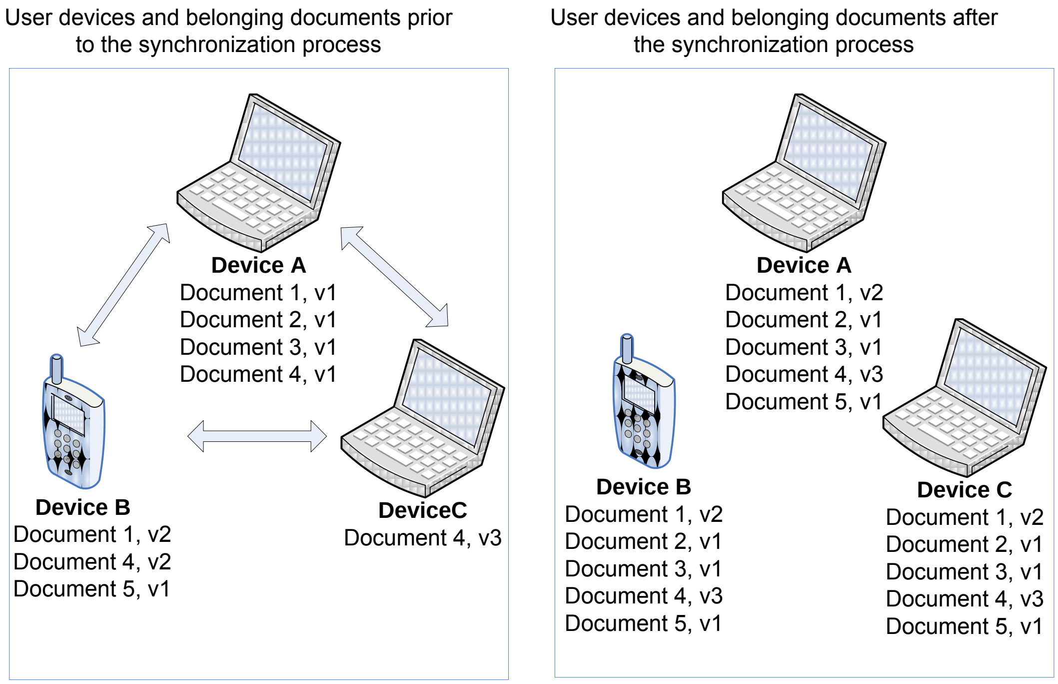


**Figure 1.1: User data synchronization aim and purpose. Before the synchronization process each devices stores a different subset of user documents while after the synchronization process all documents in the user network are stored on all of the involved devices and have equal content.**

The synchronization environment usually includes numerous devices like powerful desktop machines (computationally very fast and efficient personal computers equipped with sufficient memory and relatively fast network connections) as well as mobile devices with limited capabilities and resources. Limited processing capabilities and memory resources of user devices can obstruct extensive synchronization computation and data replication. Moreover, due to numerous devices involved, which work in both, the offline and the online mode, different synchronization conflicts may occur as shown in Figure 1.2. These problems lead to data inconsistencies and can make the realization of a reliable data synchronization protocol a complex problem to solve. Synchronization protocols in force implement different, more or less complex mechanisms which dictate system behavior if synchronization conflicts arise. The efficiency of these mechanisms and procedures varies. It is not unusual that after synchronization is achieved users notice different faults and unexpected results. Synchronization faults embrace problems like data loss, data record duplication, and invalid data record content.

## 1.2. Article orientation

Our aim is to present a new paradigm that embraces possible user data synchronization improvements. First we give a presentation of the synchronization process. We give special attention to principles which determine how data reconciliation is achieved and how conflicts are resolved. Further on we present the introduction of unique global document identification for reliable user data synchronization. Different replications of a particular document have all the same unique global identification mark. During the synchronization process, valid replications are determined according to their timestamp value. Using timestamps and global identification at the same time, simple and exclusive reconciliation procedures are enabled. The presented solution is efficient and reliable even in the environment which involves numerous user devices.

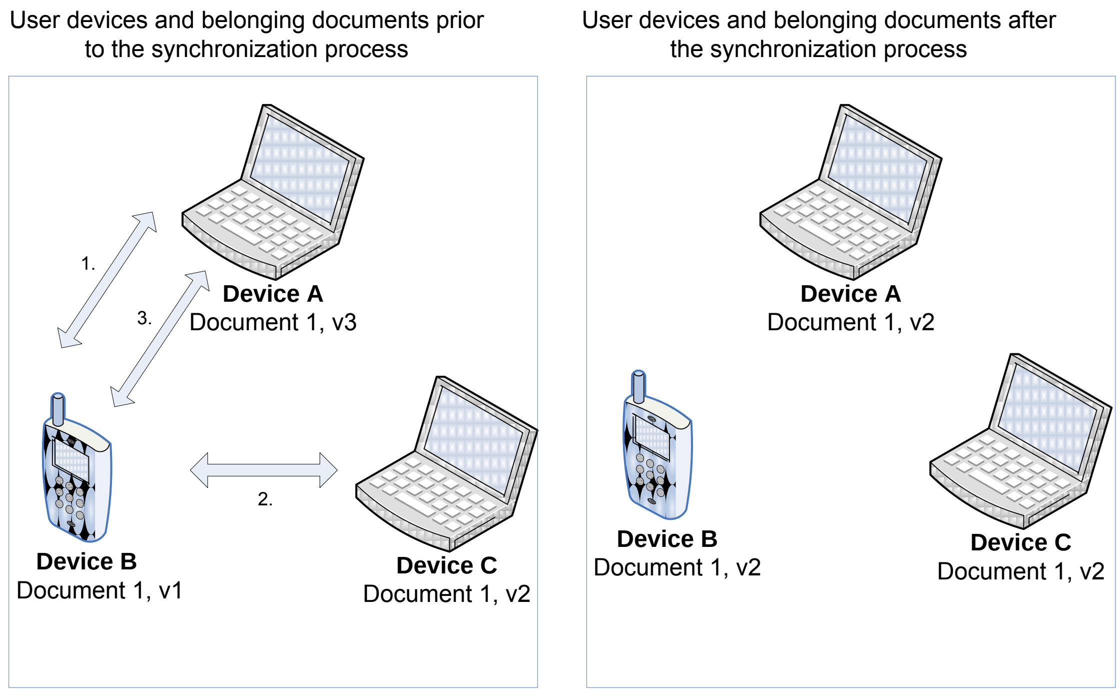


**Figure 2.2: User data synchronization error due to conflict occurrence. In the presented example three devices synchronize among themselves, where Devices A and C replications of Document 1 have been changed since last synchronization was performed. These changes occurred is such time sequence, that Document 1 v3 is the newest version and should therefore be propagated to other devices included in the network (Devices B and C). The result of the first synchronization is unambiguous as only one replication (Document 1 v3) is changed and established to be valid. As no conflicts arise, Devices B version of Document 1 is replaced with Devices A version of the same document. During the second synchronization process however, the two cooperating devices, Devices C and B can not establish which version of Document 1 is to be treated valid as both of the replications have been changed since the last synchronization between these devices. In the presented case, such synchronization mode was chosen, that in the case of conflicts, Devices B replications are replaced by replications on Device C. If now Devices B and A should again reconcile their documents sets, a synchronization error would occur as devices A replication content is replaced with the invalid content originated on Device C.**

## 2. USER DATA SYNCHRONIZATION

Synchronization protocols dictate the rules for devices communication, user data distribution and reconciliation, and conflict resolution. Implemented rules affect protocol performances - network scalability, time, memory and computation efficiency, their openness and wideness of applicability. We focus of synchronization mechanisms that enable change detection and reconciliation and conflict resolution.

### *2.1. Detection of data record differences*

Synchronization protocols relay on special mechanisms for establishing which data records have been changed since the last synchronization. These include status flags, timestamps and algebraic set reconciliation. Maintaining change logs on both the server and the client side is also possible but is rarely sufficient. When no change detection mechanisms are considered all applications data records get overridden by records from another application data set. Having in mind possible extensiveness of data set, the realization of such a concept is usually inefficient from the bandwidth usage, latency and energy consumption point of view.

Supporting status flags is the most

common mechanism that different protocols implement. These flags are used to indicate if a special data record has been modified, inserted or deleted since the last synchronization has been accomplished. Usage of single status flags considers maintenance of one flag for each record in an application's data set. In general, upon a synchronization request all data records whose status flags have been changed get transmitted from the client to the server device. The server then establishes what the proper synchronization result is and sends its corrections back to the client. While computation, memory, data transmission load and central point of failure are shown as strengths of the single status flag network size is considered to be a problem. This particular usage of status flags limits the synchronization process in a way that a device can synchronize with a device with which it synchronized last.

Usage of multiple status flags per record supports multi device synchronization. For each data record as many flags are maintained as is the number of devices in the user network with which that device synchronizes – a flag per record with reference to each device. Devices can therefore synchronize with different devices each time but the synchronization procedure still does not scale with network size due to time and memory problems of maintaining a vast number of flags.

Change detection can also be achieved using timestamp record denotation. Synchronization is then achieved by transferring all data records marked with time later than the time when previous synchronization was accomplished. Some argue the efficiency of timestamps and the protocols that implement this mechanism as they [1] "poorly adapt to dynamic situations where devices are frequently added or removed from the network".

