Gloria Phillips-Wren
Lakhmi C. Jain
Kazumi Nakamatsu
Robert J. Howlett
Editors

SMART INNOVATION, SYSTEMS AND TECHNOLOGIES 4



Advances in Intelligent Decision Technologies

Proceedings of the Second KES International Symposium IDT 2010

Springer

Gloria Phillips-Wren, Lakhmi C. Jain, Kazumi Nakamatsu, and Robert J. Howlett (Eds.)

Advances in Intelligent Decision Technologies

Smart Innovation, Systems and Technologies 4

Editors-in-Chief

Further volumes of this series can be found on our homepage: springer.com

Vol. 1. Toyoaki Nishida, Lakhmi C. Jain, and Colette Faucher (Eds.)
Modeling Machine Emotions for Realizing Intelligence, 2010
ISBN 978-3-642-12603-1

Vol. 2. George A. Tsihrintzis, Maria Virvou, and Lakhmi C. Jain (Eds.)
Multimedia Services in Intelligent Environments –
Software Development Challenges and Solutions, 2010
ISBN 978-3-642-13354-1

Vol. 3. George A. Tsihrintzis and Lakhmi C. Jain (Eds.)
Multimedia Services in Intelligent Environments –
Integrated Systems, 2010
ISBN 978-3-642-13395-4

Vol. 4. Gloria Phillips-Wren, Lakhmi C. Jain,
Kazumi Nakamatsu, and Robert J. Howlett (Eds.)
Advances in Intelligent Decision Technologies –
Proceedings of the Second KES International
Symposium IDT 2010, 2010
ISBN 978-3-642-14615-2

Gloria Phillips-Wren, Lakhmi C. Jain,
Kazumi Nakamatsu, and Robert J. Howlett (Eds.)

Advances in Intelligent Decision Technologies

Proceedings of the Second KES International Symposium IDT 2010

Springer

Prof. Gloria Phillips-Wren
Sellinger School of Business and Management
Loyola University Maryland
4501 N.Charles Street
Baltimore. MD 21210
USA
Email: gwren@loyola.edu

Prof. Kazumi Nakamatsu
School of Human Science and Environment
University of Hyogo
1-1-12 Shinzaike-honcho
Himeji, Hyogo 670-0092
Japan
Email: Nakamatsu@shse.u-hyogo.ac.jp

Prof. Dr. Lakhmi C. Jain
School of Electrical and
Information Engineering
University of South Australia
Adelaide
Mawson Lakes Campus
South Australia SA 5095
Australia
Email: Lakhmi.jain@unisa.edu.au

Prof. Robert J. Howlett
KES International
P.O. Box 2115
Shoreham-by-Sea BN43 9AF
UK
Email: rjhowlett@kesinternational.org

ISBN 978-3-642-14615-2 e-ISBN 978-3-642-14616-9

DOI 10.1007/978-3-642-14616-9

Smart Innovation, Systems and Technologies ISSN 2190-3018

Library of Congress Control Number: 2010930917

Typesetting: Scientific Publishing Services Pvt. Ltd., Chennai, India.

Printed on acid-free paper

9 8 7 6 5 4 3 2 1

springer.com

Foreword

KES International (KES) is a worldwide organisation that provides a professional community and association for researchers, originally in the discipline of Knowledge Based and Intelligent Engineering Systems, but now extending into other related areas. Through this, KES provides its members with opportunities for publication and beneficial interaction.

The focus of KES is research and technology transfer in the area of Intelligent Systems, i.e. computer-based software systems that operate in a manner analogous to the human brain, in order to perform advanced tasks. Recently KES has started to extend its area of interest to encompass the contribution that intelligent systems can make to sustainability and renewable energy, and also the knowledge transfer, innovation and enterprise agenda.

Involving several thousand researchers, managers and engineers drawn from universities and companies world-wide, KES is in an excellent position to facilitate international research co-operation and generate synergy in the area of artificial intelligence applied to real-world 'Smart' systems and the underlying related theory.

The KES annual conference covers a broad spectrum of intelligent systems topics and attracts several hundred delegates from a range of countries round the world. KES also organises symposia on specific technical topics, for example, Agent and Multi Agent Systems, Intelligent Decision Technologies, Intelligent Interactive Multimedia Systems and Services, Sustainability in Energy and Buildings and Innovations through Knowledge Transfer. KES is responsible for two peer-reviewed journals, the International Journal of Knowledge based and Intelligent Engineering Systems, and Intelligent Decision Technologies: an International Journal.

KES supports a number of book series in partnership with major scientific publishers.

Published by Springer, 'Smart Innovative Systems and Technologies' is the KES flagship book series. The aim of the series is to make available a platform for the publication of books (in both hard copy and electronic form) on all aspects of single and multi-disciplinary research involving smart innovative systems and technologies, in order to make the latest results available in a readily-accessible form.

The series covers systems that employ knowledge and intelligence in a broad sense. Its focus is systems having embedded knowledge and intelligence, which may be applied to the solution of world industrial, economic and environmental problems and the knowledge-transfer methodologies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of scientific and technological disciplines.

Examples of applicable areas to be covered by the series include intelligent decision support, smart robotics and mechatronics, knowledge engineering, intelligent multi-media, intelligent product design, intelligent medical systems, smart industrial products, smart alternative energy systems, and underpinning areas such as smart systems theory and practice, knowledge transfer, innovation and enterprise.

The series includes conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes take responsibility for ensuring that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Professor Robert J. Howlett
Executive Chair, KES International
Visiting Professor, Enterprise: Bournemouth University
United Kingdom

Preface

Intelligent Decision Technologies (IDT) seeks an interchange of research on intelligent systems and intelligent technologies which enhance or improve decision making in industry, government and academia. The focus is interdisciplinary in nature, and includes research on all aspects of intelligent decision technologies, from fundamental development to the applied system.

The field of intelligent systems is expanding rapidly. Advances in artificial intelligence (AI) and connectivity have delivered exciting new applications. Networks have integrated the Internet and wireless technologies to enable communication and coordination between dispersed systems. Intelligent decision making now means that technology assists the human decision maker in everyday tasks and complex environments. The field of intelligent decision systems is interdisciplinary in nature, bridging computer science with its development of artificial intelligence, information systems with its development of decision support systems, and engineering with its development of technology.

It is therefore an honor to publish the research of scholars from the Second KES International Symposium on Intelligent Decision Technologies (KES IDT'10), hosted and organized by the Sellinger School of Business and Management, Loyola University Maryland, USA, in conjunction with KES International. The book contains chapters based on papers selected from a large number of submissions for consideration for the symposium from the international community. Each paper was peer reviewed by at least two independent referees. The best papers were accepted based on recommendations of the reviewers and after required revisions had been undertaken by the authors. The final publication represents the current leading thought in intelligent decision technologies.

We wish to express our sincere gratitude to the plenary speakers, invited session chairs, delegates from all over the world, the authors of various chapters and reviewers for their outstanding contributions. We express our sincere thanks to Dr. Karyl Leggio, Dean of the Sellinger School, and to Loyola University Maryland for their sponsorship and support of the symposium. We would like to thank Peter Cushion of KES International for his help with organizational issues. We thank the editorial team of Springer-Verlag and Heather King for their support in production of this volume. We sincerely thank Jean Anne Walsh, Katlyn Good, Jessica Ross, Brian Hatcher, Pat Donohue and students (Nathan Hill, Mary Kiernan, Pat Moran, JT Laue) at Loyola University Maryland for their assistance.

We hope and believe that this volume will contribute to ideas for novel research and advancement in intelligent decision technologies for researchers, practitioners, professors and research students who are interested in knowledge-based and intelligent engineering systems. We invite you to join us at a future symposium.

Baltimore, Maryland, USA
July 28-30, 2010

Gloria Philips-Wren
Kazumi Nakamatsu
Lakhmi C. Jain
Robert J. Howlett

KES IDT 2010 Organization

Symposium Co-chairs

Professor Gloria Phillips-Wren	Loyola University Maryland, USA
Professor Lakhmi C. Jain	University of South Australia, Australia

General Session Chair

Professor Kazumi Nakamatsu	University of Hyogo, Japan

Executive Chair

Professor Robert J. Howlett	Bournemouth University, KES International, UK

KES IDT 2010 International Program Committee Members

Adrian Barb	USA
Ana Respicio	Portugal
Anca Ralescu	USA
Andreas Tolk	USA (Country Liaison)
Angel P. del Pobil	Spain
Anne Hakansson	Sweden (Country Liaison)
Anne Snowden	Canada (Country Liaison)
Bala M. Balachandran	Australia
Barbara Catania	Italy
Beniamino Murgante	Italy
C.P. Lim	Malaysia (Country Liaison)
Carlo Sansone	Italy
Chei-Chang Chiou	Taiwan
Chia-Feng Juang	Taiwan
Chih-Yung Chang	Taiwan
Chung-Nan Lee	Taiwan
Daniel Power	USA
Daniela Godoy	Argentina

Despina Filippidis	Australia
Eizo Kinoshita	Japan
Elena Zudilova-Seinstra	Netherlands
Enrique H. Viedma	Spain (Country Liaison)
Eyke Hüllermeier	Germany
Feng-Tse Lin	Taiwan
Frada Burstein	Australia
Francisco Chiclana	UK
Fumiaki Takeda	Japan
G.A.Vijayalakshmi Pai	India
George A. Tsihrintzis	Greece
Georgia D. Tourassi	USA
Gregoris Mentzas	Greece
Hisao Shiizuka	Japan
Hu Xiangpei	China
Huey-Ming Lee	Taiwan
Ines Couso	Spain
Isabelle Bichindaritz	USA
Jaeseok Choi	Korea
Jaeseok Choi	Korea
Jens Pohl	USA
Jian-Bo Yang	UK
Jie Lu	Australia (Country Liaison)
Jose Manuel Ferrandez	Spain
Juan Antonio Botía Blaya	Spain
Junzo Watada	Japan (Country Liaison)
Kazuo Hatakeyama	Brazil
Leonardo Garrido	Mexico
Li-Ling Hung	Taiwan
Lihui Chen	Singapore
Lily Lin	Taiwan
Lluís Belanche	Spain
Luiz Fernando Capretz	Canada
Luís Moniz Pereira	Portugal (Country Liaison)
Manuel Mora	Mexico (Country Liaison)
Margarita Sordo	USA
Maria Virvou	Greece

Martin R. Stytz	USA
Matilde Santos Peñas	Spain
Maurice Mulvenna	UK
Mika Sato-Ilic	Japan
Miroslav Karny	Czech Republic (Country Liaison)
P.C. Yuen	Hong Kong
Pascale Zarate'	France
Patrick Brenzillon	France
Pedro Paulo Balestrassi	Brazil
Pei Liu	Taiwan
Pierre Beauseroy	France
Pierre Morizet-Mahoudeaux	France
Robert Cierniak	Poland
Ronald Hartung	USA
Roumen Kountchev	Bulgaria (Country Liaison)
Roy Rada	USA
Shangming Zhou	UK
Sheila B. Banks	USA
Stefania Montani	Italy
Thomas Sudkamp	USA
Thong Nguyen	Australia
Toyohide Watanabe	Japan
Valentina Emilia Balas	Romania (Country Liaison)
Vicenc Torra	Spain
Yukio Ohsawa	Japan

