

Ontology optimisation–Problematics & methodology, with a first step of formalism

Truong My DUNG¹ and Nguyen Dinh NGOC²

¹VNU in HCMC

²ThangLong Univ,Hanoi

ABSTRACT

After a short tentative of phenomenology to sketch out the background context of Ontology-based Semantic Web, the authors overview the problematics related to the ontology optimisation, then make some suggestions for a somewhat new methodology's approach, taking into account the impact of evolving Automation technology on Semantic Web design and deployment. Then the paper proposes a tentative of ontology optimisation formalism, and open source ontology software optimisation. After pointing the Road ahead for Vietnam, the paper concludes by a call for cooperation/collaboration on related education and training in Vietnam, and continuing R&D on Ontology optimisation.

KEYWORDS

Computational Mathematics, Formalism, Interactive Broadband Internet 3G/4G Technology, Methodology, natural language generation, Ontology, Open Source Software, Optimality Criteria, Optimisation, Problematics, Quality-of-Interaction, Semantic Web Industry, Vietnamese WORDNET/ SENSUS, WEB-Engineering/Services.

1 Introduction

The Semantic Web (SW), emerging from about 2000, has been linked to Computational Linguistics (CL) and Cognitive Sciences (CS), especially Cognitics, as Informatics for CS, and has rapidly evolved scientifically and technologically and becomes now a promising “Semantic Web Industry” (SWI) in the New Internet Industry (NII), driving a mutation of Ontology, from an esoteric branch of Metaphysics as the study of nature and meaning of Existence and Being (“Je pense, donc je suis!”) to an active research domain of Computational Machine Intelligence (CMI), tightly linked to Computational Linguistics (CL) and Cognitive Sciences (CS), especially Cognitics. [3–15]

By reviewing attentively the related Literature, we

encountered a bit often the notions such as: optimal, optimisation, better and best, but at least till now (July, 2005) almost no formal and formalised settings of a “regular optimisation problem” as in classical and modern operations research.

Practically, in a contract with MITANI,¹ the first author, Truong My Dung, and her IT– Master students, by using an Enhanced Vietnamese Adaption of ORACLE-TEXT in Designing a tailored (for MITANI). Search Engine (SE) by Vietnamese Keywords, have to face with an Ontology-Building issue, on their 90 GB Text repository for Vietnamese Search Experimenting. [1], [2], [45], [46]

From his study of Knowledge Management for Power System Protection (PSP), the second author has also to understand how an Ontology-based Knowledge – Integration - Dissemination (KID) Clustering could

Received July 29, 2005 ; Revised October 6, 2005 ; Accepted October 7, 2005.

¹tmdung@fit.hcmuns.edu.vn, tmdung@az.com.vn

²nguyen.dinh.ngoc@thanglong.edu.vn, ngocnd@msn.com

DOI : 10.2201/NiiPi.2005.2.7

¹ A Japanese company cooperating with the Faculty of Information Technology (FIT), University of Natural Sciences, Vietnam National University at HOCHIMINH–City, Vietnam.

be enhanced to become Real-Time/Just-in-Time for a better Decision-Support-System (DSS), in Large-Scale Power System Protection (LSPSP).

So, the two authors decided to join efforts to understand – at some extent – the notion, feasibility and formalism of ontology optimisation, under the high-lighting of classical mathematical optimisation guidelines.

Here following, after sketching the context by a short tentative of Phenomenology, we will examine the Problematics, leading to some suggestions for a somewhat new methodology's approach, a first step of ontology optimisation formalism, related open source software issues, and the Road ahead for Vietnam, then we conclude by a call for cooperation-collaboration on education and training—at the Master of S&T level—in ontology technology and ontology-based/guided Semantic Web Services and Applications, with adequate Optimisation.

2 The context sketched by a tentative of phenomenology

In the DELOS² project, we can find the so-called Delos Ontology Harmonisation: in the Delos network of excellence on digital libraries. *Ontology Harmonisation Working Group* (OHWG), funded by the European Commission's IST Programme, is exploiting the Potential for Harmonisation between the CIDOC Conceptual Reference Model (CRM) and the ABC Model developed by the Harmony Group, possibly resulting in the Development of a combined "Super-Ontology."

In the article "*Different Firms, Different Ontologies, and NO one Best Ontology*" by Davel, E.O'Leary, we can read: "Increasingly, Firms are developing Best Practices Knowledge Bases as part of their Knowledge-Management System; Best Practices (or Leading Practices) Knowledge Bases provide access to enterprise process and attempt to define the best ways of doing things. These best practices knowledge bases are based on Ontologies, what the developers call "*Best Practices Ontologies*." Because one of the common reasons firms give for developing ontologies is their capacity for Reusing, it would seem that only one ontology would be necessary, and that all firms could use the same one. However, different firms have their own "*Best Practices Ontologies*" According to this, we will examine why different firms want different ontologies, and we will also present a theoretical model finding that there is no optimal ontology!