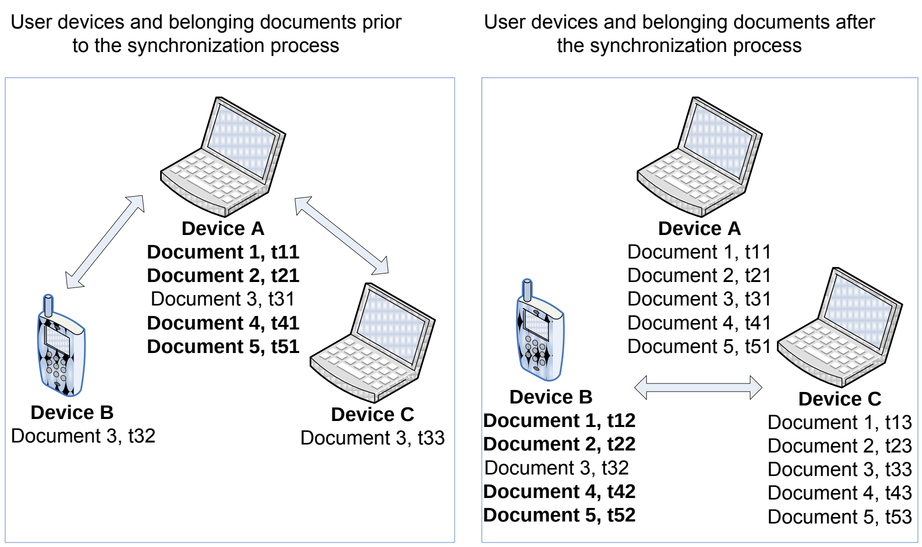


**Figure 2.1: Inefficient communication cost when synchronizing user data differences established according to the timestamp and status flags mechanism. In the presented example, Devices B and C one after another synchronize with Device A. The sequence of these two synchronizations is not important. Documents 1, 2, 4, and 5 are transmitted from Device A to the other two devices and so all three store documents which are equal in content. Let us now consider the scenario where Device B synchronizes with Device C. If timestamps are used, Documents 1, 2, 3, and 4, stored on Device B are marked with a timestamp value later then the last synchronization process between the two devices. If multiple status flags are used, these documents are also marked as changed. In either way Documents 1, 2, 4, and 5 would be again transmitted, regardless of the fact that there are no differences between the content of these two sets of documents.**

Synchronization schemes can be improved using more sophisticated computational methods [3]. Algebraic set reconciliation considers data to be represented by sets. When a pair of devices exchanges information, the devices must reconcile their respective data sets without the a priori knowledge of which data records need to be transmitted. The gain of algebraic set reconciliation algorithms typically manifests when the number of differences to be reconciled is far smaller than the size of the reconciling databases. Intensive computation provides a replacement for huge amounts of data transmitted over the network. Algebraic set reconciliation protocols do not need to maintain any state information about other devices in the network, but they are computationally intensive.

### *2.2. Conflict resolution*

Specific nodes in a network have to be able to efficiently identify data record changes and propagate updates to other devices present in the network. The problems arise when a specific data record has been inconsistently changed on both cooperating sides, the one acting as a client and on the one acting as a server. These situations represent synchronization conflicts. Resolving conflict depends on the server side synchronization algorithm. A synchronization algorithm my simply decide which changes to be treated as the valid ones: client changes always win, server changes always wins or, in the example of timestamps, latest change always wins. Some synchronization algorithms are less exclusive and take into consideration merging data records and data duplication.

## 3. NEW SYNCHRONIZATION PARADIGM PROPOSITION

### *3.1. New paradigm basis*

We propose a new paradigm for achieving efficient user data synchronization from both, the time and the processing consumption point of view. Efficient synchronization is achieved by associating documents with globally unique identification marks and with document replications differentiation referring to their timestamp values according to the following:

- In the network of user devices, only one specific document exists at a specific moment of time. A unique global document identification (GDID) value is associated with each document. Different devices obtain different replications of a particular document and so multiple replications exist. Each of these replications has the same identification mark and when synchronized, these replications contain the same user content.

- For GDID determination we consider the usage of cryptographic hash functions. The determination of the GDID value is extremely important for keeping data sets in a consistent state. Each user document must have a different GDID value. The GDID value should be a value with characteristics of a random value, long enough so we can presume that the possibility of a collision would be contentedly low. A cryptographic hash system that works with one billion 160-bit hashes has an approximate $10^{\wedge(-32)}$ collision probability.

- Each replication of a document has its own timestamp mark. Same timestamp values of two documents replications imply that they are synchronized. If the values are different, the replications are different in content and must be reconciled during the next synchronization.

- The timestamp value also enables conflict resolution. If during the synchronization process, two involved replications of a particular document (recognized by the same GDID) have different timestamp values, one that has the bigger timestamp value, we can also say - the "younger" replication - is considered to be the valid replication of that particular document. The synchronization protocol must then provide for the according modification of the invalid replication.

- A document can be hard deleted or soft deleted. When a document has been soft deleted it implies that only the record content and not the record itself have been deleted. This kind of deletion is actually treated equally as data content modification. After deletion, the data record itself remains stored in device memory and it retains its GDID. The timestamp value belonging to that, content empty, replication is updated, at the same time it does not needlessly occupy valuable device memory space. For achieving efficient synchronization of data replications that should remain deleted, each deleted record can be marked with an additional element.

- After it is achieved that two documents replications are the same in content, they must have the same timestamp value. The timestamp value of the "older" replication is overridden by the "younger" replications value.

## *3.2. The synchronization process*

The initial synchronization agreement includes the exchange of device capability information in order for devices to complete the client-server agreement. Each device maintains a list of timestamp values which represent the time of the last synchronization with each other user device in the network. Upon the initial synchronization agreement, either one either both devices, depending on the synchronization mode, create a list of data records. Data records are represented with their GDID and timestamp values. The list is created in such a way that a device compares records time-stamp values with timestamp values of the last corresponding synchronization. Data records with timestamp values smaller, or "younger", then the corresponding time-stamp value are considered to have been modified since the last synchronization and are inserted into the list. If the last synchronization timestamp is for some reason unavailable, these lists comprise GDID and last modification timestamp values of all records in a particular data set. The client device then sends this list to the server device. After receiving such a list, the server side makes a comparison of the received client list and its own list upon what it can establish replication validity. The server then requests that appropriate reconciliation is performed. The synchronization process can finish when timestamps are adjusted and all of the performed changes are reported to the server.

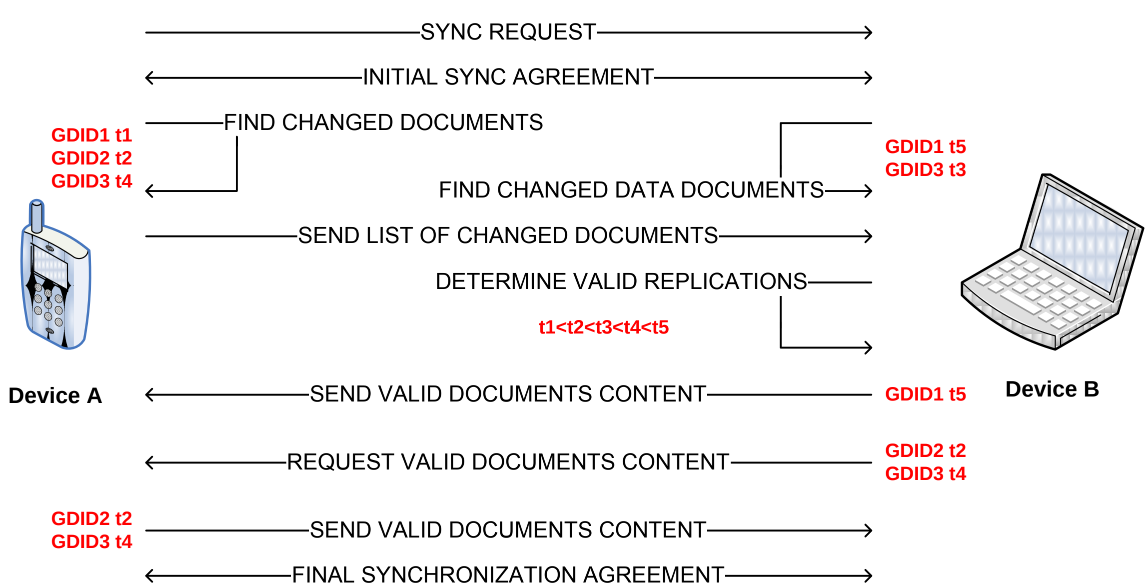


**Figure 4.1: The synchronization process according to the new paradigm proposal. After Device A creates the list of changed documents, it sends it to Device B, which is in this case acting as server. After receiving such a list, Device B can compare Devices A list with its own and determined valid replications can be exchanged between the two devices.**

### *3.3. The relation between the new synchronization paradigm and SyncML*

For the proposed paradigm implementation purposes we suggest the usage of the SyncML standard. SyncML is an industry initiative which seeks to provide an open standard for data synchronization across different platforms. This standard assumes that each client device maintains status flags for each of its records with respect to every other device on the network. Therefore, multi device synchronization is supported. SyncML synchronization enables devices to synchronize with a different device each time. More extensive reports on SyncML can be found in [1], [3], [4], and [5]. The SyncML synchronization process can be extended by defining new data elements which enables timestamps and GDID inclusion. The proposed paradigm is SyncML supplementary and is not in confrontation with any of its goals. At the same time, including the proposed mechanisms, reported SyncML weaknesses regarding mapping tables and timestamps can be avoided.

The globally unique identification concept opposes to the identification concept, with which the SyncML initiative is familiar with. SyncML uses two identifiers, the LUID (*Local Unique ID*) and the GUID (*Globally Unique ID*). LUID identifiers are associated with client data records, while the GUID identifiers represent the corresponding data on the server side. The LUID is always assigned by the client device. As the LUID and GUID of a particular data record can be different, each server has to maintain special mapping tables – one for each of the involved clients. The maintenance of mapping tables is therefore of crucial importance for achieving efficient data synchronization. As the size of these mapping tables grows linearly with the number of data records, these tables can become significant in their extent. Moreover, different SyncML clients can assign different LUID identifiers for the same data record and different clients can assign same LUID identifier for different data records. Inadequate and faulty mapping table maintenance can consequently lead to user data inconsistency, duplication and loss of content.

The proposed usage of timestamps also opposes to the usage with which SyncML is familiar with. When more than two devices are synchronizing, the use of timestamps results in an inefficient communication cost. For example let us consider the synchronization environment which includes three devices: Device A, B, and C as illustrated in Figure 4.2. Further on, let us consider that since the last synchronization was performed, Documents 1, 2, 4, and 5 were created on device A. The problem arises if Devices B and C try to synchronize with each other after they have successively performed synchronization with device A. In such an example, modification records require transmission of eight differences, whereas, in fact, there are none. The proposed representation of document replications with GDID and timestamp values, this inefficiency is avoided.