Invited Session Chairs

Anne Snowdon	Canada
Lakhmi Jain	Australia
Eizo Kinoshita	Japan
Manuel Mora, Miguel Angel-Sicilia	Mexico
Toyohide Watanabe, Koichi Asakura, Shigeki Matsubara	Japan
Lois (Yu-Cheng) Tang	Taiwan
Xiangpei Hu	China
Robert Cierniak	Poland
Bala M. Balachandran, Masoud Mohammadian, Dharmendra Sharma	Australia
Hisao Shiizuka, Hiroko Shoji	Japan

Kazumi Nakamatsu	Japan
Junzo Watada	Japan
Beniamino Murgante	Italy
Mika Sato-Ilic	Japan

Local Organizing Committee

Nathan Hill
Mary Kiernan
Pat Moran
JT Laue

Sponsoring Institutions

Loyola University Maryland, USA

Sellinger School of Business and Management, Loyola University Maryland, USA

KES International

Table of Contents

I. Keynote Papers

II. Decision Making Theory

III. Advances in Intelligent Decision Systems

IV. Intelligent Decision Technologies in Accounting and Finance

V. Optimization-Based Intelligent Techniques in Image Processing

VI. E-commerce and Logistics Management

VII. Intelligent Spatial Decision Analysis

VIII. Using Intelligent Systems for Decision Support in Health Systems

IX. Ontology-Based KMS and DMSS for Service Systems

X. Service-Oriented Innovation for Designing Intelligent Environment

XI. Applying Intelligent Decision Technology

XII. Soft Data Analysis Based Fuzzy Systems, Control and Decision Making

XIII. Kansei Communication and Value Creation in Human Mind

XIV. Future Direction of Innovative Decision Technologies

Intelligence Analysis as Agent-Assisted Discovery of Evidence, Hypotheses and Arguments

Gheorghe Tecuci[1], David Schum[1], Mihai Boicu[1], Dorin Marcu[1], and Benjamin Hamilton[2]

[1] Learning Agents Center, George Mason University, Fairfax, VA, USA
{tecuci,dschum,mboicu,dmarcu}@gmu.edu
[2] Department of Defense, USA

Abstract. This paper presents a computational approach to intelligence analysis which is viewed as mixed-initiative discovery of evidence, hypotheses and arguments by an intelligence analyst and a cognitive assistant. The approach is illustrated with the analysis of wide area motion imagery of fixed geographic locations where the goal is to discover threat events such as an ambush or a rocket launch. This example is used to show how the Disciple cognitive assistants developed in the Learning Agents Center can help the analysts in coping with the astonishing complexity of intelligence analysis.

Keywords: intelligence analysis, science of evidence, wide-area motion imagery, discovery, cognitive assistants, learning, evidence-based reasoning, mixed-initiative reasoning.

1 Introduction

Problem-solving and decision-making depends critically on accurate intelligence that needs to be discovered in an overwhelming amount of mostly <u>irrelevant</u>, always <u>incomplete</u>, usually <u>inconclusive</u>, frequently <u>ambiguous</u>, and commonly <u>dissonant</u> information with various degrees of <u>believability</u> about a highly complex and dynamic world. This is an astonishingly complex process where each analytic task is <u>unique</u> and always requires mixtures of <u>imaginative and critical reasoning</u>. Indeed, hypotheses about situations of interest must be generated by imaginative thought and then subjected to critical evidence-based analysis.

We are researching a <u>computational theory of intelligence analysis</u> which forms the basis for the development of cognitive assistants that help intelligence analysts in coping with this complexity. Part of this theory is a view of intelligence analysis as mixed-initiative discovery of evidence, hypotheses and arguments by intelligence analysts (who are capable of imaginative reasoning, have broad subject matter expertise, and have access to evidence from a wide variety of sources) and their cognitive assistants (that are capable of critical reasoning and have both domain-specific knowledge and general knowledge from the <u>Science of Evidence</u>).

In the next section we present a sample intelligence analysis problem (analysis of wide-area motion imagery) that will be used to illustrate the developed approach. After that, the following five sections present the processes of discovery of hypotheses, evidence and arguments. Then, section 8 concludes the paper with a discussion on how the Disciple cognitive assistants developed in the Learning Agents Center, which

G. Phillips-Wren et al. (Eds.): Advances in Intel. Decision Technologies, SIST 4, pp. 1–10.
springerlink.com

are capable of analytic assistance, learning, and tutoring, can help in coping with the astonishing complexity of intelligence analysis [14, 15].

2 Sample Problem: Analysis of Wide-Area Motion Imagery

Capabilities exit today to persistently monitor fixed geographic locations (such as conflict areas) as wide as 100 km^2, for long periods of time, using electro-optic sensors (see Fig. 1). This leads to the collection of huge amounts of data to be used either in real-time analysis or in forensic analysis. During real-time analysis, analysts attempt to discover impeding threat events (e.g., ambush, kidnapping, rocket launch, false check-point, suicide bomber, IED) in time to react. During forensic analysis, the analysts backtrack from such an event (e.g., an ambush) in order to discover the participants, possible related locations and events, and the specific movement patterns [2]. The problem however is that the manual analysis of these huge amounts of data would require thousands of analysts.

We will use this sample analysis problem to illustrate our approach.

Fig. 1. Wide area motion imagery (WAMI).

3 Discovery of Hypotheses

Let us consider an analyst who, while reviewing wide area motion imagery (WAMI) of a region of Iraq, notices evidence of road work at 1:17am, an unusual time for such an activity. The question is: *What possible threat does this evidence suggest?* Through a flash of insight, the analyst may abductively leap to the hypothesis H_k that there is an ambush threat at that location [6]. Attempting to justify the hypothesis, the analyst may generate the following abductive inference steps shown also in the left hand side of Fig. 2 (as we know, abductive inference indicates that something is possibly true):

E^*_i: *There is evidence of road work at 1:17am at location L_1.*
→ E_i: *It is possible that there is indeed road work at location L_1.*
→ H_a: *It is possible that the road work is for blocking the road*
→ H_c: *It is possible that there is an ambush preparation at location L_1.*
→ H_k: *It is possible that there is an ambush threat at location L_1.*

So here we have evidence in search of hypotheses where a newly discovered item of evidence searches for hypotheses that would explain it.

4 Discovery of Evidence

A great challenge in any intelligence analysis task is the massive amount of data that needs to be searched quickly, especially during the real-time use of the system. The diagram in the middle of Fig. 2 illustrates the deductive process involved in putting the generated hypothesis at work to guide the search for new relevant evidence in the WAMI data. The question is: *Assuming that the threat is real, what other events or entities should be observable?* The deductive reasoning process for answering this question successively reduces the assessment of the top-level hypothesis H_k to the assessment of simpler hypotheses, ultimately resulting in precise queries to be answered from the WAMI data (as we know, deductive inference indicates that something is necessarily true):

Let us assume H_k, *that there is an ambush threat at location L_1 after 1:17am.*
→ H_b: *L_1 should be an ambush location.*
H_c: *There should be ambush preparation at L_1 around 1:17am.*
H_q: *There should be ambush execution at L_1 (if forensic analysis).*

If this is real-time analysis occurring soon after 1:17am, then the ambush has not yet been executed and the third sub-hypothesis (H_q) will not be considered. However, if this is forensic analysis, then H_q should also be considered.

Let us now assume H_b, *that L_1 is indeed an ambush location.*
→ H_d: *L_1 should be on a route of the blue forces after 1:17am.*
H_e: *There should be cover at location L_1.*

This guides the analyst to search for the following evidence:

- *Search for evidence that L_1 is on a planned blue route after 1:17am.*
- *Search for evidence in the WAMI data that there is cover at location L_1.*

A similar analysis of the hypothesis H_c (*There is an ambush preparation at L_1 around 1:17am*) leads to the following queries for specific events and entities in the WAMI data and from other sources (shown as shaded circles in Fig. 2):

- *Search for evidence in the WAMI data that there is departure of vehicle V_1 from facility F_1 before 1:17am.*
- *Search for evidence that F_1 is a suspected terrorist facility.*

- *Search for evidence in the WAMI data that there is arrival of vehicle V_1 at location L_1 short before 1:17am.*
- *Search for evidence in the WAMI data that personnel P_1 descends from vehicle V_1 at location L_1 short before 1:17am.*

Notice that these are precise queries that can be answered very fast. Being based on evidence, the answers will be probabilistic, such as:

It is almost certain that there is arrival of vehicle V_1 at location L_1 at 1:09am.
It is very likely that personnel P_1 descends from vehicle V_1 at L_1 at 1:10am.

These probabilistic solutions and other discovered evidence will be used to assess the likelihood of the top level hypothesis H_k, as discussed in Section 5.

The above has illustrated the deductive process of hypotheses in search of evidence that leads to the discovery of new evidence that may favor or disfavor them.

Some of the newly discovered items of evidence may trigger new hypotheses or the refinement of the current hypothesis. Therefore, as indicated at the bottom of Fig. 2, the processes of evidence in search of hypotheses and hypotheses in search of evidence take place at the same time, and in response to one another.

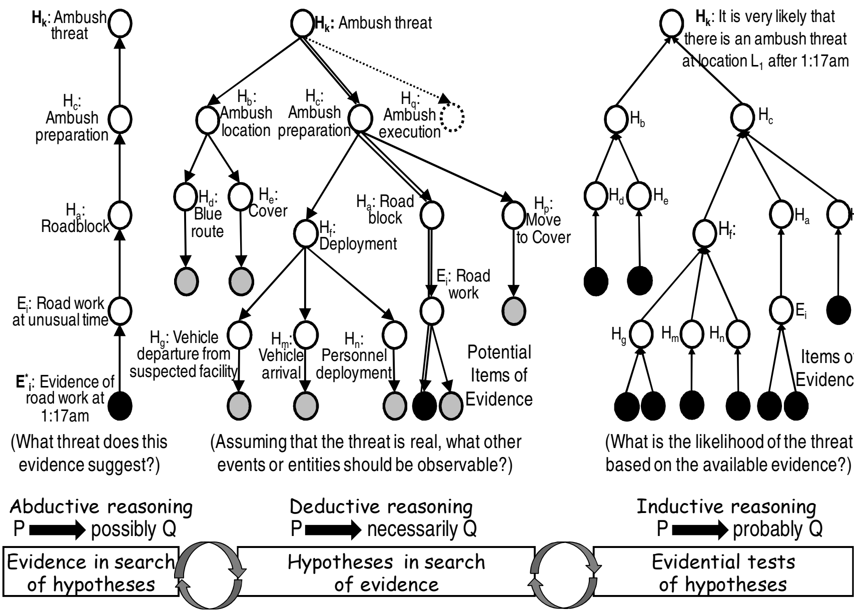

Fig. 2. Discovery of evidence, hypotheses and arguments.

5 Discovery of Arguments

The discovered evidence (shown as black circles at the right hand side of Fig. 2) can now be used to discover an argument that assesses, through inductive inference, the likelihood of the hypothesis H_k (e.g.,*"It is very likely that there is an ambush threat at* *location L_1 after 1:17am"*). As we know, inductive inference indicates that something is probably true.