It means that, unlike the IEEE's Standard Upper Ontology, the notion of "Optimal Ontology" is rather objective-oriented.

²<http://www.ercim.org/delos/>

In the article "*Approximating an Interlingua in a Principled Way*" by Eduard Hovy and Sergei Nirenburg,³ we could point out an *asymptotical* approach: "We address the problem of constructing in a principled way an ontology of terms to be used in an Interlingua for Machine Translation. Given our belief that a true Language-Neutral Ontology of Terms can only be approached asymptotically, the construction method outlined involves a step-wise folding-in, of one language at a time"...

In the PhD Thesis, "*An Algebraic Framework for the Interpretation of Ontologies*" by Prasenjit Mitra, Stanford University, August 2004,⁴ we could notice the evolution of the notion of Optimisation from 5.2. The Use of Articulation Rules for Query Answering: "*Several existing optimal algorithms can perform the rewriting and evaluation of Queries in such a setting,*" to 8.2.6. Query Rewriting and *Optimisation* in Peer-to-Peer Systems.

So, it is about query optimisation but not yet at the level of Ontology optimisation.

About *Sub-Ontologies*, we can read in an important paper by Mehul Bhatt et al. in AINA'04⁵: "As information on the Web increases significantly in size, Web Ontologies also tend to grow bigger, to such an extent that they become too large to be used in their entirety by any single application. This has stimulated our work in the area of *Sub-ontology extraction* where each user may extract optimised sub-ontologies from an existing base-ontology. Sub-ontologies are valid independent ontologies, known as materialised ontologies, that are specifically extracted to meet certain needs. Because of the size of the original ontology, the process of repeatedly iterating the millions of nodes and relationships to form an *optimized* sub-ontology can be very extensive. Therefore we identified the Needs for a Distributed approach to the Extraction process. As ontologies are currently widely used, our proposed approach for distributed sub-ontology extraction will play an important role in improving the efficiency of information retrieval."

In the Handbook on Ontologies-2004,^{6,7} we could see over the significant *Table of Contents*, (see Annex) that it has been reflected at some extent the current Research on Ontology Technology and important Domain Ontologies.

It is important to notice that the field of *Ontology*

³<http://acl.lido.upenn.edu/H/H92/H92-1052.pdf>

⁴<http://www-db.stanford.edu/~prasenj9/thesis.pm.pdf>

⁵<http://doi.ieeecomputersociety.org/10.1109/AINA.2004.1283981>

⁶S. Staab and R. Studer (Eds.) Handbook on Ontologies, International Handbooks on Information Systems, Springer 2004, ISBN 3-540-40834-7.

⁷<http://www.informatik.uni-tr+ier.de/~ley/db/books/collections/StaabS2004/html>

Optimisation has not been systematically examined with methodology and algorithms.

But in the research paper “*OntoEdu: Ontology-based Education Grid System for E-Learning*” by CuiGuangzuo⁸ we can read in the last part, *conclusion and future works*, the following claim: “Ontology is the core knowledge of education platform; Concept reuse could be realized by education ontology; Ontology is also the core technique of Semantic Grid; *Ontology Optimisation* and Logic Application **will be our Next Work.**”

So, Ontology Optimisation is also in the research scope of Prof. Cui Guangzuo’s school in Beijing University, P.R of China.

Beside the above-mentioned examples there are many other use cases in SEKT⁹ suggesting a systematic, mathematical Investigation of useful optimisation problems, either application-driven or fundamental-research-oriented domain. We could also notice that beside Ontology-Optimisation there is also a proposition on *Optimisation-Ontology*, as in GEODISE home page (Ontology Services)¹⁰:

3 A first step for problematics viewing a tentative of ontology optimisation formalism

Till now, we have not yet seen a ontology optimisation formalism in the classical understanding, with objective function, constraints, optimality criteria, etc...

In order to make a first step in this direction, we will start sketching a local formalism with an adequate case study, then investigating how it is possible to extend this formalism, in order to generalise to some more general settings.

By examining many potential situations, we have decided to begin with a typical use case, in the spirit of *user-oriented/centered information systems* as follows:

In the study of UKARL use cases for SEKT, especially for *ontology usage mining*, some methods and techniques are required:

First, means are needed to trace and store the actual usage of the system and of the underlying ontology employed. Moreover, it is indispensable to store changes that are made to the ontology in case the system under consideration employs an evolving ontology.

Second, adequate optimality criteria should be selected to enable a quality estimation of the quality of the users’ interaction with the system.