## 5. CONCLUSION

The synchronization protocol must support reliable reconciliation of replicated user data which is a demanding task as from the computation as from the communication point of view considering that the synchronization environment can include devices with different capabilities and limited network resources. Synchronization protocols rely on special mechanisms for establishing which data records have been changed since the last synchronization. These include status flags, timestamps and algebraic set reconciliation. Supporting status flags is the most common mechanism that different protocols implement. While computation, memory, data transmission load and central point of failure are shown as strengths of the single status flag network size is considered to be a problem. This particular usage of status flags limits the synchronization process in a way that a device can reconcile its records with records stored on a device with which it synchronized last.

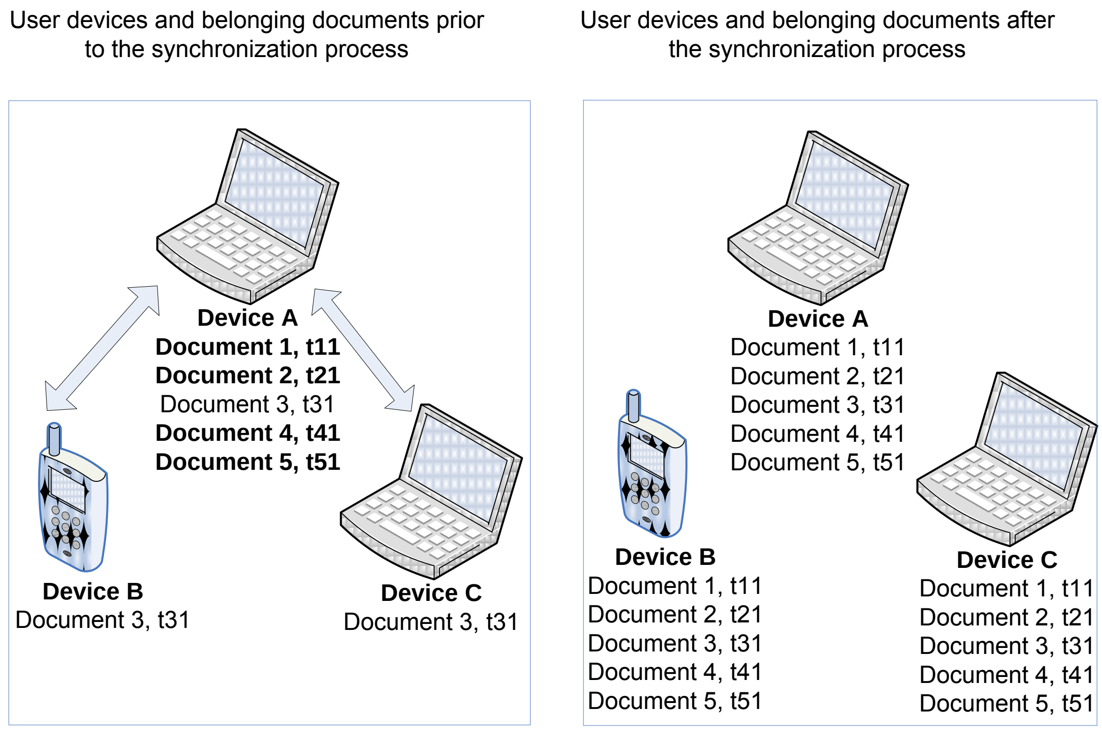


**Figure 4.2: Efficiency of timestamp usage according to the proposed paradigm: after Devices B and C successively synchronize with Device A, all replications presenting the same document on different devices are marked with the same timestamp value. If now, Devices B and C would try to synchronize, no documents would be transmitted over the network as there are no differences between these two data sets and the problem of inefficient communication cost is avoided.**

Usage of multiple status flags enables devices to synchronize with a different device each time but the synchronization procedure still does not scale with network size due to time and memory problems of maintaining a vast number of flags. Algebraic set reconciliation protocols do not need to maintain any state information about other devices in the network, but they are computationally intensive.

We presented a new synchronization paradigm which enables efficient synchronization even when connection bandwidth, user device processing capabilities and memory resources are very limited. According to the new paradigm, different replications of a particular document have all the same unique global identification mark. During the synchronization process, valid replications are determined according to their timestamp value, which presents the time of the last document modification. Using timestamps and global identification at the same time, the proposed paradigm enables simple and exclusive procedures for reconciliation of different replications of one document. For the proposed paradigm implementation purposes we suggest the usage of the open standard SyncML. Including the proposed GDID for document identification and timestamps for replication differentiation, reported SyncML weaknesses can be avoided. If the GDID is really unique, no mapping tables are required and the devices need not maintain modification information about each of its records with respect to each other device. Consequently, the memory problem is avoided and the protocol can be applicable to larger networks as well. At the same time, the protocol remains simple and efficient from the computation point of view.

## REFERENCES


[1] Argawal, S. & Starobinski, A. & Trachtenberg, A. (2002) "On the Scalability of Data Synchronization Protocols for PDAs and Mobile Devices", *IEEE Network*, Vol.16, p22-28.

[2] Minski, Y. & Trachtenberg, A. & Zippel, R. (2003) "*Set Reconciliation with Nearly Optimal Communication Complexity*", *IEEE Transactions on Information Theory*, Vol. 49, Issue 9, p2213-2233.

[3] Starobinski, D. & Trachtenberg, A. & Argawal, S. (2003) "Efficient Data synchronization", *IEEE Transactions on Mobile Computing*, Vol. 2, No. 1, p40-51.

[4] Jönsson, A. & Novak, L. (2001) "SyncML-Getting the mobile Internet in sync", *Ericsson Review,* No.3

[5] Open Mobile Alliance, "SynML specifications", *http://www.openmobilealliance.org/tech/affiliates/syncml/syncmlindex.html.*

[6] Intellisync web resources *http://www.intellisync.com*

[7] Palm HotSync web resources http://www.palm.com/uk/en/software

# Ontologies and Knowledge Representation for Information Retrieval from Knowledge Databases and the Internet

**Omerović, Sanida; Tomažič, Sašo; Radenković, Božidar; Milutinović, Veljko**

## ABSTRACT

*Recording knowledge in a common framework that would make it possible to seamlessly share global knowledge remains a central challenge for researchers. This annotated survey of the literature examines ideas about concept representation that address this challenge.*

*General Terms: Concepts, Knowledge, Meaning, Modeling, Ontology, Semantics*
*Additional Key Words and Phrases: data, relations, representation*

## 1. INTRODUCTION

The information world that we live in today presents us with a vast amount of data stored separately in books, newspapers, radio, TV, Internet, etc., all of them increasingly digitized. Moreover, there is an exponential increase in these data day to day so that the ability of an average computer-educated person to find a specific data element or subject-related useful piece of information is decreasing rapidly.

## 2. ONTOLOGY

The term ontology is taken from philosophy, where it means the study of being or existence ("What exists?", "What is?", "What am I?"). Questions about being exemplify and highlight the most basic problems in ontology: how to find a subject, a relationship, and an object to talk about. Within the more limited scope of the works cited in this paper, an ontology is a concept that groups together other (in some sense "like") concepts as shown in Figure 1. This grouping of concepts is brought under a common specification in order to facilitate knowledge sharing.

### *2.1 Ontology definition*

Even within the limited scope of information sharing, the term ontology has been defined from many different view points and with different degrees of formality. Ontologies mostly include metadata such as concepts, relations, axioms, instances, etc [NAVIGLI 2003]. Ontologies can be viewed as mediators in the acquisition of knowledge from concepts. Therefore, ontologies lie between the concepts (that they subsume) on one hand and the overarching knowledge domain (within which they are embedded) on the other. Here are some ontology definitions from the viewpoint of concepts:

- A "specification of a shared conceptualization" [GRUBER 1993].
- An arrangement of concepts that represents a view of the world that can be used to structure information [CHAFFEE 2000].
- A conceptual model shared between autonomous agents in a specific domain [MOTIK 2002].

And here are some ontology definitions from the viewpoint of knowledge:

- An organized enumeration of all entities of which a knowledge-relation system is aware [HALLADAY 2004].
- A description of the most useful, or at least most well-trodden, organization of knowledge in a given domain [CHAN 2004].

Given ontology definitions from both of these concept and knowledge perspectives, the next issue is how ontologies can be organized.

### *2.2 Ontology organization*

Modeling is used to achieve a consistent organization among and within ontologies; moreover finding (or inventing) consistent descriptive metadata for ontology modeling purposes is cited as the main obstacle to the introduction of ontology-based knowledge management applications into commercial environments [WARREN 2006]. Various approaches are suggested to address these problems.

Created from subsumed components, ontologies can unite classes, relationships, and entities that are equivalent but differently expressed. Ontologies themselves can be combined to model a certain knowledge domain [CHAN 2004] by organizing them as a set of terms and constraints in some form of ontology vocabulary. Alternatively, [MOTIK 2002] presents the organization of an ontology as a set consisting of: a relation, a sub-concept, an instance, a property, and a concept. In order to achieve interoperability between information presented in different ontologies, an application can consolidate ontologies into one, through a process of ontology mapping (Figure 1).

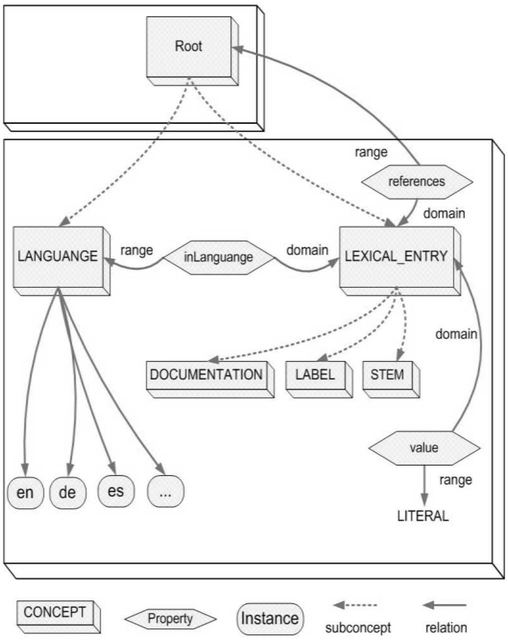


**Figure 1. A Lexical Ontology-Instance-Model Structure. Each instance of a ROOT concept may have a lexical entry which reflects various lexical properties of an ontology entity, such as a label, stem, or textual documentation. Before interpreting a model, the interpreter must filter out a particular view of the model (whether a particular model can be observed as a concept, a property, or an instance) – it is not possible to consider multiple interpretations simultaneously. However, it is possible to move from one interpretation to another - if something is viewed as a concept in one case, in another case it is possible to interpret the same thing as an instance. *This is similar to the vocabulary switching process proposed in [SCHATZ 1997].***

Ontologies can be organized as a set of hypercubes, where each hypercube represents a single concept [SCHLOOSSER 2002]. A hypercube composed of peers supporting a ***TravelService*** concept, in a peer to peer network, is presented in Figure 2.