Fig. 3 shows a Wigmorean probabilistic inference network that combines the deductive reasoning tree and the inductive reasoning tree from Fig. 2. This network has a well-defined structure, which has a grounding in the problem reduction representations developed in Artificial Intelligence [4, 10], and in the argument construction methods provided by the noted jurist John H. Wigmore [17], the philosopher of science Stephen Toulmin [16], and the evidence professor David Schum [5, 7]. This approach uses expert knowledge and evidence to successively reduce a complex hypothesis analysis problem to simpler and simpler problems, to find the solutions of the simplest problems, and to compose these solutions, from bottom-up, to obtain the solution of the initial problem. The Wigmorean network shows how evidence is linked to hypotheses through arguments that establish the relevance, believability and inferential force or weight of evidence [5, 9].

As shown in Fig. 3, the assessment of hypothesis H_k is reduced to the assessment of two simpler hypotheses: H_b and H_c. Then H_b is reduced to H_d and H_e. Each of these two hypotheses is assessed by considering both favoring evidence and disfavoring evidence. Let us assume that there are two items of favoring evidence for H_d: E_{d1} and E_{d2}. For each of them one would need to assess the extent to which it favors the hypothesis H_d. This requires assessing the relevance, believability, and inferential force or weight of evidence.

Relevance answers the question: *So what? How does this item of information bear on what the analyst is trying to prove or disprove?*

Believability (or credibility) answers the question: *Can we believe what this item of intelligence information is telling us?*

Inferential force or weight answers the question: *How strong is this item of relevant evidence in favoring or disfavoring various alternative hypotheses or possible conclusions being entertained?*

Let us assume the following solutions for the relevance and the believability of E_{d1}: *"If we believe E_{d1} then H_d is almost certain"* and *"It is likely that E_{d1} is true."*

In this example, almost certain and likely are symbolic probabilities for likelihood similar to those from the DNI's standard estimative language, but other scales for uncertainty can easily be used [18].

The relevance of E_{d1} (almost certain) is combined with its believability (likely), for example through a "min" function, to determine E_{d1}'s inferential force or weight on H_d: *"Based on E_{d1} it is likely that H_d is true."*

Similarly one assesses the inferential force of E_{d2} on H_d: *"Based on E_{d2} it is almost certain that H_d is true."*

By composing the above solutions (e.g., through "max") one assesses the inferential force of the favoring evidence (i.e., E_{d1} and E_{d2}) on H_d: *"Based on the favoring evidence it is almost certain that H_d is true."*

Similarly one assesses the inferential force of the disfavoring evidence on H_d: *"Based on the disfavoring evidence it is unlikely that H_d is false."*

Now because there is very strong evidence favoring H_d and there is weak evidence disfavoring H_d, one concludes: *"It is almost certain that H_d is true."*

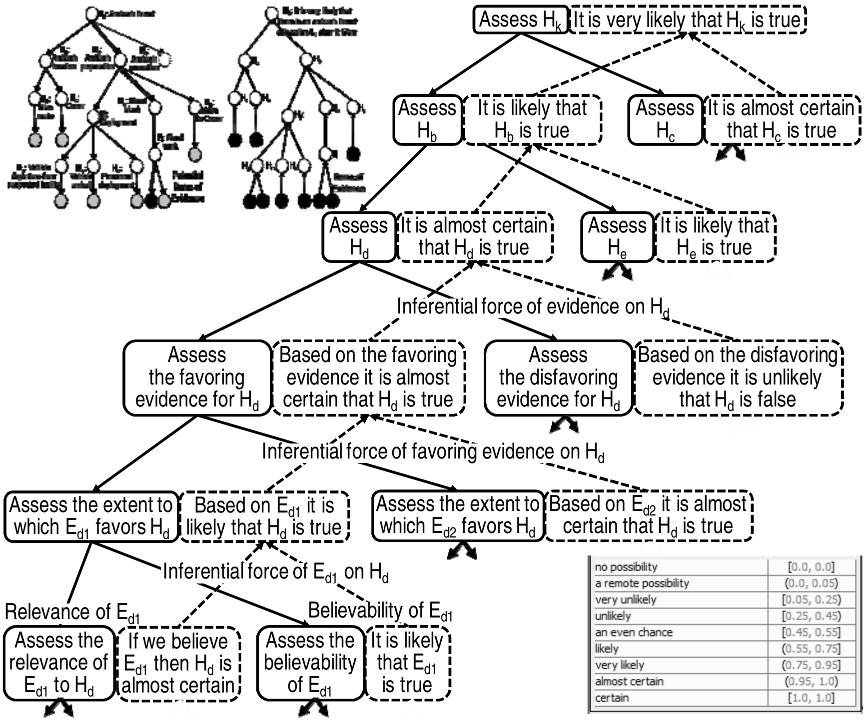

Fig. 3. Wigmorean probabilistic inference network for hypothesis assessment.

H_e is assessed in a similar way: *"It is likely that H_e is true."* Then the assessments of H_d and H_e are composed (through "min") into the assessment of H_b: *"It is likely that H_b is true."* Finally, this assessment is composed with the assessment of H_c (*"It is almost certain that H_c is true."*), through "average", to obtain the assessment of H_k (*"It is very likely that* H_k *is true."*)

6 Believability of Evidence

Above we have discussed the process of evidence-based hypothesis assessment down to the level where one has to assess the relevance and the believability of an item of evidence. In this section we will show how a Disciple agent helps in assessing the believability of evidence. This is based on its stock of established knowledge about evidence, its properties, uses, and discovery from the emerging Science of Evidence [1, 5, 7, 8], which is itself based upon 700 years of experience in the Anglo-American system of law. For example, the right-hand side of Fig. 4 shows a substance-blind classification of recurrent forms and combinations of evidence based, not on substance or content, but on the inferential properties of evidence [9].

This classification is important because each type of evidence has specific believability credentials, as well as a well-defined procedure for assessing its believability, as shown in the left hand side of Fig. 4.

In this classification, wide area motion imagery is demonstrative tangible evidence (i.e., a representation or image of a tangible thing), which has three believability attributes: authenticity, reliability, and accuracy.

Authenticity addresses the question: Is this object what it is represented as being or is claimed to be?

Reliability is especially relevant to various forms of sensors that provide us with many forms of demonstrative tangible evidence. A system, sensor, or test of any kind is reliable to the extent that the results it provides are repeatable or consistent. For example, a sensing device is reliable if it provides the same image or report on successive occasions on which this device is used.

Finally, the accuracy concerns the extent to which the device that produced the representation of the real tangible item had a degree of sensitivity (resolving power or accuracy) that allows us to tell what events were observed.

For testimonial evidence we have two basic sources of uncertainty: the competence and the credibility of the source (see bottom left-side of Fig. 4). Competence involves access and understandability. Credibility involves veracity (or truthfulness), objectivity, and observational sensitivity under the conditions of observation [9].

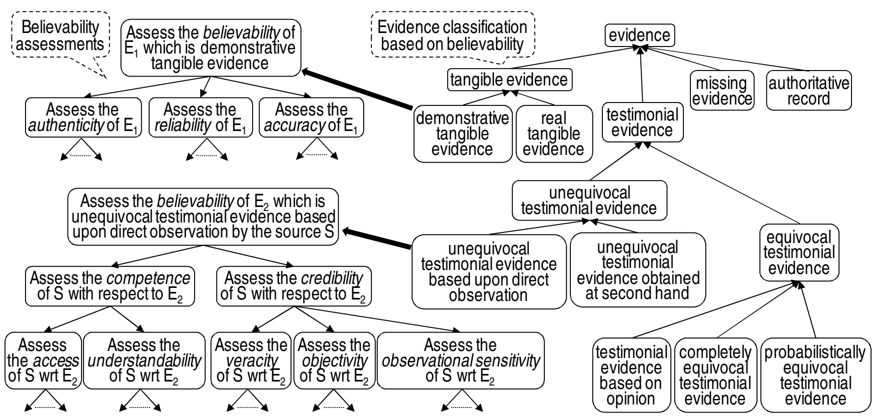

Fig. 4. Types of evidence and their believability assessments.

7 Analysis of Competing Hypotheses

Just because we have evidence of an event (e.g., E^*_i: evidence of road work at 1:17am) does not mean that the event actually occurred. Thus, as indicated in Fig. 5, we need to test two hypotheses: E_i (There is road work …) and Not E_i (There is no road work …). Similarly, for each abduced hypothesis (e.g., H_a: Roadblock), one would need to consider competing hypotheses (e.g., H_{a1}: Road repair). Moreover, for

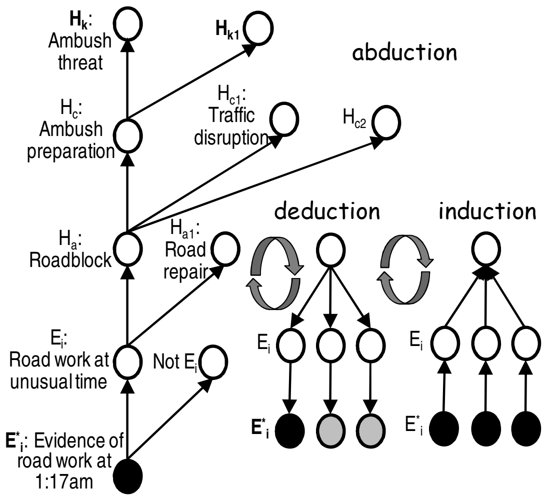

Fig. 5. Analysis of competing hypotheses.

each such competing hypothesis one has to search for relevant evidence and use this evidence to test it, as discussed in the previous sections.

8 Cognitive Assistants for Learning, Teaching, and Analysis

The researched computational theory of intelligence analysis is being implemented in Disciple cognitive assistants that synergistically integrate three complex capabilities. They can rapidly learn the analytic expertise which currently takes years to establish, is lost when analysts separate from service, and is costly to replace. They can tutor new intelligence analysts how to systematically analyze complex hypotheses. Finally, they can assist the analysts in analyzing complex hypotheses, collaborate, and share information [14, 15].

The problem solving engine of a Disciple assistant employs a general divide-and-conquer approach to problem solving, called problem-reduction/solution-synthesis, which was illustrated in Fig. 3. To exhibit this type of problem solving behavior, the knowledge base of the agent contains an ontology which describes both general concepts for evidence-based reasoning (see Fig. 4) and domain-specific concepts from an application domain. The knowledge base also includes a set of learned problem reduction and solution synthesis rules which are represented with the concepts from the ontology. A problem reduction rule expresses how and under what conditions a generic problem can be reduced to simpler generic problems. Reduction rules are applied to automatically reduce assessment problems to simpler problems, as illustrated in Fig. 3. Similarly, a solution synthesis rule expresses how and under what conditions the solutions of generic sub-problems can be combined into the solution of a generic problem. These rules are applied to automatically perform compositions such as those from Fig. 3.

The cognitive assistant also includes a complex learning engine that uses multistrategy methods (e.g., learning from examples, from explanation, and by analogy) to allow a subject matter expert to teach it in a way that is similar to how the expert would teach a person [10, 11, 14]. For instance, the expert will show the agent how to perform an analysis, as it was illustrated in Fig. 2, and will help it to understand each inference step. The agent, on the other hand, will attempt to learn a general reduction and synthesis rule from each such step and will extend its ontology. Moreover, the acquired knowledge will be pedagogically tuned [3], the agent solving new problems and explaining its reasoning similarly to how the expert taught it. This makes the agent an effective tool for teaching new intelligence analysts.