Such measures ought to reflect the users’ information needs and how those are fulfilled by the target system. When considering the improvement of system usage as an Optimisation problem, it may be feasible to apply Machine Learning Techniques (MLT) to optimise the measures mentioned and thus to refine and optimise the underlying ontology.

Third, in order to realise these ideas, of course it is necessary to access to methods to adapt the system’s structure (i.e. the ontology), with respect to usage optimality. Another required technique concerns the analysis of different users and user groups, i.e. the determination of User Profiles, for which the underlying Ontology may be tailored.

To summarise, we have to consider the main Entities in the Optimisation Problem [OP]:

- Information System [IS]
- User [U]
- Use Group [UG]
- User Profile(s) [UP]
- User-System Interaction [UIS]
- Underlying Ontology [UO]
- Use-case of Ontology-Usage Mining [OUM]
- Means [MTS] to Track and Store the Actual Usage of the System [IS] of the Underlying Ontology [UO] employed.
- Changes [OCh] that are made to the evolving Ontology [UO]
- User-System’s Interaction Quality [QoI]
- User’s Actual Information Needs [UIN]
- System Usage [SU]
- Measures [MS] of [QoI] of [SU]
- Machine Learning Techniques [MLT] to Optimise the Measures [MS] in order to Optimise the Underlying Ontology [UO]

By Combining the above-listed Entities in an *Optimisation Formalism*, we could at some extent, FORMALISE the corresponding Optimisation Problem [OP], along the Optimality Criteria [OC] allowing an Estimation [EQ] of the Quality [QoI] of [UIS].

To formalise: Given an Information System [IS] used by an User [U], we consider the improvement of System Usage as an *Optimisation Problem* [OP]. Adequate *Optimality Criteria* [OC] are needed, which allow for an Estimation [EQ] of the Quality [QoI] of the User’s Interaction [UIS] with the System [IS].

Another required Formalisation concerns the Analysis of Different Users [U] and User Groups [UG], i.e. the Determination of User Profiles [UP], for which the Underlying Ontology [UO] may be tailored. In use case of Ontology-Uuage-Mining [OUM], some

⁸E-mail cgz@pku.edu.cn

http://www.fao.org/agris/AOS/workshops/China_AOS/ppt/4-3.pdf

⁹Semantically Enabled Knowledge Technologies-University of Karlsruhe.

¹⁰<http://www.geodise.org/toolboxes/generic/ontology.htm>

more methods and techniques should be required, as sketched below, which will have to be considered formally.

First, Means [MTS] are required to track and store the actual usage of the information system [IS] and of the Underlying Ontology [UO] employed.

Second, it is indispensable to store changes [OC] that are made to the Underlying Ontology [UO] in case the Information System [IS] under consideration employs an evolving Ontology [UO].

Third, Measures [MS] of the User-System's Interaction quality [QoI] ought to reflect the User's Actual Information Needs [UIN] and how those are fulfilled by the System [IS] used.

Fourth, when considering the improvement of System Usage [SU] as an Optimisation Problem [OP], it may be feasible to apply Machine Learning Techniques [MLT] to optimise the Measures [MS] above-mentioned and thus to refine and optimally improve the Underlying Ontology [UO].

Fifth, in order to realise these Ideas, of course, it is necessary to have Method [MT] libraries to adapt the System's Structure [SS] (i.e. the Underlying Ontology [UO]) with respect to the Optimality-of-Use [OU].

From the above-sketched "informal" Holistic View of the considered Optimisation Problem [OP], we could see that it is at some extent not a classical or routine Operations Research (OR) problem. For that reason we have to be a bit more attentive regarding the methodology.

So, before going further in this direction, we would like to consider some unavoidable Methodology's Aspects, useful to the future Algorithm's Design.

4 Some suggestions for a new Methodology's approach:

First, it is important to take into account the upgrading ascension from the old HyperText to HyperMedia – AV, then to the now emerging HyperKnowledge level, from the old-traditional Web, as first generation via the second Broadband for AudioVisual Quality of Services to Semantic Web (SW) now, as a third generation, with new Possibilities-Functionalities-Facilities in Knowledge Interchange with the Knowledge Interchange Format (KIF), Sharing, Automated Processing at the Machine-Level, Knowledge Integration and Dissemination (KID) in useful KID-Clustering for many kinds of Information Flow Framework (IFF) in various imaginable Internet 3G/4G applications, including e-Science & Technology ones.

So the related methodology has to be upgraded from the classical optimisation platform to an Hyper-Knowledge-based Semantic Web optimisation plat-

form, with relevance to Automatic Knowledge Processing by the Machine / Infrastructure.

Second, in this High-End SW-Services, a permanent attention should be given to Automation-enabled and Semantically Enabled Knowledge Technologies [e.g. University of Karlsruhe (UKARL), Institut fuer Angewandte Informatik]. That is, methods from Mathematics, Informatics are invited to cooperate–collaborate with their homologues in Cognitive Science & Technology (CST), especially in Cognitics (at least from 1983).