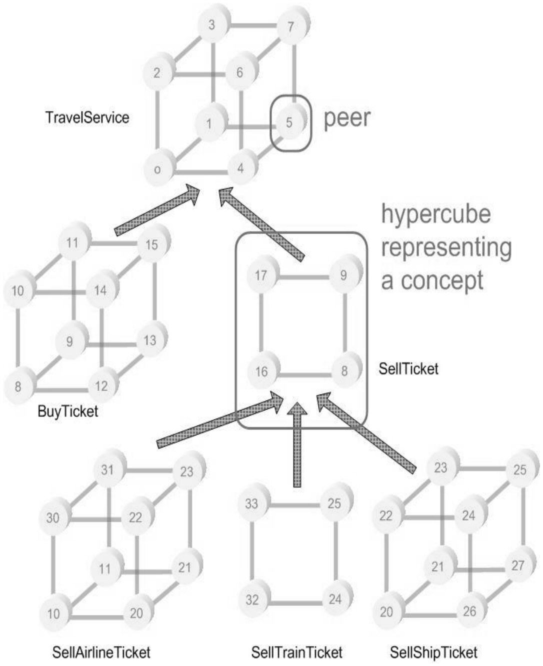


**Figure 2. A peer to peer (P2P) ontology-structured network topology representing the process of buying and selling airline, train, and ship tickets. The internal peer organization of a hypercube is structured so that the network can support queries that are logical combinations of ontology and concepts . Every peer should be able to become a root of a tree spanning all nodes in the network. Also, to maintain the network symmetry, crucial for P2P networks, any node in the network should be allowed to accept and integrate new nodes in the network. Querying the network works in two routing steps. The first step is to propagate a query to those concept clusters that contain peers that the query is aiming at. In the second step, a broadcast is carried out within each one of these concept clusters, optimally forwarding the query to all peers in the clusters. This involves shortest-path routing as well as restricted broadcast in the concept coordinate system.**

Some software tools for grouping and organizing ontologies are:

- Ontology library systems [DING 2001],
- Automatic ontology derivation [GAUCH 2002] from hierarchical collections of documents like Open Directory Project [OPEN 2002],
- Protégé [PROTEGE 2006],
- KAON [KAON 2001], etc.

One sees that the primary purpose of all of this effort, with an eye towards enterprise applications, is to so define and organize ontologies as to facilitate information sharing among originally incompatible data elements – possibly with the assistance of software tools that have been developed to automate this effort. This leads directly to a consideration of the ways that ontologies so constructed are used in applications to support information sharing.

```
<?xml version="1.0" ?>
<!DOCTYPE rdf:RDF (View Source for full doctype...)>
- <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
xmlns:owl="http://www.w3.org/2002/07/owl#" xmlns:address="http://
daml.umbc.edu/ontologies/ittalks/address#" xmlns="http://
daml.umbc.edu/ontologies/ittalks/address#" xml:base="http://
daml.umbc.edu/ontologies/ittalks/address">
- <owl:Ontology rdf:about="">
<rdfs:comment>This file describe the postal address. We are more
consider about the delivery aspect rather than accurate geographical
localtion, such as latitude and longitude.It is created by Li Ding -- http://
www.csee.umbc.edu/~dingli1/, Harry Chen -- http://www.csee.umbc.edu/
~hchen4/, Lalana Kagal -- http://www.cs.umbc.edu/~lkagal1/, Tim Finin -
- http://www.csee.umbc.edu/~finin/.</rdfs:comment>
<owl:versionInfo>$Revision: 1.1 $</owl:versionInfo>
</owl:Ontology>
- <owl:Class rdf:ID="Address">
<rdfs:label>Address</rdfs:label>
<rdfs:comment>Address</rdfs:comment>
</owl:Class>
- <owl:DatatypeProperty rdf:ID="roomNumber">
<rdfs:domain rdf:resource="#Address" />
<rdfs:range rdf:resource="http://www.w3.org/2001/
XMLSchema#string" />
</owl:DatatypeProperty>
- <owl:DatatypeProperty rdf:ID="streetAddress">
<rdfs:domain rdf:resource="#Address" />
<rdfs:range rdf:resource="http://www.w3.org/2001/
XMLSchema#string" />
</owl:DatatypeProperty>
- <owl:DatatypeProperty rdf:ID="city">
<rdfs:domain rdf:resource="#Address" />
<rdfs:range rdf:resource="http://www.w3.org/2001/
XMLSchema#string" />
</owl:DatatypeProperty>
- <owl:DatatypeProperty rdf:ID="state">
<rdfs:domain rdf:resource="#Address" />
<rdfs:range rdf:resource="http://www.w3.org/2001/
XMLSchema#string" />
</owl:DatatypeProperty>
- <owl:DatatypeProperty rdf:ID="zip">
<rdfs:domain rdf:resource="#Address" />
<rdfs:range rdf:resource="http://www.w3.org/2001/
XMLSchema#string" />
</owl:DatatypeProperty>
- <owl:DatatypeProperty rdf:ID="country">
<rdfs:domain rdf:resource="#Address" />
<rdfs:range rdf:resource="http://www.w3.org/2001/
XMLSchema#string" />
</owl:DatatypeProperty>
</rdf:RDF>
```

**Figure 3. DAML ontology for the concept of *address* in OWL [OWL04]. The concept *address* is observed as a class, with the following subclasses: roomNumber, streetAddress, city, state, zip, and country.**

## *2.3 Ontology use*

Ontologies cover a broad range of knowledge. They are variously presented in this section as applied to information systems, software agents, automatic translation process, photo descriptors, and text mining. Many different applications use ontologies to explicitly declare the knowledge embedded in them [PEREZ 2002]. As an illustration, a DARPA Agent Markup Language DAML [DAML 2007] ontology record for the concept of ***address*** is presented in Figure 3.

### *2.3.1. Information systems*

Ontologies can serve as a basis for an information system as is the case in Ontology-Driven Information Systems (ODIS) [GUARIN 1998]. ODIS system illustrates how four distinct general ontology types can be involved in building an information system in terms of Top-level, Domain, Task and Application ontologies. This unified hierarchical organization is presented in Figure 4.

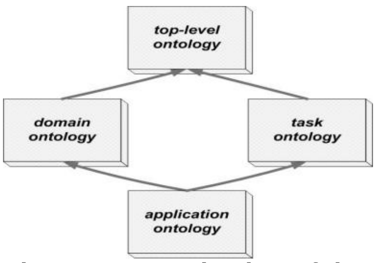


**Figure 4. Organization of the ontologies in the Ontology-Driven Information System. Top-level ontologies describe very general concepts like space, time, matter, object, event, action etc., which are independent of a particular problem or domain. Domain ontologies and task ontologies describe, respectively, the vocabulary related to a generic domain (like medicine, or automobiles) or a generic task or activity (like diagnosing or selling), by specializing the terms introduced in the top-level ontology. Application ontologies describe concepts depending both on a particular domain and tasks related to a specific application.**

### *2.3.2. Software agents*

Ontologies can also be applied to agent-based computing environments. One approach in that direction is presented in [SMIRNOV 2001] where ontologies act as multi-agents in three forms:

- An application oriented ontology - a conceptual model that describes a real-world application domain depending on the particular domain and problem.
- A resource ontology – a knowledge source description in application ontology terms.
- A request ontology – a user request description in application ontology terms.

In agent-based computing environments, devices, software agents and services are expected to seamlessly integrate and cooperate with each other in support of human objectives – i.e. anticipating needs, negotiating for the service, acting on our behalf, and delivering services in an anywhere, anytime fashion. To serve as the core for such an environment, the authors [CHEN 2003] propose an intelligent Context Broker (CB). CB has the ability to integrate and reason over retrieved information in order to maintain a coherent model of the space, the devices, the agents, and the people in it. An ontology graph based on OWL [OWL 2004] supporting the work of an intelligent CB, is presented in Figure 5.