A trained Disciple cognitive assistant can help an analyst cope with the astonishing complexity of intelligence analysis through the use of mixed-initiative reasoning, a type of collaboration between humans and automated agents that mirror the flexible collaboration between people. It consists of an efficient, natural interleaving of contributions by the analyst and the agent that is determined by their relative knowledge and skills and the problem-solving context, rather than by fixed roles, enabling each of them to contribute what it does best, at the appropriate time [12, 13]. The analyst will act as the orchestrator of the reasoning process, guiding the high-level exploration, while the agent will implement this guidance by taking into account the analyst's preferred problem solving strategies, assumptions and biases. For example, the agent discovers evidence in the WAMI data of road work at location L_1, at 1:17am, an unusual time for such an activity, and alerts the analyst. As a result, the analyst directs the agent to analyze the hypothesis that there is an ambush threat and the agent develops the reasoning tree from the middle of Fig. 2, which will guide it to search for additional relevant evidence in the WAMI data and from other sources. The identified evidence is then used by the agent to evaluate the likelihood of the considered hypothesis, as was discussed in Section 5 and illustrated in Fig. 3. This reasoning tree makes very clear the analysis logic, what evidence was used and how, what assumptions have been made, and what is not known. This allows the analyst to critically evaluate the reasoning process, to accept parts of it, to modify other parts, and to produce an analysis which s/he would consider her/his own. The emphasis is on enhancing analyst's creativity, relying on the human to take the most critical decisions, and only to critique and correct the more routine ones that are proposed by the agent.

Acknowledgments. We are very grateful to Phil Hwang, Don Kerr, Joan McIntyre, Kelcy Allwein, Keith Anthony, Cindy Ayers, Susan Durham, Sharon Hamilton, Jim Homer, David Luginbuhl, Bill Nolte, George Stemler, and Ben Wible for their suggestions and support. This research was performed in the Learning Agents Center and was partially supported by several agencies of the U.S. Government, including the National Geospatial-Intelligence Agency, the Department of Defense, and the National Science Foundation (0750461). The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or position of any agency of the U.S. Government.

References

1. Anderson, T., Schum, D.A., Twining, W.: Analysis of evidence. Cambridge U. Press, Cambridge (2005)
2. Desai, M.: Persistent stare exploitation and analysis system, PerSEAS (2009), http://www.darpa.mil/ipto/solicit/baa/BAA-09-55_PIP.pdf (Accessed 18 April 2010)
3. Le, V., Tecuci, G., Boicu, M.: Agent shell for the development of tutoring systems for expert problem solving knowledge. In: Woolf, B.P., Aïmeur, E., Nkambou, R., Lajoie, S. (eds.) ITS 2008. LNCS, vol. 5091, pp. 228–238. Springer, Heidelberg (2008)
4. Nilsson, N.J.: Problem solving methods in artificial intelligence. McGraw-Hill, New York (1971)
5. Schum, D.A.: Evidence and inference for the intelligence analyst. University Press of America, Lanham (1987)
6. Schum, D.A.: Species of abductive reasoning in fact investigation in law. Cardozo Law Rev. 22(5-6), 1645–1681 (2001)
7. Schum, D.A.: The evidential foundations of probabilistic reasoning. Northwestern University Press (2001)
8. Schum, D.A.: Science of Evidence: Contributions from Law and Probability. Law Probab Risk 8, 197–231 (2009)
9. Schum, D.A., Tecuci, G., Boicu, M., Marcu, D.: Substance-blind classification of evidence for intelligence analysis. In: Proceedings of the Conference Ontology for the Intelligence Community, George Mason University, Fairfax, Virginia, October 20-22 (2009)
10. Tecuci, G.: Building intelligent agents: An apprenticeship multistrategy learning theory, methodology, tool and case studies. Academic Press, London (1998)
11. Tecuci, G., Boicu, M., Boicu, C., Marcu, D., Stanescu, B., Barbulescu, M.: The Disciple-RKF learning and reasoning agent. Comput. Intell. 21(4), 462–479 (2005)
12. Tecuci, G., Boicu, M., Cox, M.T.: AI Mag, Special issue on mixed-initiative assistants. 28(2) (2007)
13. Tecuci, G., Boicu, M., Cox, M.T.: Seven aspects of mixed-initiative reasoning: An introduction to the special issue on mixed-initiative assistants. AI Mag. 28(2), 11–18 (2007)
14. Tecuci, G., Boicu, M., Marcu, D., Boicu, C., Barbulescu, M.: Disciple-LTA: Learning, tutoring and analytic assistance. J. Intell. Community Res. Dev. (2008), http://lac.gmu.edu/publications/2008/Disciple-LTA08.pdf (Accessed 18 April 2010)
15. Tecuci, G., Schum, D.A., Boicu, M., Marcu, D.: Coping with the complexity of intelligence analysis: cognitive assistants for evidence-based reasoning. Research Report 7 (2010); Learning Agents Center, August 2009, updated April 2010, http://lac.gmu.edu/publications/2009/CompIA.pdf (Accessed 20 April 2010)
16. Toulmin, S.E.: The uses of argument. Cambridge University Press, Cambridge (1963)
17. Wigmore, J.H.: The science of judicial proof. Little, Brown & Co., Boston (1937)
18. Weiss, C.: Communicating uncertainty in intelligence and other professions. Int. J. Intell. Count. Intell. 21(1), 57–85 (2008)

Intelligent Software for Ecological Building Design

Jens Pohl[1], Hisham Assal[1], and Kym Jason Pohl[2]

[1] Collaborative Agent Design Research Center (CADRC)
California Polytechnic State University (Cal Poly)
[2] CDM Technologies, Inc.
San Luis Obispo, California, USA

Abstract. Building design is a complex process because of the number of elements and issues involved and the number of relationships that exist among them. Adding sustainability issues to the list increases the complexity of design by an order of magnitude. There is a need for computer assistance to manage the increased complexity of design and to provide intelligent collaboration in formulating acceptable design solutions. Software development technology today offers opportunities to design and build an intelligent software system environment that can serve as a reliable intelligent partner to the human designer. In this paper the authors discuss the requirements for an intelligent software design environment, explain the major challenges in designing this environment, propose an architecture for an intelligent design support system for sustainable design and present the existing technologies that can be used to implement that architecture.

Keywords: agents, architectural design, collaboration, design, ecological design, ontology, representation, service-oriented architecture (SOA), sustainability.

1 Introduction

Typically, design requires decisions to be made among several imperfect alternatives. It is in the nature of those decisions that designers will often find the need to supplement logical reasoning with intuitive feelings about the problem situation that can lead to creative solutions and new knowledge. As a rule such new knowledge cannot be logically deduced from the existing available knowledge and is validated only after the solution has been discovered and tested. In this respect design is not unlike the decision making activities that occur in a wide range of complex problem situations that have to be dealt with in many professional fields such as management, economics, medicine, law, transportation planning, and military command and control.

2 The Inherent Complexity of Building Design

Design is the core activity in the field of architecture. The design of even a relatively simple low-rise building can be a complex task involving critical issues related to macro and micro climatic conditions, building loads and structural system selection, site planning, internal space layout, heating and cooling, ventilation, lighting, noise

G. Phillips-Wren et al. (Eds.): Advances in Intel. Decision Technologies, SIST 4, pp. 11–28.
springerlink.com

control and room acoustics, construction materials and finishes, security, privacy, construction duration and cost, labor and product availability, and aesthetics. Since many of these design issues tend to conflict in different ways, it is not just the number of issues involved but in particular the relationships among the issues that are the core cause of design complexity.

To come to terms with such a complex problem solving environment architects pursue an iterative path of analysis, synthesis, and evaluation that requires the design problem to be decomposed into multiple sub-problems (Pohl 2008). Typically, they will select what they consider to be the most important issues and analyze those largely in isolation from the other issues. The results of this analysis are then synthesized into narrow solutions, which are evaluated in the context of both the selected and the remaining issues. When the narrow solutions fail to adequately cater for some of the issues the entire analysis, synthesis, and evaluation cycle is laboriously repeated with the objective of generating better narrow solutions.

3 Increased Complexity of Ecological Design

Based on current and historical building construction and occupancy experience it is quite difficult to imagine the design and operation of a building that is not in some measure destructive to the natural environment. Typically: the site is graded to provide convenient vehicular access and suit the layout of the building and its immediate surroundings; the construction materials and components are produced from raw materials that are extracted from nature and consume a great deal of energy during their production; the materials and components are transported to the site consuming more energy in transit; on-site construction generates waste in terms of packaging material and the fabrication of footings, walls, floors, and roof; during the life span of the building energy is continuously consumed to maintain the internal spaces at a comfortable level and power multiple appliances (e.g., lights, communication and entertainment devices, food preservation and preparation facilities, and security systems); despite some concerted recycling efforts much of the liquid and solid waste that is produced during the occupancy of the building is normally collected and either treated before discharge into nature or directly buried in landfills; and finally, at the end of the life span when the building is demolished most, if not all, of the construction materials and finishes are again buried in landfill sites.

Let us consider the other extreme, a building that has been designed on ecological principles and is operated as a largely self-sufficient micro-environment. Ecological design has been defined in broad terms as being in symbiotic harmony with nature (Van Der Ryn and Cowan 1996, Kibert 2005). This means that the building should integrate with nature in a manner that is compatible with the characteristics of natural ecosystems. In particular, it should be harmless to nature in its construction, utilization, and eventual demolition. The implementation of ecological design concepts in architecture has gained momentum over the past two decades with the increasing adoption of sustainability as a primary design criterion.

In the context of the built environment sustainability is the overarching concept that acknowledges the need to protect the natural environment for future generations. It proposes that anything that we build today should be sustainable throughout its life span. Furthermore, at the end of its life span it should be amenable to deconstruction and the reuse of all of its materials in some form. For a building to meet the full intentions of sustainability it would need to:

- be constructed only of materials and products that are reusable in some form or another at the time of deconstruction of the building and, by implication, most of these materials would already contain recycled ingredients;
- be constructed of materials and products that used as little energy (i.e., embodied energy) as possible during their manufacture;
- be constructed of materials that are not subject to toxic off-gassing;
- be as close to energy self-sufficiency as possible subject to climatic and technology limitations;
- employ water harvesting, treatment and reuse strategies to reduce its freshwater draw to the smallest possible amount (i.e., about 10% of existing usage based on current predictions); and,
- incorporate a waste management system that is capable of recycling most, if not all, of the dry and wet waste produced in the building.

The overarching impact of such stringent sustainability-based design and occupancy requirements adds an order of magnitude of complexity to an already very complex and time consuming building design process. How will the architect be able to cope with the increasing complexity of the building design process under these exacting ecological design principles based on sustainability criteria? Clearly, this is not just a matter of academic preparation and experience, but will depend on the ability of the designer to apply sufficient technical depth and breadth to the development of the design solution. Such an ability will increasingly depend on the availability of an arsenal of readily accessible and seamlessly integrated design tools. What is required amounts to an intelligent design environment that seamlessly assists the designer in finding and gaining access to the required information, generating and evaluating narrow solutions on the basis of simulations, identifying and resolving conflicts as narrow solutions are merged into broader solutions, and continuously monitoring the progress of the overall design solution within a dynamically interactive and collaborative software environment.