Third, special attention to the paradigm shift to pervasiveness is needed, with pervasive computing on the new generation Internet 3G/4G, or even Peer-to-Peer 100 Mbit/sec Intranet/Internet as in current R&D at the ICT Department of the South-East University, Nanjing, P.R. of China, since 2001.

Hereafter is a very short *study framework* of the potential impact of the evolving automation technology - in the large - on Semantic Web Deployment. ([44], to be revisited in VICA 7, 2007)

- *Introduction*. The Context: Automation Technology on Submicronics - along the ITRS, Pervasive Internet 3G/4G and Semantic Web for Hyper-Knowledge under Globalisation.
- *Instrumentation and Control* on the Frontiers of a New Millenium: Timing scale evolution: from hundred-millisecond Synchronisation to Submillisecond Synchronising; Intel Open Control Technology for Intel Online Services and the "Customer Pod."
- *New Technology, New Approaches*: Pervasive Automation Technology, a paradigm shift; The Automation/Commercial Implications of the Convergence of Grid Computing, Semantic Web Services and Self-Managing Systems.
- *The Technology "LAWS"* (revisited) that define Exponential Advances of Technology:
 - MOORE: The Number of Transistors on a Chip Doubles every Two years.
 - MOORE-HWANG: driven by the Mobile Revolution, semiconductor capacity would double every twelve months.
 - ROCK: The Cost of Semiconductor Tools Doubles every Four years.
 - METCALFE: A Network's Value grows in proportion to the Square of the Number of its Users.
 - WIRTH: Software Execution is Slowing Faster than Hardware is Accelerating.
 - GILDER: Bandwidth grows at least Three times Faster than Computer Power.
 - MURPHY^{11,12}:

- *The Pervasive Internet - Smart Power*: Pervasive Internet Technology in a Changing Energy Market; *Distributed-Grid-Pervasive Internet Computing* in HP's CoolTown, with Emerging Pervasive Workplace.
- *Intelligent Connected Appliances* with Distributed Autonomous Intelligence.
- *Wireless Sensor Networks*: Wireless Systems in the Automation & Control Environment at the Field Level; ISA-SP100 Committee for Wireless Standards.
- *Networked, Intelligent Input/Output*: Ten-Gigabit Ethernet versus Broadband Internet 3G/4G for I/Os; Virtual Devices Network (VDN) and Networked / Internet-enabled Monitoring Systems.
- *The Truly Distributed Control Revolution*: The Visibility and Management Control of Individual Items *without* the Human Interaction will revolutionise the way Items are produced, warehoused and distributed.
- *Towards Fully Automated-Integrated Factory*: To Automate or not to Automate, a Scalability Issue.
- *Nanotechnology and Self-Organising-Systems*: Stephen WOLFRAM's "New Science" with Cellular Automata, theory and application.
- *Intelligent Robots will be Everywhere*: Computational Machine-Intelligence for Hyper-Knowledge Semantic Web-based Online Learning, even for mobile intelligent robots.

As a conclusion to the exponential Advances of Technology, there is a hard challenge to overcome for Semantic Web Design and Deployment, from automated reasoning in Computational mathematics to automated system design in Automation & Control (A&C) via Mathematical Ontology of A&C Engineering,[20–22] **but this will be a worthy Opportunity to catch up.**

The model issue

Unfortunately, till now at least, we have not a widely-adopted relevant Open Systems Interconnection model as the well-known classical Standardised 7-level/layer ISO/OSI [Physical, Link, Network (IP), Transport (TCP/IP), Session, Presentation, Application (in Information-Communications-Technology (ICT), with notions on Applet/JAVA, Servlet/JINI, now complemented by Contentlet/SEKT].

A layered "Tower" of 4 to 5 levels has to be added on the top of the classical 7-level/layer ISO/OSI ICT-model, according to the considered main Application, with future-proof attention to Scalability (vertical/horizontal), Reusability, Inter-Operability, 2-level Semantics (denotational on operational by a Lambda-Calculus, as a compilation language), and, last-but-not-least, Multilingual Access (e.g. by Unicode), either by written or voice I/O s, with special kind attention to unfortunated-disabled people (blind, deaf, dumb, etc.).

For some problems, the **W3C Semantic Web Stack Diagram** could be useful, at least for the Problem settings: **1–URI, Unicode; 2–XML, Namespaces; 3–RDF M&S; 4–RDF Schema; 5–Ontology; 6–Rules; 7–Logic Framework; 8–Proof; 9–Trust; Vertically: 3–8–Signature, Encryption.**

5 A tentative of Ontology Optimisation Formalism, as future plan

Let us analyse the above-mentioned use-case to see to point out the key elements/actors of the corresponding *Optimisation Problem*, by using the notations already foreseen in Section "A first step for problematics viewing a tentative of ontology optimisation formalism."