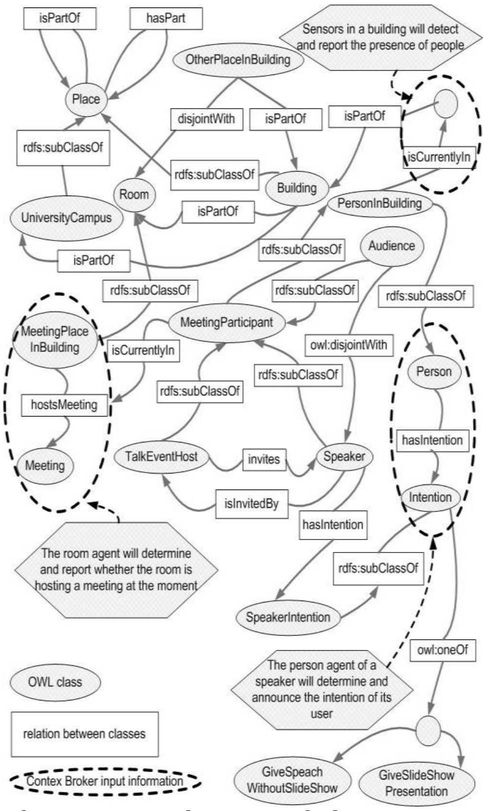


**Figure 5. Ontology graph for context broker processing support. It consists of 17 classes and 32 property definitions. Each one of the classes and properties are used to describe "Person", "Place", and "Intention" from retrieved data. The "Person" class defines the most general properties about a person in an intelligent space (i.e., conference room, office room, and living room). The "Place" class defines the containment relationship properties (i.e., isPartOf, and hasPartOf) and naming properties of a place (like fullAddressName). The "Intention" class defines the notion of user intentions; for example, a speaker's intention to give a presentation and an audience's intention to receive a copy of the presentation slides and handouts. Each oval with a solid line represents an OWL [OWL 2004] class. Each oval with a broken line indicates the kind of information that CB will receive from other agents and sensors in the environment.**

"Semantic interoperability" represents a type of communication between two software agents that work in overlapping domains, whether they use the same or different notations and vocabularies. Ontology-based agents such as given by OntoMerge and OntoEngine [DOU

2004] are offered as one possible solution to implement such “semantic interoperability”. Like its name, OntoMerge is a tool for ontology merging, and OntoEngine is a tool to automate reasoning over merged ontologies. Here, the merger of two related ontologies is obtained by taking the union of the terms and the axioms defining them. Reasoning is automated by means of an inference mechanism that uses a dataset of several ontologies as input and automatically projects the conclusions into one or several target ontologies, as output. Another example of automated reasoning mechanism on top of ontologies is Ontology Inference Layer. It uses the Fast Classification of Terminologies [HORROC 2003] as a system to provide reasoning support for ontology design, integration, and verification [FENSEL 2001].

### *2.3.3. Natural language automatic translation*

OntoLearn [NAVIGLO 2003], a system that automatically associates multiword English and Italian terms, is a practical example of the use of ontologies in automatic translation. This automated learning system extracts relevant domain terms from a corpus of text, relates them to appropriate concepts in a general-purpose ontology, and detects taxonomic and other semantic relations among concepts. The main features of semantic interpretation in OntoLearn are:

- A determination of the right concept (finding the sense) behind each component of a complex term (semantic disambiguation).
- An identification of the semantic relations holding among concepts to build a complex concept.

An example of semantic net in OntoLearn is presented in Figure 6.

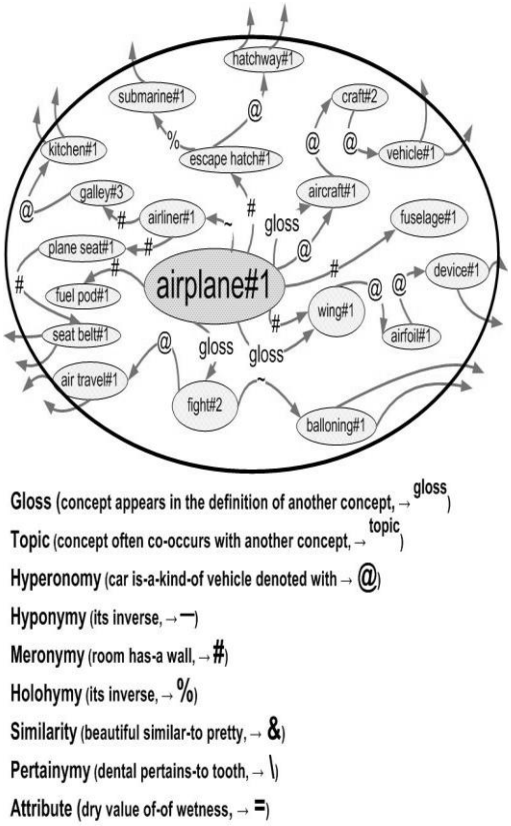


**Figure 6. OntoLearn semantic net for the concept *airplane* (sense number 1, airplane#1). The system automatically builds semantic nets by using the following lexicosemantic relations: Hyperonomy, Hyponymy, Meronymy, Holohymy, Pertainymy, Attribute, Similarity, Gloss, and Topic.**

### *2.3.4. Photo descriptors*

Ontologies can also be used as a tool for describing photos, in order to help in the photo retrieval process. In [SCHREIBER 2001], the ontology-based photo annotation tool consists of the following two features:

- A photo annotation ontology and
- A subject matter vocabulary.

The photo annotation tool provides the description template for annotation construction and consists of the following features:

- A subject matter feature - what does the photo depict?
- A photograph feature - how, when, and why was the photo made?
- A medium feature - how is the photo stored?

The subject matter vocabulary is a domain-specific ontology for the animal domain (basically describing photo's subject matter). It consists of the following four elements:

- An agent (for example: "an ape"),
- An action (for example: "eating"),
- An object (for example: "a banana"), and
- A setting (for example: "in a forest at down").

### *2.3.5. Text mining*

Text mining usually reefers to a process of automatically obtaining information from texts. An example of a text mining software tool based on ontologies is Artequakt [ALANI 2003]. It has implemented the Knowledge Extraction Tool (KET), which searches online documents and extracts knowledge that matches the given classification structure. Artequakt links KET with ontologies in order to achieve efficient information extraction from Web pages. An example of the Artequakt knowledge extraction process is presented in Figures 7a and 7b.

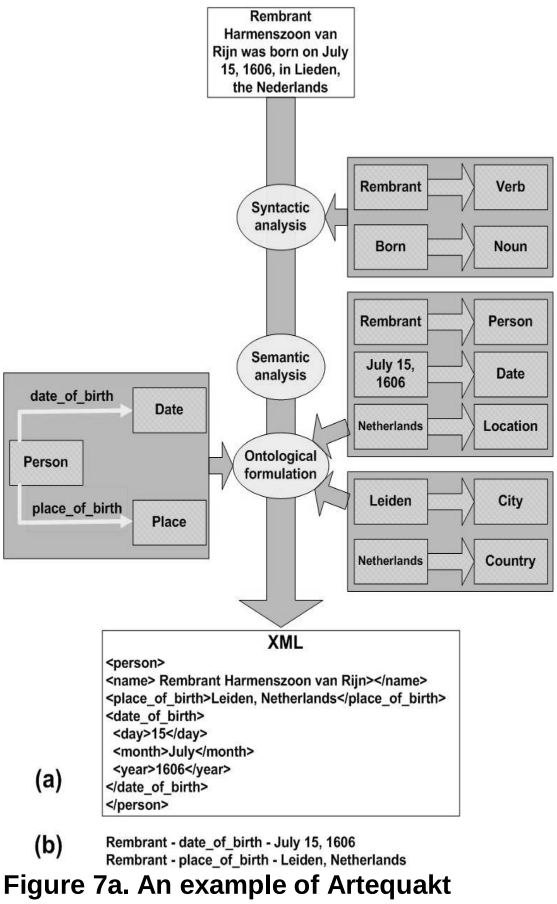


**Figure 7a. An example of Artequakt knowledge extraction from a Web page.**

**When a Web page is recognized to match an input query, it is further processed in a form of syntactic analysis, semantic analysis, and ontological formulation. Outputs are extracted knowledge triplets from the web page in XML syntax, as shown in example (a). After the web page extracted information is presented in a form of XML, it is further processed in a form of ontology, with corresponding instances and relationships, as shown in example (b).**

This section has presented various scientific contributions related to ontology definition, organization, and use. All seek to provide a common means of knowledge representation suitable for further processing. However, the thing that most of the ontologies lack (generally) is atomicity. They usually start in the middle of the problem, not at the beginning and, as a result, tend to be domain specific. They can find concepts in mammography, but not in seismological reports. As such, they can be considered to be content theories about the types of objects, properties of objects and relations between objects that are possible in a specific knowledge domain [CHANDRASEKARAN 1999]. Working up from this basis, the final section focuses on knowledge representation as bringing together both concepts and ontologies.

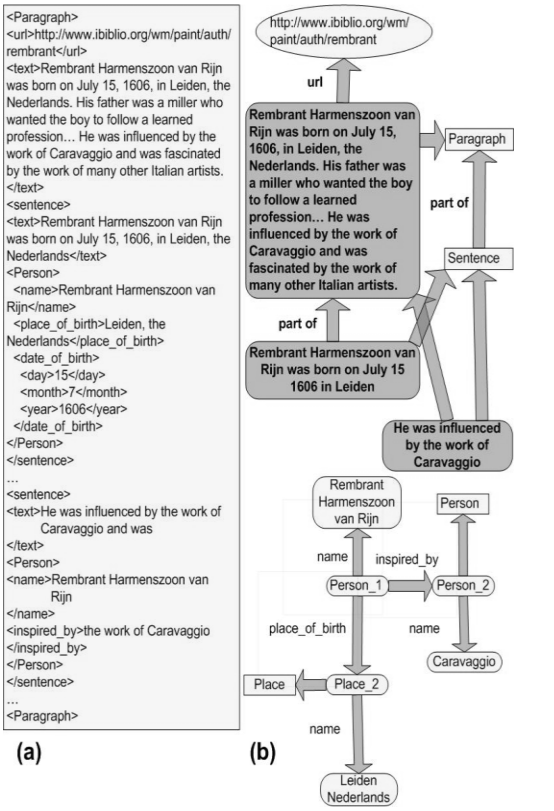


**Figure 7b. Automatic Artequakt ontology population process. Based on the XML file of extracted information from the web page (a), the corresponding instances and relations are made (b), supported by Protégé [PROTEGE 2006].**

## 3. KNOWLEDGE REPRESENTATION

A person can experience knowledge as information at its best. Loosely stated, it is information in support of or in conflict with some hypothesis or it serves to resolve a problem or to answer some specific question. This kind of knowledge that may be expected as the outcome of information processing – or it may be something new and surprising. Information as initially gathered is often fragmented and unstructured and in that form is not suitable for further exchange and processing across different systems. Moreover, a priori one does not usually have a firm grip on what the atoms of knowledge are, how they are connected, how populated, and how one can retrieve or deduce new knowledge from them. In order to answer some of these important questions, the next section begins by examining different definitions of knowledge followed by a discussion of knowledge organization and concluding with practical applications of how knowledge representation uses both concepts and ontologies.