4 Desirable Capabilities of an Intelligent Design Environment

Some importance is attached to the term environment in preference to the more conventional nomenclature that would refer to a related set of software components that are intended to interoperate as a system. The use of the term environment is intended to convey a level of integration of capabilities that is seamless and transparent to the user. In other words, while engaged in the design activity the designer should not be conscious of the underlying software and inter-process communication infrastructure that is necessary to support the operation of the environment. The objective is for the designer to be immersed in the design activity to the extent that both the automated capabilities operating mostly in background and the capabilities explicitly requested by the user at any particular time operating in foreground are an integral part of the process.

From a general point of view there are at least two overriding requirements for an intelligent computer-based design environment. The first requirement relates to the representation of information within the environment. The software must have some level of understanding of the information context that underlies the interactions of the human user with the environment. This is fundamental to any meaningful human-computer interaction that is akin to a partnership. The level to which this understanding can be elevated will largely determine the assistance capabilities and essentially the value of the software environment to the human designer.

The second requirement is related to the need for the designer to be able to collaborate. In a broad sense this includes not only the ability to interact with human users who play a role in the design process, such as members of the design team, specialist consultants, material and product vendors, contractors and subcontractors, the building owners and their representatives, and local building authorities, but also non-human sources of information and capabilities. All of these interactions between the designer, other human participants in the design process, data sources, and software-based problem solving capabilities, must be able to be performed seamlessly without the user having to be concerned about access protocols, data formats, or system interoperability issues.

While these overall requirements would at first sight appear to be utopian compared with the state of computer-based environments that exist today (2010), the technology needed for the creation of such environments has been rapidly emerging during the past decade and is now largely available. However, before addressing the technical software design aspects it will be necessary to delve more deeply into the functional requirements of the postulated intelligent design environment.

Emphasis on partnership: A desirable computer-aided design environment is one that assists and extends the capabilities of the human designer rather than replaces the human element. Human beings and computers are complementary in many respects. The strengths of human decision makers in the areas of conceptualization, intuition, and creativity are the weaknesses of the computer. Conversely, the strengths of the computer in computation speed, parallelism, accuracy, and the persistent storage of almost unlimited detailed information are human weaknesses. It therefore makes a great deal of sense to view a computer-based design environment as a partnership between human and computer-based resources and capabilities.

This is not intended to suggest that the ability to automate functional sequences in the computer-based environment should be strictly confined to operations that are performed in response to user actions and requests. Apart from the monitoring of problem solving activities, the detection of conflicts, and the execution of evaluation, search and planning sequences, the computer-based environment should be able to undertake proactive tasks. The latter should include not only anticipation of the likely near-term need for data from sources that may be external to the design environment and need to be acquired by the environment, but also the exploration of alternative solution strategies that the environment considers promising even though the user may be currently pursuing another path.

It follows that the capabilities of the computer-based environment should be designed with the objective of assisting and complementing the user in a teaming role. Such tools are interactive by nature, capable of engaging in collaboration with the user to acquire additional information to help better understand the situation being analyzed. These tools are also able to provide insight into the reasoning processes that they are applying, thereby allowing the designer to gain confidence in their inferencing capabilities as well as make subtle adjustments in the logic being applied. The authors' past experience with multi-agent decision-support applications has shown that tools that are engineered for collaboration with each other and the human user provide opportunities for augmenting their capabilities through user interaction during execution (Pohl et al. 1997). It is therefore suggested that these kinds of tools better assist designers in dealing with the complexity of design. In other words, a collaborative approach affords the necessary visibility and agility to deal with the large number of considerations across a far reaching set of domains that characterizes the design activity.

Collaborative and distributed: Design or complex problem environments in general normally involve many parties that collaborate from widely distributed geographical locations and utilize information resources that are equally dispersed. A computer-based design environment can take advantage of the distributed participation by itself assuming a distributed architecture. Such an architecture typically consists of several components that can execute on more than one computer. Both the information flow among these components and the computing power required to support the system as a whole can be decentralized. This greatly reduces the potential for communication bottlenecks and increases the computation speed through parallelism.

An open architecture: The high degree of uncertainty that pervades complex problem environments, such as design, extends beyond the decision-making activity of the collaborating problem solvers to the configuration of the computer-based environment itself. The components of a design environment are likely to change over time, through modification, replacement, deletion, and extension. It should be possible to implement these changes in a seamless fashion through common application programming interfaces and shared resources.

Tools rather than solutions: The computer-based design environment should offer a set of tools rather than solutions to a predetermined set of problems. The indeterminate nature of design problems does not allow us to predict, with any degree of certainty, either the specific circumstances of a future problem situation or the precise terms of the solution. Under these circumstances it is far more constructive to provide tools that will extend the capabilities of the human designer in a highly interactive problem solving environment.

Expressive internal representation: The ability of the computer-based environment to convey a sense of having some level of understanding of the meaning of the data and in particular the concepts being processed is the single most important prerequisite for a collaborative design environment (Assal et al. 2009). An expressive representation of the real world entities and concepts that define the problem space forms the basis of the interactions between the users and the design environment and, also, the degree of intelligence that can be embedded within its components. To the designer a building consists of real world entities such as rooms, walls, windows, and doors, as well as related concepts such as accessibility, energy conservation, and structural efficiency. Each of these notions has properties and relationships that determine their behavior under certain conditions. These semantic descriptors form the basis of collaboration among human problem solvers and are therefore likewise the fundamental subject matter being discussed in a computer-based design environment.

Embedded knowledge: The computer-based design environment should be a knowledge-based environment. In this context knowledge can be described as experience derived from observation and interpretation of past events or phenomena, and the application of methods to past situations. Knowledge-bases capture this experience in the form of rules, case studies, standard practices, and descriptions of objects and object systems that can serve as prototypes. Problem solvers typically manipulate these prototypes or patterns through adaptation, refinement, mutation, analogy, and combination, as they apply them to the solution of current problems (Gero et al. 1988, Pohl 2008).

Decentralized decision-making: The computer-based design environment need not, and should not, exercise centralized control over the problem solving process. Much of the design activity will be localized and performed in parallel involving the collaboration of different members of the design team. In this regard building design is neither a rigidly controlled nor a strongly disciplined activity, but more aptly described as a process of information seeking, discovery, and subsequent processing. For example, intelligent and dynamically interactive design tools that are responsible for pursuing the interests of real world objects, such as spaces and other building elements (Pohl 1997) and management personnel in commercial and industrial applications (Pan and Tenenbaum 1991), can achieve many of their objectives through employing services and engaging in negotiations that involve only a few nodes of the design environment. This greatly reduces the propensity for the formation of communication bottlenecks and at the same time increases the amount of parallel activity in the computer-based environment.

Emphasis on conflict identification: The capabilities of the computer-based design environment should not be bound by the ultimate goal of the automated resolution of conflicts, but rather the identification of the conflict and presentation to the human designer as much of the related context as possible. This notion gains in importance as the level of complexity of the design problem increases. The resolution of even mundane conflicts can provide subtle opportunities for advancing toward design solution objectives. These opportunities are more likely to be recognized by a human designer than a computer-based agent. The identification of conflicts is by no means a trivial undertaking. It includes not only the ability to recognize that a conflict actually exists, but also the determination of the kind of conflict and the relationships and related context that describe the conflict and what considerations appear relevant to its resolution.

Adaptability and agility: Traditionally, software tools categorized as intelligent were engineered for specific scenarios. Consequently, the successful application of these tools depended largely on the degree to which the characteristics of a particular problem component aligned with situations that the tool had been design for. This rigidity has tended to prove quite problematic when these tools were applied to even slight variations of the scenarios that they had been developed or trained for.

In contrast, what the experience of the authors has shown is that intelligent tools not only need to support variation, but that these tools should be engineered with such adaptation as a core criterion. Much of this ability to effectively deal with variation is due to the ability of these tools to decompose complex problems into much more manageable components without losing the relationships that tie the components together. To accomplish this, the reasoning capabilities of the tools can be organized as discrete fragments of logic capable of addressing smaller components of the larger problem. If these components are described within an expressive, relationship-rich representation then the connections between the decomposed components are maintained automatically. The effects of addressing each individual component are automatically propagated across the entire expanse of the problem space due to the extensive set of relationships represented within the model that retains their connections and context.

The human-computer interface: The importance of a high degree of interaction between the human members of the design team and the various intelligent components of the computer-based design environment is integral to most of the principles and requirements described above. This interaction is fundamentally facilitated by the information-centric representation core of the environment through which the interacting software components are able to maintain some level of understanding of the current context of the design activity. However, there are other aspects of the user-interface that must be provided in support of the human-computer interactions. These include two-dimensional and three-dimensional graphical representation capabilities, explanation facilities, and a context-sensitive help system with semantic search support.

At a minimum the graphical capabilities must be powerful enough to include the accurate representation of the analysis results of the progressively evolving design solution in terms of the environmental factors that are involved in building design, such as: shadows based on sun path projections; daylighting and artificial lighting simulations within the building interior to the extent that adverse conditions such as glare can be readily perceived by the human designer; structural behavior based on the simulation of static dead and live loads, as well as dynamic wind and seismic loads; animated air movement and heat flow simulations; and, pedestrian traffic visualization. Technology permitting, the ultimate aim of the design environment should be to provide a virtual reality user-interface that allows the human designer to become fully immersed in the physical and emotional aspects of the design experience.

The authors' experience with decision-support systems over the past two decades has lent credence to the supposition that the need for the proposed design environment to be able to explain how it arrived at certain conclusions increases with the sophistication of the inferencing capabilities embedded in the software environment. At the very least, the intelligent components of the environment should be able to explain their results and methods of analysis. In this regard retrospective reasoning that is capable of providing answers to *what*, *how*, and *why* questions is the most common type of explanation facility available in multi-agent systems. A *what* question requires the explanation or definition of a fact. For example, in the context of architectural design the user may ask: *What are the characteristics of the window in the north wall of the conference room?* In the past, expert system methodologies based on *format templates* would have allowed the appropriate answer to be collected simply through template values when a match is made with the facts (i.e., window, north, wall, conference) contained in the question (Myers et al. 1993). Today, with the application of ontology-based reasoning capabilities more powerful and direct methods based on the ability of an ontology to represent concepts are available. A *how* question requires an analysis of the sequence of inferences or reasoning that produced the fact. Continuing with the above example, the designer may ask: *How can the intrusion of external noise into the conference room be mitigated?* The answer would require a sequence of inferences by the Noise Agent. This sequence can be preserved and presented to the designer.

Why questions are more complicated. They require reference to the sequence of goals that have driven the progression of inferences (Ellis 1989). In large collaborative systems many agents may have contributed to the inference sequence and will need to participate in the formulation of the answer. This third level of explanation, which requires a summary of justification components, has received considerable attention over the past 30 years. For example: text summary systems such as Frump (Dejong 1982) and Scisor (Jacobs and Rau 1988); fast categorization techniques such as Construe (Hayes and Weinstein 1991); grammatical inference that allows inductive operators to be applied over the sequences of statements produced from successive justifications (Michalski 1983); explanation-based learning (Mitchell et al. 1991); and, case-based reasoning (Shank et al. 1990).