For the moment, in absence of better models, we will try to adapt the **W3C-9stack-model** recalled above, from the **Fifth Stack** : [23–35] (underlying) ontology.

- 1) For every not-too-short-term ontological problem, it is often recommended to examine at least *three* levels: deep, intermediate, surface ontologies. For the deep level, like SUO (IEEE Standard – Upper - Ontology) somewhat temporally stable, the *Optimality Criterion* could be "the smallest sub-ontology of SUO." For the surface level one, domain-dependent and application-driven, the *maximal effectiveness* for the QoI (Quality-of-Interaction) between a user (defined by U) and the information system (called IS) could be for some purpose an acceptable optimality criterion. For the intermediate level, by learning from the good example from IBM database practice QBE-QMF-SQL (Query-by-Example/Query Management Facilities/Structural-Query-Language), we adopt as *Optimality Criterion* the maximal effectiveness of ontology-based/guided query-management.
- 2) For the **sixth stack** Rules, it could be the *maximal effectiveness* of Rules software application.
- 3) For the **seventh stack** Logic Framework, there could be many different choices, but from the computational logic view-point, and experience

¹¹ <http://www.spectrum.ieee.org/careers/careerstemplate.jsp?ArticleId=n120403>

¹² <http://www.geocities.com/murphylawsite/technologylaws.html>

in logic programming, we would like to choose *minimal logic computing time*.

- 4) For the **eighth stack** Proof, the simplest-shortest proof process is a good candidate.
- 5) For the highest **ninth stack** Trust, the *maximal trust degree* is acceptable at least qualitatively, although sometimes quantitatively questionable.

So we have projected the Optimality Criteria on the five upper-level stacks, by admitting that the IS has been, at some extent, classically *optimised* for the four bottom-up stacks as First URI-Unicode, Second XML-Namespaces, Third RDF M&S, and Fourth RDF-Schema.

For the two vertically managing **3-8 stacks** of Signature-Encryption, we would like to adopt a min-max approach: *minimal complexity/maximal security*.

According to this Optimality - Criteria - Analysis, many QoI -based / guided *Objective Functions* could be conceived and represented adequately.

For the *Constraints* (C), both qualitative and quantitative, we have examined the specific features of the already chosen use-case, especially the Semantically Enabled Knowledge Technologies (SEKT) ones, always - to be consistent - under the highlight of the **W3C-9stack-Model**.

First, for Ontology: a) Constraint(s) on the Underlying Ontology's Changes - Evolution Possibilities, b) Constraint(s) on its User-Profile(s) -tailored Refinement-Improvement(Optimisation) Possibilities.

Second, Constraint(s) on Machine - Learning - Techniques (MLT) to optimise the QoI-Measures of U-IS Interaction.

Third, Constraint(s) on determination of User Profiles, User Modeling.

Similarly, other *Constraints* on the stacks 6-7-8 Rules, Logic Framework, Proof, and vertically from stacks 3-to-8, on Signature, Encryption are to be discussed and determined accordingly.

It remains to *combine* the above-conceived notions of *Objective-Function* (OF), *Constraints* (C) to conceive and design **Algorithms** for the Evolving-Ontology-based Information-System (IS) and the Interaction-Protocol between IS and the user U, to control the Optimality according to the Optimality-Criteria (OC). This new kind of Algorithmics-study will be examined formally in one of our subsequent research papers.

We could also tackle the problem from an Adaptive Optimal Control view-point: an "Intelligent" IS could be able to *determine* the User Profile by User-Modelling, then *adapt* the Interaction-Protocol then Optimise the QoI. This will be examined in a subsequent paper, in preparation for VICA 7 (2007), in Hanoi, SR. of Vietnam.

For the coming VICA 7 further research works have been focussed on Fuzzy Optimisation of Fuzzy Ontologies, under the highlights of some new concepts and theoretical findings of Professor L.Zadeh and his school, as introduced briefly in a Special Contribution to the No 1 of "Progress in Informatics," [47] introducing the notion of "Word Computing," useful and applicable to the Ontology optimisation problem.

A first section of our coming paper related to this research will be on "Fuzzification" of the 9-stack Semantic Web Model recalled above, section 4 : Stack 5: Fuzzy Ontology, stack 6: Fuzzy Rules, stack 7: Fuzzy Logic Framework, stack 8: Fuzzy Proof, and even - sometimes, for some specific applications- Fuzzy Trust, as in Fuzzy Vaults ;¹³ also in.¹⁴

A second unavoidable one is on Fuzzy Optimisation related pertinently to Fuzzy Ontology. After these two sections, the most difficult one will be "A First Step of Fuzzy Formalism" leading to some Word Computing Algorithms, a la L. Zadeh et al.