### *3.1 Knowledge definition*

*Knowledge* and *concept* are among most abstract terms in human vocabulary. Just as stated with the term of *concept*, all of the characteristics of knowledge cannot be captured within a single definition. Therefore, we start with some abstract definitions of knowledge:

- The content of all cognitive subject matter [MERRILL 2000].
- A critical resource for any activity [SMIRNOV 2001]: enterprise activity [YOON 2002], intelligent systems [GUO 2005] etc.
- A net made of entities and relationships [MILLIGAN 2003] where relationships between entities provide meaning, and entities derive their meaning from their relationships.

Some more concrete definitions of knowledge related to both concepts and ontology are:

- Conceptual models of information items or systems, including principles that can lead a decision system to resolution or action [HALLADAY 2005].
- In scale-free networks only two types of nodes exist: a few nodes with many connections, and many nodes with very few connections. Concept organization in a scale-free network can be considered as knowledge [HALLADAY 2004].

Because knowledge based on entities and relationships (upgraded in the form of concepts and ontologies) represent the foundation for many intelligent systems, this introduces the problem of how to organize the knowledge in a uniform manner to make it suitable for further processing. In order to provide answers to this important question we next discuss knowledge organization issues.

### *3.2 Knowledge organization*

Generally speaking, knowledge organization is directly related to a particular way of thinking [YOON 2002]. There are many ways to characterize this. [MERRIL 2000] describes process of thinking consisting of knowledge objects. These knowledge objects variously describe the subject matter, the content or what is to be taught. From this perspective, four types of knowledge objects are essential for knowledge organization:

- Entities - things or objects.
- Actions - procedures that can be preformed by a learner on/to/with entities or their parts.
- Processes - events that occur often as a result of some action.
- Properties - qualitative or quantitative descriptors for entities, actions, or processes.

[HALLADAY 2004] simulates the acquisition of knowledge that has been previously organized for education purposes by introducing the concept of clusters in knowledge objects. As the subject matter of an area is learned, the relevant clusters undergo a phase transition among the connections that make up the way the cluster was originally formed.

Starting from another point of view, [PEREZ02] specifies knowledge organization using five components: concepts, relations, functions, axioms, and instances.

In [LAND01] the organization of knowledge is distinguished by the level of formality and by the level of individuality. Formality levels can be expressed as:

- Implicit knowledge, i.e., not well structured, and cannot be easily articulated, or
- Explicit knowledge, i.e., formally represented using a precise and sufficiently formal knowledge representation scheme.

(While it is possible to conceptually distinguish between explicit and implicit knowledge, in practice these are seldom separated, because new knowledge is created through the dynamic interaction and combination of both types.)

Knowledge can also be organized according to the level of individuality as:

- Individual knowledge - as resides in the brains of the individual, and is owned by the individual, or
- Collective knowledge - distributed and shared among members within same team, different teams, and organizations.

Database organization is a common form of explicit knowledge representation that facilitates both mathematical analysis and computer processing. To establish a database organization [ZELLWEGER 2003] uses navigation structures like a network of topic lists, topic data and data relationships (such as one-to-one, one-to-many, many-to-many and many-to-one). Such a database structure is presented in Figure 8a. A structure of semantic relationships database nodes is illustrated in 8b.

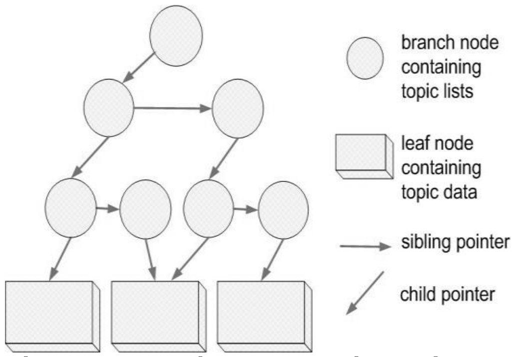


**Figure 8a. Database network graph structure. Data flow starts from the root node and progresses downwards through one or more branch nodes to form paths that link to the leaf nodes. Each branch node has a sibling pointer and a child pointer. The sibling pointer creates a list of topics and the child pointer connects each list to a successor node (either another branch or leaf node). The advance is that any number of topic lists can link to the same topic data. *It is analogous to the situation where different words can link to the same concept.***

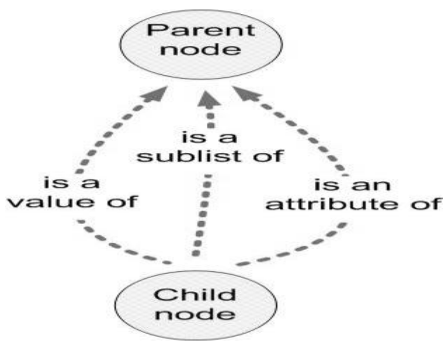

**Figure 8b. Relationships between parent-child nodes within a database organization. As a knowledge structure, each parent-child node pair represents a semantic relationship like "is a sublist of", "is an attribute of", and "is a value of".**

Now that several proposed ways of characterizing knowledge organization are presented, the following section focuses on various possibilities in knowledge use which combine with both concepts and ontologies use.

## *3.3 Knowledge use*

Clearly, a main use of knowledge is in problem-solving [MERRIL 2000]. A successful knowledge structure incorporates the information required for an interested party to solve a particular problem. If the required knowledge components and their relationships are incomplete, then the party will not be able to use the available information efficiently. Problem solving by computer requires not only an appropriate knowledge representation, but also proper algorithms or heuristics to manipulate the knowledge components. A successful problem-solving sequence passes through the following phases:

- knowledge integration,
- knowledge modeling,
- knowledge storage, and
- knowledge retrieval.

### *3.3.1 Knowledge integration*

To begin the information at hand must be cleaned up to remove redundancy, subsumption and contradiction between different knowledge entities – which is the task of knowledge integration [GUO 2005]. [MEDSKER 1995] cites the benefits of knowledge integration in one expert system:

- Existing knowledge can be reused.
- Knowledge acquired from different sources usually has a better validity and is more comprehensive than from only one source.
- Knowledge integration by computer can build a knowledge-based system faster and less expensively than can human experts.

In order to integrate different knowledge sources, the relationships between these different sources must be made explicit. In the process of integrating knowledge hidden relationships may be uncovered that reveal new knowledge. To assure that all such relations remain consistent both before and after knowledge integration one requires a knowledge modeling process – which is the next step.

## *3.3.2 Knowledge modeling*

Knowledge modeling takes the way one thinks about data, information and knowledge from the real world (a human cognition process) and combines it with knowledge models from the information world [WEIQI 2004]. As a consequence, knowledge models incorporate the set of information entities such as data, ontology, rules, logic, and propositions. An example of such a knowledge modeling process is presented in Figure 9.

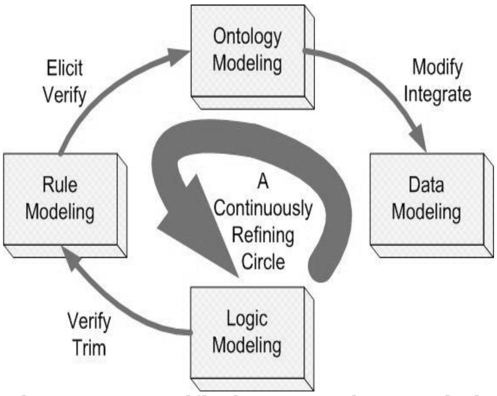


**Figure 9. A Unified Knowledge Modeling process consisting of knowledge models: data, ontology, rule, and logic, forming an inner and outer circle. In the inner circle processes are carried out as follows: data can be used to build ontologies, rules can be formed on the top of these ontologies, and logic can be inferred from these rules. Each knowledge model forms the underlying base for the next model, in contrast to the outer cycle. In the outer cycle each newly built model can be useful to the previously built model: the ontology model can be used in modifying and integrating a data model, a rule model can be used in eliciting and verifying an ontology model, and a logic model can be used in verifying and trimming a rule model.**

[CHAN 2004] presents another possible way to model knowledge by using Knowledge Modeling Systems (KMS) based on an Inferential Modeling Technique (IMT). IMT first models the domain objects and relations before deciding what tasks are involved and what problem-solving method to adopt. Thus KMS consists of two primary modules:

- A class module – gathers user knowledge on classes of objects, the attributes and values associated with each class, and the relationships between the classes, all related to the problem-solving domain.
- A task module – represents an organized structure or a sequence of activities that is performed to accomplish some objective in the problem-solving process.

The main benefit of building a KMS is to gain a shared and reusable knowledge base. The shared and reusable knowledge base paradigm leads us into the next section where our paper discusses how to store knowledge in such a knowledge base, once it is modeled in a uniform manner.

### *3.3.3 Knowledge storage*

There is still no machine that can simulate the efficient way that the human brain stores its data and thinks about them, but generally a person does not even have many static records in his or her head. Over time, mankind has invented increasingly sophisticated means to store knowledge – by writing on stone, papyrus, and paper, later adding recorded speech and film. Now, all kinds of knowledge records are stored digitally in machines. With advances in computer science, knowledge stored in knowledge basis has begun to serve as the foundation for intelligent systems [GUO 2005]. But the lack of consistency in the vast amount of implicit knowledge poses a particular storage problem. [LAND 2001] has designed a specialized conceptual framework to first capture, organize and finally store implicit knowledge in the domain of software engineering. The two phases of this proposed conceptual framework are:

- Knowledge Capture (KC) and
- Knowledge Organization (KO)

as indicated in Figure 10.

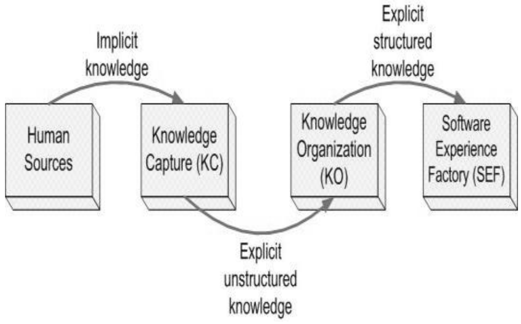


**Figure 10. The process of storing implicit knowledge. Knowledge Capture (KC) extracts implicit knowledge (related to software development) residing in the minds of the parties involved, with other mechanisms such as anecdotes, case studies, lessons learned, best practices, failures, successes, etc. The knowledge retrieved with KC is explicit, but it lacks structure and organization, thus Knowledge Organization (KO) is necessary. KO usually includes transcription (translation from voice or video formats to written form), summarization (production of the main points from transcribed data), and coding (assigning symbols to transcribed data). The output of KO is an explicitly structured knowledge, suitable for further exchange and comparison in a computer system; and serves to populate the Software Experience Factory (SEF). SEF represents the storage of explicit and structured knowledge related to software development.**

After knowledge from different sources has been integrated, modeled in a uniform manner, and stored in a knowledge base, the next step is the purpose of it all: the extraction of knowledge, or knowledge retrieval.