While existing computer-aided design systems typically support only factual searches, the proposed intelligent design environment should provide semantic search capabilities that can deal with inexact queries. Due to the complexity of the problem space the designers will not always know exactly what information they require. Often they can define only in conceptual terms the kind of information that they are seeking. Also, they would like their query to be automatically broadened with a view to discovering additional information that may be relevant to their current problem solving focus.

The desirability of the design environment to be able to deal with inexact search requests warrants further discussion. A flexible query capability, such as the human brain, can generate best guesses and a degree of confidence for how well the available information matches the query. For example, let us assume that the designer is searching for a window unit of something like the *double-hung* window type. The flexible query facility would presumably include a *something like or similar to* operator capable of matching in a partial sense. Windows that have a movable part are something like the *double-hung* window type. Windows that have their movable part in the vertical direction are more like *double-hung* than windows that have their movable part in the horizontal direction. Windows that open by rotation are even less like *double-hung* than windows that are simply fixed. In other words each of the *something like* information items would be validated by a degree of match qualification.

5 The Technical Approach

The desired capabilities of the proposed intelligent design environment outlined in the previous section call for a distributed system architecture that can be accessed from any physical location, is highly flexile, and totally transparent to the human user. In particular, the user must be shielded from the many protocols and data and content exchange transformations that will be required to access capabilities and maintain

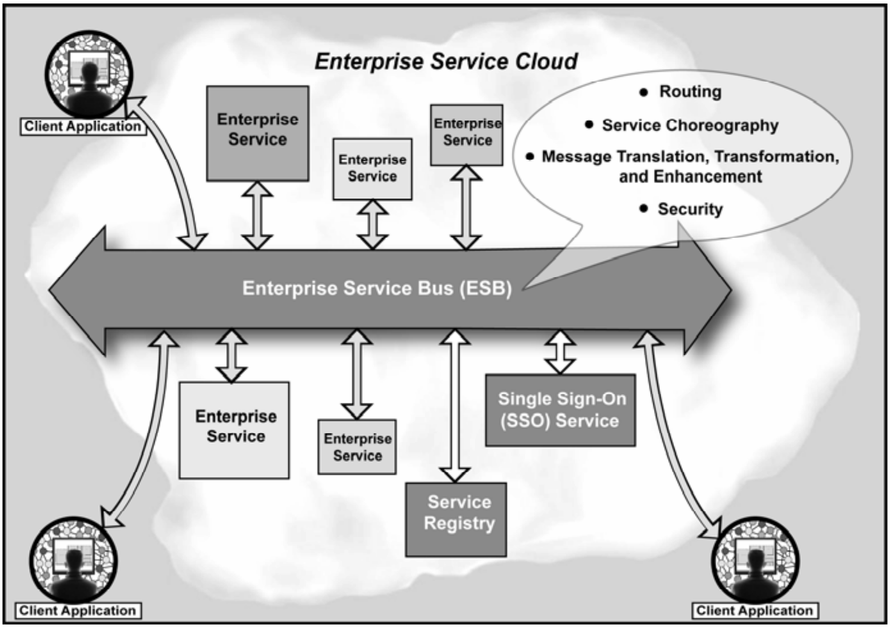

Fig. 1. Principal SOA Components

seamless interoperability among those capabilities. Any member of the design team, once authenticated during the single sign-on point of entry, should be able to access those capabilities (e.g., intelligent design tools and data sources) that are included in the authentication certificate. The focus of the designer should not be on systems, as it still is mostly today, but on the capabilities or services that the computer-based environment can provide.

Service-oriented architecture (SOA): In the software domain these same concepts have gradually led to the adoption of Service-Oriented Architecture (SOA) principles. While SOA is by no means a new concept in the software industry it was not until Web services became available that these concepts could be readily implemented (Erl 2008, Brown 2008). In the broadest sense SOA is a software framework for computational resources to provide services to customers, such as other services or users. A fundamental intent that is embodied in the SOA paradigm is flexibility. To be as flexible as possible a SOA environment is highly modular, platform independent, compliant with standards, and incorporates mechanisms for identifying, categorizing, provisioning, delivering, and monitoring services. The principal components of a conceptual SOA implementation scheme (Figure 1) include a Enterprise Service Bus (ESB), one or more portals to external clients with single sign-on facilities, and the enterprise services that facilitate the ability of the user community to perform its operational tasks.

The concept of an Enterprise Service Bus (ESB) greatly facilitates a SOA implementation by providing specifications for the coherent management of services. The ESB provides the communication bridge that facilitates the exchange of messages

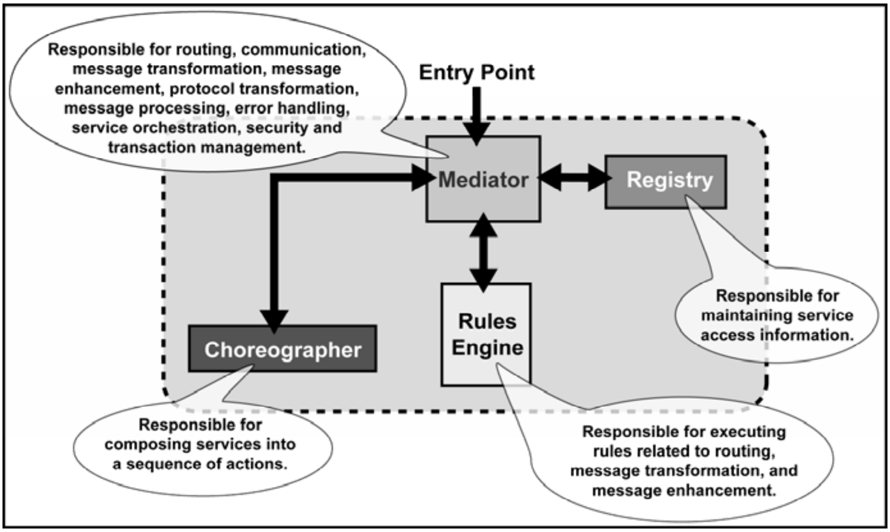

Fig. 2. Principal ESB Components

among services, although the services do not necessarily know anything about each other. According to Erl (2008) ESB specifications typically define the following kinds of message management capabilities: routing; protocol transformation; message transformation; message enhancement; service mapping; message processing; process choreography and orchestration; transaction management; and, access control and security.

There are quite a number of commercial off-the-shelf ESB implementations that satisfy these specifications to varying degrees. A full ESB implementation would include four distinct components (Figure 2): Mediator; Service Registry; Choreographer; and, Rules Engine. The Mediator serves as the entry point for all messages and has by far the largest number of message management responsibilities. It is responsible for routing, communication, message transformation, message enhancement, protocol transformation, message processing, error handling, service orchestration, transaction management, and access control (security).

The Service Registry provides the service mapping information (i.e., the location and binding of each service) to the Mediator. The Choreographer is responsible for the coordination of complex business processes that require the participation of multiple service providers. In some ESB implementations the Choreographer may also serve as an entry point to the ESB. In that case it assumes the additional responsibilities of message processing, transaction management, and access control (security). The Rules Engine provides the logic that is required for the routing, transformation and enhancement of messages.

Information-centric representation: The methods and procedures that designers utilize to solve design problems rely heavily on their ability to identify, understand and manipulate objects. In this respect, objects are complex symbols that convey meaning by virtue of the explicit and implicit context information that they encapsulate within their domain. For example, architects develop design solutions by reasoning about neighborhoods, sites, buildings, floors, spaces, walls, windows, doors, and so on. Each of these objects encapsulates knowledge about its own nature, its relationships with other objects, its behavior within a given environment, what it requires to meet its own performance objectives, and how it might be manipulated by the designer within a given design problem scenario. This knowledge is contained in the various representational forms of the object as factual data, algorithms, rules, exemplar solutions, and prototypes (Pohl 2008, 59-62).

It is therefore apparent that a critical requirement for effective human-computer interaction in the proposed intelligent design environment is the appropriate representation of the evolving design solution model. This can be accomplished utilizing an ontology. The term ontology is loosely used to describe an information structure that is rich in relationships and provides a virtual representation of some real world environment. The elements of an ontology include objects and their characteristics, different kinds of relationships among objects, and the concept of inheritance (Assal et al. 2009). While an ontology is expressed in object-oriented terms, it is more than an object model. It is designed to describe the entities, concepts, and related semantics of some subject matter domain. Software that incorporates an internal information model, such as an ontology, is often referred to as information-centric software. The information model is a virtual representation of the real world domain under consideration and is designed to provide adequate context for software agents (typically rule-based) to reason about the current state of the virtual environment.

Within a SOA-based system environment the various information-centric tools that are available to the designer will exist as an integrated collection of clients (i.e., users of the ontology), typically referred to as services. These services can communicate directly or indirectly via message translation, in terms of the real world objects and relationships that represent the contextual framework of the evolving design solution. To reduce the amount of work (i.e., computation) that the computer has to accomplish and to minimize the volume of information that has to be transmitted within the system, two strategies can be readily implemented. First, since the services involved in a particular collaboration are stateful in nature (i.e., they retain a working knowledge of the various aspects of the evolving design solution that they are concerned with) only the changes in information need to be communicated. For example, an agent that is monitoring the layout of spaces during the design of a building may have an extensive set of information concerns or interests relating to various aspects of the evolving design solution. These interests will likely include the location, geometric parameters and functional characteristics of a particular space. If the designer changes the locations of this space then only that aspect should be transmitted to interested parties.

Second, to further economize on communication traffic as well as increase the timeliness and efficiency with which components (i.e., agents, etc.) interact, an asynchronous notification facility (i.e., subscription service) can be employed where parties can indicate their respective information interests. When entries with such subscription profiles are satisfied, respective users are asynchronously notified allowing them to take whatever action they see fit. By allowing relevant information to be automatically pushed to interested parties, the subscription service obviates the need for repetitive queries and thereby greatly reduces the amount of work the computer has to perform.

5.1 Design Tools as Agents Types

On the assumption of an information-centric software architecture that incorporates an ontology-based high level representation of the design problem context, the intelligence of the proposed design environment will be largely contributed by the design tools that are available to the human designer. Most of these design tools will be in the form of invocable services or self-initiating agents. There is a behavioral distinction between services and agents. Services are invoked to perform a discrete activity, returning to their original inactive state after the activity has been completed. Agents on the other hand may be active on a continuous basis, taking the initiative opportunistically whenever they determine that the situation warrants an action. Often these agent actions will invoke services.

There are many types of software agents, ranging from those that emulate symbolic reasoning by processing rules, to highly mathematical pattern matching neural networks, genetic algorithms, and particle swarm optimization techniques. While all of these have capabilities that are applicable to an intelligent design environment, only symbolic reasoning agents that can interact directly with the ontology-based design context model will be discussed in this paper. For these rule-based agents the reasoning process relies heavily on the rich representation of entities and concepts provided by the ontology.