Following, after the Computational Mathematics use-case, is an important related problem: optimisation of ontology-software, in particular, **open-source** ones, which will be necessary to "install" Computational Mathematics on the Semantic Web platform.

6 Computational Mathematics on the Semantic Web: Feasibility and Application Perspectives

Another useful use-case worth to be examined is how to enhance Intranet/Extranet to Semantic Web, with Collaborative/Cooperative High Performance Scientific Computing (HPSC), with Computational Mathematics in the newest understanding, that is with Computer-Aided Formal Reasoning (CAFR), Automated Reasoning (AR), Theorem Proving (TP), Large Software Package Correctness Proof, etc...

It is feasible due to the 9-stack model recalled above, especially with the four stacks 5-6-7-8: Ontology - Rules - Logic Framework - Proof, compatible or adaptable to current Mathematical System Software as Axiom, Maple, Mathematica, MuPad, etc... for collaborative/cooperative Grid Computing or Internet-enabled Parallel Computing.

Application Perspectives are easily seen on the Scopes of FoCM International Conferences from 1995 to 2005, as well as on the Contents of the new and ascending Journal of Computational Mathematics, since 2001; see also the official Web site of the Association for Computational Mathematics.

¹³ <http://www.cacr.math.uwaterloo.ca/conferences/2002/isw-tenth/juels.html>

¹⁴ eprint.iacr.org/2002/093.ps

7 Open Source Ontology Software Optimisation (OSOSO)

It is not question here to discuss *in-extenso* on OSOSO. We would like only to examine the main “ingredients” in doing this. When starting out on an Ontology - Project , the main task is to find / design a suitable Ontology-Software-Tool. [36–43] The Ontology - Tools can be clustered according to the Tasks they are performing as follows: 1-Ontology Development Tools, 2- Ontology Merge and Integration Tools, 3- Ontology Evaluation Tools, 4- Ontology-based Annotation Tools, 5- Ontology Storage and Querying Tools, 6- Ontology Learning Tools. So **OSOSO will be Task-Oriented**.

The main important issues to consider when choosing to Use or to Build an Ontology-Environment are:

- Collaboration and Distributed Workforce Support
- Platform Interoperability
- Scale (Scalability in Size of Ontologies as well as number of simultaneous Users)
- Versioning (ability to support many Versions of Ontologies)
- Security (to support both Read and Write Access)
- Analysis (to support Acquisition, Evolution, and Maintenance of Ontologies)
- Life Cycle Issues (Update and Merge Ontologies over many years)
- Ease of Use (Training materials, Tutorials, Conceptual Modelling Support, Graphic Browsing Tools, etc....)
- Diverse User Support
- Presentation Style (Extensive detail/Pruned information/Abstractions)
- Extensibility (Adaptable along with the Needs of the Users and the Projects, with other Functionalities/Modules)
- Tool Architecture (stand-alone, client/server, n-tier application)
- Ontologies Storage/Retrieval (databases, text files, etc....)

Especially, for *Ontology Editors*, it is mandatory to consider carefully the following features: developers specifications, current release (date), availability (open source, freeware, license), software architecture, extensibility (plug-ins), ontology storage, backup management, interoperability with, imports from/ exports to - languages, KR paradigm of Knowledge model (DL, Frames + FOL, Frames + FOL + Metaclasses), Axiom language (DAML+OIL, Flogic, PAL), built-in Inference engine (FaCT, OntoBroker, PAL), other attached inference engines (JESS,...), constraint/consistency checking, automatic classification, exception handling, graphical taxonomy, graphical prunes (views), Zooms, Collaborative working,

ontologies libraries.

Naturally, all the above have to be done on the General Background – Methodology of Software Optimisation, and Open-source Software Optimisation.

8 The road ahead for S.R. of Vietnam

The International Semantic Web Conference SWC’ 2005 [16–19] is an event to point out the MegaTrends to pay attention for. But, in our SR. of Vietnam, we should be aware of the existing - since long – and growing Digital Gap, either in Internet Industry, Content Industry, as well as in Natural (Vietnamese) Language Generation (NLG), more and more important in the emerging Semantic Web Industry (SWI), as a sine-qua-non Reliable Companion for the Globalised e-Commerce, a “must” for Vietnam’s Integration into AFTA-APEC-ASEM-WTO: *to be “integrated” or not to be!*

“**Things to be done diligently**” (if not immediately) could be:

- 1) Vietnamese Wordnet/SENSUS as a Lexical Upper Ontology based on, for example, Princeton Wordnet/SENSUS, Chinese Wordnet/ SENSUS the IEEE’s SUO (Standard Upper Ontology) and CLO (Taiwan’s Chinese Lexical Ontology), because of the Sino-Vietnamese Radicals/Terminology in many important Domains in Vietnamese language, eg. scientific terminology.
- 2) Education and Training in Digital Content Creation Technology and Engineering (DCCTE), learning from the Japanese 10-year DCC National Programme, since about 1997, including Automatic Intelligent Translation in MultiLingual Internet [(MLI), CICC, Tokyo, Japan], especially for Asian Languages, thus Vietnamese included.
- 3) Education-Training-Research (or R&D) in Cognitive Sciences, especially Cognitics, as a Knowledgeability Improving Synergy (KIS) between the Leading Research Institutes: Mathematics, Informatics, Physics, Automatics, Linguistics, of the Vietnamese Academy of Science & Technology (VAST), and Vietnam National Universities, from North to South.
- 4) Promotion of Regional and International Cooperation/Collaboration in the above-mentioned Domains, especially with Native Vietnamese, for Adequate and Appropriate Technology Transfer (e.g. TOKTEN), in order not to “reinvent the wheel” too often, and modestly “be good pupil to become master ,” and this - unfortunately - even in Natural

- Vietnamese Language Analysis and Generation.
- 5) Last-but-not-least, to REVIVIFY the Scientific and Technological Terminology R&D National Programme !

The Knowledgeability Battle is ahead: to fight for surviving or not to fight ? (Only the Knowledgeable can survive!)

The (two) authors have compiled, South and North, a growing (Knowledge-) Repository on the above-discussed subjects, becoming aware of the Scope and related Efforts to be done and promoted. But at least, “Three Trees to make a (small) Mountain,” so we would like respectfully to make a “Call for Cooperation/Collaboration,” at least in Education & Training of young Vietnamese Master-level Students for the Up-to-date State-of-the-Art of an emerging Vietnamese Cognitics.

9 As a concluding remark

This paper overviewed the Ontology Optimisation issue in the current literature, to assess at what extent how related problems have been solved. We noticed that there was not yet a good “consensus” on Concepts, Problematics and Methodology, as well a lack of adequate Formalism as usual in classical Operations Research. Furthermore, we introduced our own viewpoint and a tentative of optimisation formalism. Although this has been discussed on a specific SEKT-UKARL use case, it is encouraging to continue the research experiment to explore further in this potentially important field, expecting some useful findings for the emerging Semantic Web Deployment issues.

References

- [1] T. M. Dung, “Report Project Mitani 1: Vietnamese Search Engine,” VNU-HCMC, Nov / 2003.
- [2] T.M. Dung and D.H. Cuong, “Designing & Developing a Search Engine by Vietnamese keywords,” *J. Science & Technology Development, Vietnam National University – HCM City*, vol.7, no.4&5, pp.83–88, 2004.
- [3] <http://fr.wikipedia.org/wiki/cognitique>
- [4] http://fr.wikipedia.org/wiki/Jean-Michel_TRUONG
- [5] <http://fr.wikipedia.org/wiki/IdC>
- [6] <http://www.-u-bordeaux2.fr>
- [7] <http://www-eirisco.onecert.fr/>
- [8] <http://www.u-ordeaux3.fr/Actu/colloques/2004/25-22052005.htm>
- [9] <http://repositories.cdlib.org/brie/BRIEWP137/>
- [10] <http://www.intel.com/pressroom/>
- [11] <http://www.intel.com/update/departments/desktop/dt03001.pdf>
- [12] <http://www.euro6ix.net>
- [13] <http://jmichelcornu.free.fr/textes/2003/200304dnac.CR4.html>
- [14] <http://lesrapports.ladocumentationfrancaise.fr/BRP/044000022/0000.pdf>
- [15] <http://www.telecom.gouv.fr/rnrt/reseau/doc/pdf/etude-roadmap-rapport.pdf>
- [16] <http://iswc2002.semanticweb.org>
- [17] <http://iswc2003.semanticweb.org>
- [18] <http://iswc2004.semanticweb.org>
- [19] <http://iswc2005.semanticweb.org>
- [20] <http://realizat.teknowledge.com/IJCAI01/#reference>
- [21] <http://www.aiai.ed.ac.uk/project/euroknow/>
- [22] <http://www.teknowledge.com/>
- [23] <http://suo.ieee.org/SUO/SUO-4D/index.html>
- [24] <http://www.cyc.com/SUO/opencyc-ontology.txt>
- [25] <http://suo.ieee.org/SUO/SUMO/index.html>
- [26] <http://suo.ieee.org/SUO/LatticeOfTheories/index.html>
- [27] <http://suo.ieee.org/SUO/resolutions.html#LatticeOfTheories>
- [28] <http://meganesia.ict.gu.edu.au/~phmartin/WebKB/doc/MSO.html>
- [29] <http://suo.ieee.org/SUO/SUO-Procedures/index.html>
- [30] <http://suo.ieee.org/SUO/Ontology.refs.html>
- [31] <http://suo.ieee.org/SUO/links.html>
- [32] <http://suo.ieee.org/SUO/expertcontributions.html>
- [33] <http://suo.ieee.org/SUO/KIF/index.html>
- [34] <http://suo.ieee.org/SUO/KIF/suo-kif.html>
- [35] <http://www-user.uni-bremen.de/~bateman>
- [36] http://www.xml.com/2002/11/06/Ontology_Editor_Survey.html
- [37] http://www.xml.com/2004/07/14/examples/Ontology_Editor_Survey_2004_Table_Michael_Denny.pdf
- [38] <http://www.rdf-logic@w3.org>
- [39] <http://www.bell-lab.com/topic/swdist/>
- [40] <http://www.gmod.org/argos/argos.shtml>
- [41] <http://www.oreilly.com.html>
- [42] EGTDC Software Documentation Pages: **Protege on BioLinux**
envgen.nox.ac.uk/envgen/software/archives/000487.html
- [43] EGTDC Software Documentation Pages: 10.)
Ontology Software ...
envgen.nox.ac.uk/envgen/software/archives/cat_10_ontology_software.html
- [44] N. D. Ngoc, The Future of Automation Technology , VICA VI-2005, April 12–14, Hanoi, ThangLong Univ., REV-VAIP- IEEE&IEEE ComSoc.
- [45] T.M. Dung, Multilingual Ontology-based Upgrading Extensions of an OracleText--based Search Engine, FIT/UNS/VNU-HOCHIMINH-City, preprint. (submitted to FAIR 2005, HCMC, Vietnam).
- [46] T. M. Dung, Towards an Internet 3G/4G-based University Semantic Web Portal, FIT/UNS/VNU at HOCHIMINH-City, preprint, (submitted to FAIR 2005, HCMC, Vietnam).