### *3.3.4 Knowledge retrieval*

Knowledge discovery in databases or data mining refers to the nontrivial extraction of implicit, previously unknown, and potentially useful information from the data stored in Databases [FRAWLEY 1992]. Two types of queries and answers for efficient knowledge retrieval in the database domain are cited by [HAN 1996]:

- A simple data query – to find a stored data item in the database (which corresponds to a basic retrieval statement in a database system).

- A knowledge query – to find a certain rules and other kinds of complex knowledge in observed database.

The answering to a database query can take two forms:

- Direct answers that are simple examples of data or knowledge from a database.
- An answer to a query using intelligence – by first analyzing the intent of the query and then providing generalized, neighborhood, or associated information relevant to the query (by means of data summarization, concept clustering, rule discovery, etc).

One possible way to increase efficiency in a Web pages domain knowledge retrieval process is to collect user feedback from the pages visited (so that in future iterations, user searches can better refine and match the system searches). [TSAI 2003] presents a multi-agent framework that iteratively collects user's feedback and updates the user Web page profile. Its task cycle is presented in Figure 11.

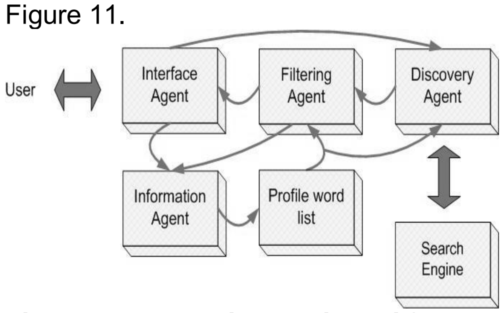


**Figure 11. A Multi-agent based framework for efficient knowledge retrieval from World Wide Web. The framework consisting of agents involves the agents' task cycle as follows: An interference agent receives a user's query and redirects it to existing search engines. Then, an Information agent analyzes the Web pages chosen by the user and derives a temporary user profile. A Discovery agent, based on a user profile, performs query expansion and modification. A Filtering agent ranks the retrieved Web pages correspondingly to the user profile and recommends the most relevant web pages in future queries. The user labels useful pages which are then further processed in profile updating. The knowledge retrieval procedure continues iteratively until a user terminates the search.**

Knowledge conceptualization is a special form of knowledge retrieval processing. The knowledge conceptualizing tool [FUJIHARA 1997] uses concept clustering and ranking techniques by applying a Vector Space Model [SALTON 1975] and a Probability Ranking Principle [ROBERTSON 1997]. An interview transcript, containing several question and answer pairs and consisting basically of unstructured conversational sentences, represents the system input. After processing, the outputs are a set of rules and facts extracted from the input data, thereby forming a new knowledge representation.

This section has presented a list of papers related to knowledge definition, organization, and use. Problem-solving in some sense is the final goal of every knowledge use. Therefore, most of the papers presented focused on how to get there through knowledge representation as a stepwise layered process consisting of knowledge integration, knowledge modeling, knowledge storage, and knowledge retrieval. In this way it is possible to combine different knowledge representations and merge them in order to answer a particular question or, some more general, problem-solving issue.

## 4. CONCLUSION

The research efforts presented here are focused on knowledge representation by ontologies populated with concepts. Concepts, ontologies, and knowledge representation are almost impossible to separate in practice, since there is no clear distinction where the use of concepts stops and use of ontologies begins in knowledge representation. Therefore, most of the research efforts presented are a combination of all tree topics. Thus the survey can be viewed as an annotated guide to this literature.

This paper sheds more light on a selected number of different avenues leading to the same future goal of knowledge retrieval based on conceptual queries, as opposed to the current state of the art based on semantic queries. As indicated in this paper, statements "I am a PhD" and "I have a doctorate," are two different semantical entities, but they both represent the same concept. Therefore, a semantic query (e.g., focused on only one of the above two statements) will be able to retrieve only a subset of relevant knowledge, while a conceptual query (focused on both statements above, as well as all other statements supporting the same concept) would retrieve the full set of relevant knowledge. A trivial solution to the problem is, for each relevant concept, to create a case structure that includes all statements supporting that particular concept. This solution is based on exhaustive approaches, and has no practical value. Practical value lies in the many sophisticated approaches discussed in this survey paper.

The authors believe that this survey will benefit both those who want to enter the field of knowledge retrieval quickly, and those who would like to extend the state-of-art. To the best of our ability, all of the relevant work up to the present has been cited and discussed. For those who are concerned with implementation, there are examples of numerous working systems. Quite clearly there is no overarching "Killer Ap"; the results achieved so far in this domain remain both tentative and incomplete. Much work remains to be done.

## ACKNOWLEDGMENTS

The authors would like to thank Charles Milligan of Sun Microsystems, USA, and Gerald O'Nions of StorageTek, France, who initiated research interest in this exciting field. Also, to Dr. Tom Lincoln, University of Southern California, USA, Dr. Roger Shannon, Duke University, USA, Dr. William Robertson, Dalhousie University, Canada, Ognjen Šćekic, B.Sc., University of Belgrade, Serbia, and Djordje Popović, B.Sc., IPSI Belgrade, Serbia, for constructive feedback on this paper.

## REFERENCES

[ALANI 2003] Alani, H., Kim, S., Millard, D., Weal, M., Hall, W., Lewis, P., Shadbolt, N., "Automatic Ontology-Based Knowledge Extraction from Web Documents," IEEE, Intelligent Systems, Jan-Feb 2003, pp. 14- 21 Vol. 18, Issue 1.

[CHAFFEE 2000] Chaffee, J., Gauch, S., "Personal Ontologies form Web navigation," ACM Press New York, USA, 2000

[CHAN 2004] Chan, C., "The Knowledge Modelling System and its Application," Canadian Conference on Electrical and Computer Engineering, 2-5 May 2004, pp. 1353 - 1356 Vol.3.

[CHANDRASEKARAN 1999] Chandrasekaran, B., Josephson , J., Benjamins, V., "What are Ontologies, and why do we need them?," IEEE, Intelligent Systems and Their Applications, Jan-Feb 1999, pp. 20-26 Vol. 14, Issue 1.

[CHEN 2003] Chen, H., Finin, T., "An Ontology for Context Aware Pervasive Computing Environments," Cambridge University Press, September 2003, Vol. 18, Issue 03.

[COMP 1988] "Principles of Database and Knowledge Base Systems," Computer Science Press, Rockville, Maryland, USA, 1988, Vol. 1.

[DAML 2007] DAML ontologies, DARPA, USA, www.daml.org/ontologies

[DING 2001] Ding, Y., Fensel, D., "Ontology Library Systems: The key to successful Ontology Re-use," Proceedings of the 1st Semantic Web working symposium, Stanford, CA, USA, 30 July – 1 August 2001.

[DOU 2004] Dou, D., McDermott, D., Qi, P., "Ontology Translation by Ontology Merging and Automated Reasoning," Yale University, New Haven, USA, 2004.

[FENSEL 2001] Fensel, D., Harmelen, F., Horrocks, I., McGuinness, D., Patel-Schneider, P., "OIL: An Ontology Infrastructure for the Semantic Web," IEEE, Intelligent systems, Mar-Apr 2001, pp. 38- 45, Vol.16, Issue 2.

[FRAWLEY 1992] Frawley, W., Piatetsky-Shapiro, G., Matheus, J., "Knowledge Discovery in Databases: An Overview," The American Association for Artificial Intelligence, USA, 1992.

[FUJIHARA 1997] Fujihara, H., Simmons, D., "Knowledge Conceptualization Tool," IEEE Transactions on Knowledge and Data Engineering archive, March 1997, pp. 209 – 220 Vol. 9, Issue 2.

[GAUCH 2002] Gauch, S., Madrid, J., Induri, S., Ravindran, D., Chadalavada, S., "KeyConcept: A conceptual Search Engine,"Information and Telecommunication Technology Center, Technical Report: ITTC-FY2004-TR-8646-37, University of Kansas, USA, 2002.

[GRUBER 1993] Gruber, T., "A Translation Approach to Portable Ontologies," Knowledge Acquisition, Nol. 5, No. 2, 1993, pp. 199–220.

[GUARIN98] Guarino, N., "Formal Ontology and Information Systems," Proceedings of FOIS'98, Trento, Italy, 6-8 June 1998.

[GUO 2005] Guo, P., Fan, L., Ye, L., Cao, J., "An Algorithm for Knowledge Integration and Refinement," Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, Guangzhou, 18-21 August 2005.

[HALLADAY 04] Halladay, S., Milligan, C., "The Application of Network Science Principles to Knowledge Simulation," Proceedings of the 37th Annual Hawaii International Conference on System Sciences, Hawaii, 5-8 Jan. 2004.

[HALLADAY 05] Halladay, S., Milligan, C., "Knowledge VS. Intelligence, IPSI Belgrade, Proceedings of the IPSI-2005 Montengro conference, Sveti Stefan, Montenegro, 2005.

[HAN 1996] Han, J., Huang, Y., Cercone, N., Fu, Y., "Intelligent Query Answering by Knowledge Discovery Techniques," IEEE Transactions on Knowledge and Data Engineering, June 1996, pp. 373-390 Vol. 8, No. 3.

[HORROC 2003] Horroc, I., "Fast Classification of Terminologies," School of Computer Science, University of Manchester, UK, 2003.

[KAON 2001] KAON, University of Karlsruhe, Germany, 2001, http://kaon.semanticweb.org.