Agent tools can be characterized as being autonomous because they can act without the direct intervention of human operators, even though they allow the latter to interact with them at any time. One important aspect of autonomy in agent applications is the ability of agents to perform tasks proactively whenever such actions may be appropriate. This requires agents to be opportunistic, or continuously looking for an opportunity to execute. In this context opportunity is typically defined by the existence of sufficient information. For example, as the location of a particular space is defined by the designer within the evolving floor plan, several agents may become involved automatically to undertake analyses (e.g., thermal, lighting, acoustics) appropriate to their capability domains.

Planning Agents: Planning Agents are complex agents that reason about the problem state and produce a plan based on the current state of the design in conjunction with the applicable constraints and objectives. This planning process involves matching the latter with the available resources to produce a course of action that will satisfy the desired objectives. The complexity of the process can be reduced by distributing the basic planning tasks among a set of agents, as follows: identify the constraints and objectives; identify the available resources; note the unavailability of resources; identify the available set of actions or characteristics; and, generate a plan for satisfying the objectives.

Plan or solution generation is the actual planning activity in the above list of tasks. Many planning systems use specialized search algorithms to generate plans according to given criteria (Blum and Furst 1997). Re-planning, which is also commonly referred to as continual planning, involves the re-evaluation of parts of an existing plan or solution because of a change in the information that has been used in the creation of that plan. This is a common situation in architectural design, where the designer is continuously adapting the evolving design solution during the iterative analysis-synthesis-evaluation cycle (Pohl 2008, 47-52).

Service Agents: Agents that are designed to be knowledgeable in a specific domain, and perform planning or assessment tasks in partnership with other agents (i.e., human agents or software agents) are often referred to as Service Agents (Durfee 1988, Pohl et al. 1997). The manner in which they participate in the decision-making activities depends on the nature of the situation. Service Agents can be designed to react to changes in the problem state spontaneously through their ability to monitor information changes and respond opportunistically. They should be able to generate queries dynamically and access resources automatically whenever the need arises.

In the proposed intelligent design environment both Service and Planning Agents will constitute the principal design tools by providing analysis, solution generation and evaluation capabilities for the full range of knowledge domains that impact an ecologically based design solution, namely: site analysis; building orientation; space layout optimization; structural system selection; deconstructability assessment; thermal design determinates; passive solar system analysis; mechanical heating, ventilating and air-conditioning solution generation and evaluation; daylighting and artificial lighting design; alternative energy analysis and solar system alternatives; room acoustics and noise insulation; building hydrology analyses; closed-loop material selection; embodied energy analysis; waste disposal and recycling; life cycle cost analysis; construction cost estimation; and so on.

What is of particular significance is that unlike the manual design process, which requires these related design factors to be considered in an essentially sequential manner, the various agents will be able to operate in parallel in the proposed design environment. Furthermore, the ability of the agents to collaborate will allow the relationships among the different knowledge domains to be pursued dynamically. Since the complexity of the building design activity is due to the large number of relationships among the domains, the proposed design environment embodies the potential for dealing with a highly complex problem situation in a holistic manner.

Mentor Agents: A Mentor Agent is a type of agent that is based on the agentification of the information entities and concepts that are intrinsic to the nature of each application. In the proposed design environment these are the primary building elements and concepts that the architect reasons about and that constitute the foundations of the internal representation (i.e., ontology) of the problem situation within an information-centric software system (Pohl 1996). For example, a Mentor Agent may attend to the needs of a specific building space (i.e., an entity) or pursue energy conservation objectives (i.e., a concept) that govern the entire design solution. The concept of Mentor Agents brings several potential benefits.

First, it increases the granularity of the active participants in the problem solving process. As agents with collaboration capabilities, agentified design elements can pursue their own objectives and perform a significant amount of local problem solving without repeatedly impacting the communication and coordination facilities utilized by the higher level components of the distributed system. Typically, a Mentor Agent is equipped with communication capabilities, process management capabilities, information about its own nature, global objectives, and some focused problem solving tools.

Second, the ability of Mentor Agents to employ services greatly increases the potential for concurrent activities. Multiple Mentor Agents can request the same or different services simultaneously. If necessary, Service Agents responding to multiple service requests can temporarily clone themselves so that the requests can be processed in parallel. Third, groups of Mentor Agents can negotiate among themselves in the case of matters that do not directly affect other higher level components or as a means of developing alternatives for consideration by higher level components.

Fourth, by virtue of their communication facilities Mentor Agents are able to maintain their relationships to other aspects of the evolving design solution. In this respect they are the product of *decentralization* rather than *decomposition*. In other words, the concept of Mentor Agents overcomes one of the most serious deficiencies of the rationalistic approach to problem solving; namely, the dilution and loss of relationships that occurs when a complex problem is decomposed into sub-problems. In fact, the relationships are greatly strengthened because they become active communication channels that can be dynamically created and terminated in response to the changing state of the problem situation.

In the realm of building design it would seem desirable to implement building spaces as agents. Since Mentor Agents have communication capabilities a conference room Space Agent, for example, would be able to collaborate with other agents such as Service and Planning Agents. If the conference room Space Agent is interested in determining where it is located in respect to any surrounding sources of noise it could invoke the services of a Noise Agent to identify any relevant noise sources. This example illustrates two distinct benefits: only the most necessary computation has been performed; and, the information that forms part of the fundamental description of the results can be held anywhere in the system (as long as it is available to any other authorized agent). Second, by distributing the collaborating parties, as well as the information that is generated as a result of the servicing of the requests, the communications involved with both the current interactions and any future use of the relevant information have been likewise distributed. Accordingly, the potential for the occurrence of a communication bottleneck has been effectively reduced.

5.2 Agent Collaboration and Conflict Management

In previous multi-agent design and military decision-support systems developed by the authors (ICADS 1991, AEDOT 1992, Diaz et al. 2006) conflicts arose when agents either disagreed among themselves or with a decision made by the designer. For example, the placement of a window in a particular space might provoke the latter type of conflict. If the designer places the window in the west wall of a conference room and a loud noise source such as a freeway runs parallel to the west boundary of the site, then the Noise Agent (a Service Agent) would insist on the removal of the window. The designer is able to resolve the conflict by relocating or deleting the window or, alternatively, overruling the Service Agent altogether. The conference room, as a passive entity, is involved in the conflict resolution process only as an information source that is used by the Service Agent in its deliberations. In other words, while the validation of the design decision is entirely dependent on the knowledge encapsulated in the informational entity the latter is unable to actively participate in the determination of its own destiny.

The situation is somewhat analogous to a scenario common in real life when one or more persons feel compelled to make decisions for another person, although the latter might be more competent to make those decisions. The outcome is often unsatisfactory because the decision makers tend to use general domain information where they lack specific knowledge of the other person. In other words, the individuality of the problem situation has been usurped by the application of generalizations and, as a result, the quality of the decisions that have been reached are likely to be compromised.

In the example of the window in the west wall of the conference room, if the conference room is a Space Agent then much of the decision-making can be localized within the knowledge domain of the agent. As soon as the window has been placed in the wall by the designer the conference room Space Agent could pose two specific questions to the appropriate Service Agents (i.e., in this example the Noise Agent and the Lighting Agent): What is the expected background noise level in the room due to the window? and What is the spatial distribution of daylight admitted through the window? The answers to these questions can be compared by the conference room Space Agent directly to what it knows about its own acoustic and lighting needs. The development of alternative strategies for resolving the noise problem can now take place within the context of all of the information in the conference room Space Agent's knowledge domain. For example, the possibility of' relocating itself to a quieter wing of the building can be explored by the agent (with or without the active collaboration of the designer) as a direct consequence of its own deliberations.

There is another kind of conflict resolution scenario that becomes possible with the availability of Mentor Agents. An agent may develop a solution to a sub-problem in its own domain that redirects the entire design solution. In the conference room example the Space Agent may resolve the noise control problem by adopting an expensive window unit (e.g., triple glazing) solution, and then continue to search for a more effective solution as the design solution continues to evolve. The search may continue into subsequent stages of the design process, during which the conference room might progressively be governed by a Mentor Agent representing the entire floor or even the building as a whole. These higher level agents may now impose certain conditions on the Space Agent for the greater good of the larger community. However, the Space Agent, persevering in its search finally comes up with a method of noise control that utilizes a novel type of wall construction in combination with background masking sound. The proposed wall construction may even be contrary, yet still compatible, to that adopted for the external west wall of the building by both the Floor and Building Agents.

First, it is significant that this alternative solution has been found at all. If the conference room had been a passive data object there would not have been any desire on the part of the system to pursue the problem after the initial conflict resolution. Second, having found the alternative the conference room Space Agent is able to communicate its proposal and have the noise control issue reconsidered. It could engage in a discourse with, in order of authority, the Floor Agent and the Building Agent. At each of the agent levels there is the opportunity for wider consultation and interaction with the designer. Finally, if the proposal has been rejected at all higher agent levels, the conference room Space Agent may appeal directly to the designer. The designer has several alternative courses of actions available: also reject the proposal; require one or more of the higher level agents to explain their ruling; reset certain parameters that allow the higher level agents to reconsider their ruling; overrule the higher level agents and accept the proposal; or, capture the current state of the design solution as a recoverable snapshot and use the Space Agent's proposal as the basis for the exploration of an alternative solution path.

5.3 System Architecture

The proposed system consists of a number of components in a SOA-based environment, and a client application that serves as the user-interface for all of the user-service communications. The system components include:

A client application with support for computer-aided drawing (CAD) capabilities and a Building Information Model (BIM) interface. The use of BIM captures the design information in a standard way, which can be communicated to other system components. A BIM model representation is typically in Extensible Markup Language (XML) format, which supports the hierarchical structure of design elements. The client application is the only user-interface in the system. It provides the user with tools to access the other services and presents the information generated by services (i.e., service results) within the CAD application. The client application also includes a Business Process Management (BPM) component to allow the user to describe a collaborative workflow, which may involve other human users (e.g., external structural consultant) and system services. The BPM component takes a user description of a process and hands it to the SOA-based environment, namely the Enterprise Service Bus (ESB), for execution. The client application also displays any information received from the services as the result of analysis, recommendations, or warnings.

- A CAD service, which is responsible for communicating between the CAD environment and the ontology environment.
- A translation service that translates the BIM model into the system ontological representation to allow the higher level inferences to take place. This service is made part of the workflow through the user settings in the client application. (Taylor et. al. 2009, Pohl 2008).
- An ontology service that builds, maintains, and handles the communication of the ontology with the other services. The ontology service contains the subscription service described previously, which registers the interests of other components in ontology changes. The ontology service also builds additional relationships into the model, which was exported from BIM. The additional relationships are inferred based on the existing ones and they provide enhanced context for the inference services.
- An inference service that is made up of a number of agent communities. An agent community is a collection of agents in a given domain (e.g., energy efficiency, water use, recycling, etc.). Each agent examines the design from its perspective and produces an assessment of the quality of the design elements in that perspective. Agents may make recommendations or enhancements to the design elements and communicate the recommendations back to the ontology. The inference service is connected to the ontology service and monitors changes there, through the ontology subscription service.