- [47] L.A. Zadeh, “From search engines to question-answering systems—The role of fuzzy logic”, *Prog. Informatics*, no. 1, pp. 1–3, 2005.

ANNEX

S. STAAB and R. STUDER(Eds.), *Handbook on Ontologies*, International Handbooks on Information Systems, Springer 2004, ISBN 3-540-40834-7.

Table of Contents :

- I. Ontology Representation and Reasoning:
 1. Description Logics
 2. Ontologies in F-Logic
 3. The Resource Definition Framework (RDF) and its Vocabulary Description Language RDFS
 4. Web Ontology Language OWL
 5. An Ontology Composition Algebra.
- II. Ontology Engineering :
 6. On-To-Knowledge Methodology (OTKM)
 7. Building a Very Large Ontology from Medical Thesauri
 8. An Overview of OntoClean
 9. Ontology Learning
 10. Knowledge Patterns
 11. Ontology and the Lexicon
 12. Ontology Reconciliation
 13. Ontology Evaluation
 14. Ontology Engineering Environments
- III. Ontology Infrastructure:
 15. An Extensible Ontology Software Environment
 16. Ontologies in Support of Problem Solving
 17. Ontologies in Agent Architectures
 18. Tools for Mapping and Merging Ontologies
 19. Ontology Matching
 20. Retrieving and Exploring Ontology – based Information
 21. Supporting User Tasks through Visualisation of Light-Weight Ontologies.
- IV. Ontology Applications:
 22. Ontologies for Knowledge Management

23. Ontology-based Content Management in a Virtual Organization
24. Ontology-based Recommender Systems
25. The Knowledge Portal “OntoWeb”
26. Ontologies and HyperText
27. Semantic Layering with Mag-Pie
28. Ontology and Metadata for e-Learning
29. Ontology and the Process Specification Language
30. The Role of Ontologies in e-Commerce
31. An Ontology-based Platform to Semantic Inter-Operability
32. Ontologies in BioInformatics



Nguyen Dinh NGOC

Nguyen Dinh NGOC graduated from ENSMH, ENSGM, ENST; IPG(PhD in Geophysics), IHP (DrSc. in Mahematical Sciences). [Paris,France], Senior member of IEEE&IEEE ComSoc. he is Past VP of VAIP, Present VP of REV, and Full Professor-

Member of the Board of ThangLong University, Hanoi, SR of Vietnam, (Major General- retired).



Truong My DUNG

Truong My DUNG graduated from the University of Saigon, the Institute of Mathematics/VNASt (PhD in Mathematics). She is Vice-Dean of the Faculty of ICT/ Univ. of Natural Sciences/ The National University in HOCHIMINH-City,

SR.of Vietnam, Member of the Central Boards of VAIP(SRVN), and HCA(HCM-City).