[LAND 2001] Land, L., Aurum, A., Handzic, M., "Capturing Implicit Software Engineering Knowledge," IEEE Computer Society, 13th Australian Software Engineering Conference, 2001, pp. 108-114.

[MEDSKER 95] Medesker, L., Tan, E. Turban "Knowledge Acquisition from Multiple Experts: Problems and Issues," Expert Systems with Applications, 1995, pp.35-40 Vol.9, Issue 1.

[MERRIL 2000] Merrill, M., "Knowledge Objects and Mental Models," Proceedings of the International Workshop on Advanced Learning Technologies, Palmerston North, New Zealand 12 Apr. - 12. June 2000, pp. 244-246.

[MILLIGAN 2003] Milligan, C., Halladay, S., "The Realities and Facilities Related to Knowledge Representation," IPSI Belgrade, Proceedings of the IPSI-2003 Montengro conference, Sveti Stefan, Montenegro, 2003.

[MOTIK 2002] Motik, B., Maedche, A., Volz, R., "A Conceptual Modeling Approach for Semantic-driven Enterprise Applications," Springer Berlin / Heidelberg, Book on the Move to Meaningful Internet Systems 2002: CoopIS, DOA, and ODBASE: Confederated International Conferences CoopIS, DOA, and ODBASE 2002. Proceedings 2002 Vol. 2519.

[NAVIGLI 2003] Navigli, R., Velardi, P., Gangemi, A., "Ontology Learning and Its Application to Automated Terminology Translation," IEEE, Intelligent Systems, 2003, pp. 22-31.

[OPEN 2002] "The Open Directory Project," 2002, http://www.dmoz.org.

[OWL 2004] OWL, Web Ontology Working Group, 2004, http://www.w3.org/2004/OWL/

[PEREZ 2002] Gomez-Perez, A., Corcho, O., "Ontology Languages for the Semantic Web," IEEE, Intelligent Systems, Jan-Feb 2002, pp. 54-60 Vol.17, Issue 1.

[PROTEGE 2006] Protégé, Stanford University, USA, 2006, http://protege.cim3.net/cgi-bin/wiki.pl?ProtegeOntologiesLibrary.

[ROBERTSON 1994] Robertson, S., Jones, K., "Simple, Proven Approaches to Text Retrieval," Technical report TR356, Cambridge University Computer Laboratory, UK, 1994.

[ROBERTSON 1997] Robertson, S., "The Probability Ranking Principle in Information Retrieval," Morgan Kaufmann Publishers, Readings in information retrieval archive 1997 pp. 281 – 286.

[SALTON 1975] Salton, G., Wong, A., "A Vector Space Model for Automatic Indexing," Communications of the ACM, 1975, pp. 613 - 620 Vol. 18, Issue 11.

[SCHATZ 1997] Schatz, B., "Information Retrieval in Digital Libraries: Bringing Search to the Net," Science, 17 January 1997 pp. 327-334 Vol. 275.

[SCHLOSSER 2002] Schlosser, M., Sintek, M., Decker, S., Nejdl, W., "HyperCup – Hupercubes, Ontologies and Efficient Search on P2P Networks," International Workshop on Agents and Peer-to-Peer Computing, Bologna, Italy, July 2002.

[SCHREIBER 2001] Schreiber, A., Dubbeldam, B., Wielemaker, J., Wielinga, B.,"Ontology-Based Photo Annotation," IEEE, Intelligent Systems, May-Jun 2001, pp. 66- 74 Vol. 16, Issue 3.

[SMIRNOV 2001] Smirnov, A., Pashkin, M., Chilov, N., Levashova, T., "Ontology Management in Multi-agent System for Knowledge Logistics," Proceedings of the International Conferences on Info-tech and Info-net, Beijing, China 2001, pp. 231-236 Vol.3.

[TSAI 2003] Tsai, T., Lee, W., "An Interactive Agent-based System for Concept-based Web Search," Expert Systems with Applications, May 2003, pp. 365-373 Vol. 24, Issue 4.

[WEIQI 2004] WeiQi, C., JuanZi, L., KeHong, W., "CAKE: The Intelligent Knowledge Modeling Web Services for Semantic Web," The 8th International Conference on Computer Supported Cooperative Work in Design Proceedings 26-28 May 2004 Xiamen, China, pp. 209-216.

[W3C 2007] World Wide Web consortium, http://www.w3.org

[YOON 2002] Yoon, T., Fujisue, K., Matsushima, K., "The progressive Knowledge Reconstruction and its Value Chain Management," Engineering Management Conference, 2002. IEMC '02. 2002 IEEE International, 2002, pp. 298- 303 Vol.1.

[ZELLWEGER 2003] Zellweger, P., "A Knowledge–based Model to Database Retrieval," Proceedings of the International Conference on Integration of Knowledge Intensive Multi-Agent Systems, 30 Sept.- 4 Oct. 2003, pp. 747- 753.

# Reviewers by Countries

***Argentina***
Olsina, Luis; National University of La Pama
Ovando, Gabriela P.; Universidad Nacional de Rosario
Rossi, Gustavo; Universidad Nacional de La Plata
***Australia***
Abramov, Vyacheslav; Monash University
Begg, Rezaul; Victoria University
Bem, Derek; University of Western Sydney
Betts, Christopher; Pegacat Computing Pty. Ltd.
Buyya, Rajkumar; The University of Melbourne
Chapman, Judith; Australian University Limited
Chen, Yi-Ping Phoebe; Deakin University
Hammond, Mark; Flinders University
Henman, Paul; University of Queensland
Palmisano, Stephen; University of Wollongong
Ristic, Branko; Science and Technology Organisation
Sajjanhar, Atul; Deakin University
Sidhu, Amandeep; University of Technology, Sydney
Sudweeks, Fay; Murdoch University
***Austria***
Derntl, Michael; University of Vienna
Hug, Theo; University of Innsbruck
Loidl, Susanne; Johannes Kepler University Linz
Stockinger, Heinz; University of Vienna
Sutter, Matthias; University of Innsbruck
Walko, Zoltan
***Brazil***
Parracho, Annibal; Universidade Federal Fluminense
Traina, Agma; University of Sao Paulo
Traina, Caetano; University of Sao Paulo
Vicari, Rosa; Federal University of Rio Grande
***Belgium***
Huang, Ping; European Commission
***Canada***
Fung, Benjamin; Simon Fraser University
Grayson, Paul; York University
Gray, Bette; Alberta Education
Memmi, Daniel; UQAM
Neti, Sangeeta; University of Victoria
Nickull, Duane; Adobe Systems, Inc.
Ollivier-Gooch, Carl; The University of British Columbia
Paulin, Michele; Concordia University
Plaisent, Michel; University of Quebec
Reid, Keith; Ontario Ministry og Agriculture
Shewchenko, Nicholas; Biokinetics and Associates
Steffan, Gregory; University of Toronto
Vandenberghe, Christian; HEC Montreal
***Croatia***
Jagnjic, Zeljko; University of Osijek
***Czech Republic***
Kala, Zdenek; Brno University of Technology
Korab, Vojtech; Brno University of technology
Lhotska, Lenka; Czech Technical University
***Cyprus***
Kyriacou, Efthyvoulos; University of Cyprus
***Denmark***
Bang, Joergen; Aarhus University
Edwards, Kasper; Technical University Denmark
Orngreen, Rikke; Copenhagen Business School
***Estonia***
Kull, Katrin; Tallinn University of Technology
Reintam, Endla; Estonian Agricultural University
***Finland***
Lahdelma, Risto; University of Turku
Salminen, Pekka; University of Jyvaskyla
***France***
Bournez, Olivier
Cardey, Sylviane; University of Franche-Comte
Klinger, Evelyne; LTCI – ENST, Paris
Roche, Christophe; University of Savoie
Valette, Robert; LAAS - CNRS
***Germany***
Accorsi, Rafael; University of Freiburg
Glatzer, Wolfgang; Goethe-University
Gradmann, Stefan; Universitat Hamburg
Groll, Andre; University of Siegen
Klamma, Ralf; RWTH Aachen University
Wurtz, Rolf P.; Ruhr-Universitat Bochum
***Greece***
Katzourakis, Nikolaos; Technical University of Athens
Bouras, Christos J.; University of Patras and RACTI
***Hungary***
Nagy, Zoltan; Miklos Zrinyi National Defense University
***India***
Pareek, Deepak; Technology4Development
Scaria, Vinod; Institute of Integrative Biology
Shah, Mugdha; Mansukhlal Svayam
***Ireland***
Eisenberg, Jacob; University College Dublin
***Israel***
Feintuch, Uri; Hadassah-Hebrew University
***Italy***
Badia, Leonardo; IMT Institute for Advanced Studies
Berrittella, Maria; University of Palermo
Carpaneto, Enrico; Politecnico di Torino
***Japan***
Hattori, Yasunao; Shimane University
Livingston, Paisley; Linghan University
Srinivas, Hari; Global Development Research Center
Obayashi, Shigeru; Institute of Fluid Science, Tohoku University
***Mexico***
Morado, Raymundo; University of Mexico
***Netherlands***
Mills, Melinda C.; University of Groningen
Pires, Luís Ferreira; University of Twente
***New Zealand***
Anderson, Tim; Van Der Veer Institute

Authors of papers are responsible for the contents and layout of their papers.

## Welcome to IPSI BgD Conferences and Journals!

http://www.internetconferences.net

http://www.internetjournals.net

**VIPSI-2010 AMALFI**
**Hotel Santa Caterina**
**Amalfi, Italy**
**March 4 to March 7, 2010**

**VIPSI-2010 TIVAT**
**Tivat, Montenegro**
**June 4 to June 7, 2010**

**VIPSI-2010 TIVAT**
**Tivat, Montenegro**
**October 1 to October 4, 2010**

**VIPSI-2010 MILOCER**
**Milocer, Montenegro**
**December 30 to December 31, 2010**

**CIP – Katalogizacija u publikaciji**
**Narodna biblioteka Srbije, Beograd**
ISSN 1820 – 4511 =
The IPSI BGD Transactions on Advanced Research
COBISS.SR - ID 119128844

ISSN 1820-4511
9 771820 451006