6 Conclusions

Design for sustainability combines the complexity of traditional architectural design with the complexity of considering a host of environmental issues that are based on ecological principles, in the evolving design solution. Management of this compound complexity requires the assistance of an intelligent software system environment. There are two main requirements for such an environment. One is a rich contextual representation of design information. The second is collaboration between the human user and the software environment. The current state of technology in software development offers opportunities for developing a distributed, collaborative, intelligent design support system. Service-oriented architecture (SOA) concepts provide the framework and the guiding principles for developing distributed, service-based systems. The field of ontological representation offers a direction for the expressive modeling of domain knowledge, which forms an enabling foundation for intelligent agents as autonomous, collaborative software tools that can monitor the evolving design, participate in problem solving in specific domains, gather and present relevant information to the designer, and communicate with the user when necessary.

References

AEDOT, AEDOT Prototype 11: An Implementation of the ICADS Model, Technical Report CADRU-07-92. Collaborative Agent Design Research Center, Cal Poly, San Luis Obispo (1992)

Assal, H., Pohl, K., Pohl, J.: The Representation of Context in Computer Software. In: Pre-Conference Proceedings, Focus Symposium on Knowledge Management, Inter. Symp.-2009, Baden-Baden. Collaborative Agent Design Research Center, Cal Poly, San Luis Obispo (2009)

Blum, A., Furst, M.: Fast Planning Through Planning Graph Analysis. Artificial Intelligence 90, 281–300 (1997)

Brown, P.: Implementing SOA: Total Architecture in Practice. Addison-Wesley, Reading (2008)

Dejong, G.: An Overview of the Frump System. In: Lehnert, W.G., Ringle, M.H. (eds.) Strategies for Natural Language Processing, pp. 149–176. Lawrence Erlbaum, New Jersey (1982)

Diaz, C., Waiters, W., Pickard, J., Naylor, J., Gollery, S., McGraw, P., Huffman, M., Fanshier, J., Parrott, M., ODriscoll-Packer, S., Pendergrast, P., Sylvester, E.: ICODES: Technical and Operational Description, Technical Report CDM-20-06. CDM Technologies Inc., San Luis Obispo (2006)

Durfee, E.: Coordination of Distributed Problem Solvers. Kluwer Academic, Boston (1988)

Ellis, C.: Explanation in Intelligent Systems. In: Ellis, C. (ed.) Expert Knowledge and Explanation: The Knowledge-Language Interface. E Horwood, Chichester (1989)

Erl, T.: SOA: Principles of Service Design. Prentice Hall, New Jersey (2008)

Gero, J., Maher, M., Zhang, W.: Chunking Structural Design Knowledge as Prototypes, Working Paper, The Architectural Computing Unit, Department of Architectural and Design Science, University of Sydney, Sydney (1988)

Hayes, P., Weinstein, S.: Construe-TIS: A System for Content-Based Indexing of a Database of News Stories. In: Rappaport, A., Smith, R. (eds.) Innovative Applications of Artificial Intelligence 2, pp. 47–64. AAAI Press, Menlo Park (1991)

ICADS, ICADS Working Model Version 2 and Future Directions, Technical Report CADRU-05-9. Collaborative Agent Design Research Center, Cal Poly, San Luis Obispo (1991)

Jacobs, P., Rau, L.: A Friendly Merger of Conceptual Analysis and Linguistic Processing in a Text Processing System. In: Proceedings of the Fourth IEEE AI Applications Conference, pp. 351–356. IEEE Computer Society Press, Los Alamitos (1988)

Kibert, C.: Sustainable Construction: Green Building Design and Delivery. Wiley, Hoboken (2005)

Michalski, R.: A Theory and Methodology of Inductive Learning. Artificial Intelligence 20, 111–161 (1983)

Mitchell, T., Allen, J., Chalasani, P., Cheng, J., Etzioni, O., Ringuette, M., Schlimmer, J.: Theo: A Framework for Self-Improving Systems. In: Van Lehn (ed.) Architectures for Intelligence, Twenty-Second Carnegie Mellon Symposium on Cognition, pp. 323–355. Lawrence Erlbaum, Hillsdale (1991)

Myers, L., Pohl, J., Cotton, J., Snyder, J., Pohl, K., Chien, S., Aly, S., Rodriguez, T.: Object Representation and the ICADS-Kernel Design, Technical Report CADRU-08-93. Collaborative Agent Design Research Center, Cal Poly, San Luis Obispo (1993)

Pan, J., Tenenbaum, J.: Toward an Intelligent Agent Framework for Enterprise Integration. In: Proc. Ninth National Conference on Artificial Intelligence, San Diego, vol. 11, pp. 206–212 (1991)

Pohl, J.: Cognitive Elements of Human Decision-Making. In: Jain, L., Wren, G. (eds.) Intelligent Decision Making: An AI-Based Approach. Springer, New York (2008)

Pohl, J., Chapman, A., Pohl, K., Primrose, J., Wozniak, A.: Decision-Support Systems: Notions, Prototypes, and In-Use Applications, Technical Report, CADRU-11-97. Collaborative Agent Design Research Center, Cal Poly, San Luis Obispo (1997)

Pohl, K.: A Translation Engine in Support of Context-Level Interoperability. Special Issue on Ontology Driven Interoperability for Agile Applications Using Information Systems: Requirements and Applications for Agent Mediated Decision Support. Intelligent Decision Technologies 2(1), 73–81 (2008)

Pohl, K.: KOALA: An Object-Agent Design System. In: Pohl, J. (ed.) Proceedings Focus Symposium on Advances in Cooperative Environmental Design Systems, Inter. Symp. 1996, Collaborative Agent Design Research Center, Cal Poly, San Luis Obispo, pp. 81–92 (1996)

Schank, R., Osgood, R.: Content Theory of Memory Indexing, Technical Report 2. The Institute for the Learning Sciences, Northwestern University (1990)

Taylor, D., Assal, H.: Using BPM to Support Systems Interoperability. The International C2 Journal 3(1) (2009)

Van Der Ryn, S., Cowan, S.: Ecological Design. Island Press, Washington (1996)

Issues in Aggregating AHP/ANP Scales

William C. Wedley

Faculty of Business Administration
Simon Fraser University
Burnaby, B.C., Canada, V5A 1S6

Abstract. Additive synthesis of ratio scales requires the scales to be in a common unit of measure. Unlike regular ratio scales, the unit of measure for relative ratio scales is not readily identifiable. That obscurity complicates the problem of achieving commensurability before multiple scales are synthesized. Examples are given of how conventional AHP may fail to aggregate commensurable values. Several techniques are presented that address the issue of commensurability. The analysis is then extended to more complex forms of aggregation such as benefit/cost analysis and the ANP.

1 Introduction

The main feature that distinguishes the Analytic Hierarchy and Network Process (AHP/ANP) from other Multiple Criteria Decision Making (MCDM) methods is the use of ratio scales. For many MCDM problems such as picking the best alternative, an ordinal answer is sufficient -- all we seek is the top-ranked alternative. For other problems we need ratio answers. In resource allocation, for example, relative measures allow us to apportion different percentages of the total budget to each project. Having stronger properties, ratio scales can handle a wider range of problems.

This paper investigates the nature of ratio scales and how they are combined into composite measures. In the next section a distinction is made between regular and relative ratio scales. Section 3 establishes an example for evaluating methods and reviews conventional AHP techniques that seem to ignore the need for commensurability. Section 4 reviews various methods that do address commensurability. Section 5 investigates the particular problem of aggregating benefit and cost priorities that represent positive and negative measures. Section 6 then looks at ANP aggregation and how it captures interdependent relationships. The paper concludes with a discussion of how relative scales can be tricky.

2 The Nature of Regular and Relative Ratio Scales

Scientific measurement and the establishment of a ratio scale involve the estimation or discovery of the ratio between a magnitude of a continuous quantity and a unit magnitude of the same kind of property (Michell, 1997, 1999). With n objects bearing such a property within a limited range of the continuous quantity, a column vector of

G. Phillips-Wren et al. (Eds.): Advances in Intel. Decision Technologies, SIST 4, pp. 29–42.
springerlink.com

the scale can be represented by $V=(v_1/v_u, v_2/v_u, \ldots v_n/v_u)^T$, or simply $V=(v_1, v_2, \ldots v_n)^T$ since $v_u=1$. While the simpler notation is permissible, the existence and definition of the unit is crucial for interpreting and aggregating the numeric values.

Ratio scales are distinguished by three important characteristics:

1. The magnitude for each object has an origin representing no intensity of the property being considered (e.g. absolute zero on the Kelvin scale where particles have zero kinetic energy).
2. Multiplication by a positive constant ($b>0$, $b\neq1$) is equivalent to specifying a new unit of measure. If $b>1$, transformed values will be larger numeric values with a unit object of smaller magnitude (e.g. 2.2 x the kilogram scale = the pound scale). If $b<1$, the new unit will be an object with larger magnitude.
3. A proportional transformation does not change the equality of ratios. For example: $(bv_i/bv_u)/(bv_j/bv_u) = (v'_i/v'_u)/(v'_j/v'_u)$. Although values change and the unit is different, the ratios are invariant.

When a ratio scale has a well-defined unit of measure, we call that scale a "regular ratio scale". A similarity transform, however, can set the unit to an object that is not well defined. For most AHP situations, the problem is to estimate a ratio scale for subjective factors such as importance, preference, or likelihood (Saaty & Sodenkamp, 2008) where no well-defined unit is established beforehand. Generally, AHP normalizes the derived scale for n objects to sum to unity within the 0-1 range. Such scales are called "relative ratio scales" although all ratio scales are relative to something. Because the unit of a "relative ratio scale" is an abstract object that possesses the total intensity of all n objects, they fall into the classification of "non-standard ratio scales" that have an obscure or unspecified unit of measure. Although non-standard and relative ratio scales are tricky and more challenging to use, they do provide useful information in the form of relative performances. What remains unclear is "relative to what?"

Relative ratio scales do not create problems when we are considering just one property. If we add or remove an object, re-normalization to maintain the sum-to-one convention does not upset relative ratios. As will be shown below, problems can arise if the situation is multi-criteria. If we add or delete objects and renormalize, the relative ratios within a criterion are maintained but the unit of measure has changed. When we aggregate across criteria, we are no longer using the same norms.

Saaty (2004) has a somewhat different perspective of relative ratio scales. After observing that no mathematical definition says that a scale must have a unit or an origin with zero value, Saaty asserts that relative ratio scales do not need them. In many respects he is correct. Beforehand, we do not know absolute zero or the unit, nor do we need them to derive a scale. Instead, we pairwise compare the magnitude of objects, using the objects themselves as temporary and non-specific units. The result of each comparison is an invariant value selected from a fundamental scale of absolute values. The derived sum-to-unity priorities are relative to each other (not their absolute value from zero) and no one object is the unit.

Although relative ratio scales do not display zero and a unit value, it is maintained herein, that it is useful to consider priorities as originating from absolute zero and with a unit of measure. A basic premise of additive synthesis of ratio scales is that they be commensurate. As Saaty (2004, 2008) recognizes, scale values for intangibles are transient and can change if objects are added or deleted. In other words, they become relative to a different set. If we do not try to identify the unit in order to assure commensurability upon such changes, then we are liable to make mistakes.

3 Conventional AHP: Distributive and Ideal

To test aggregation techniques, we shall use regional sales data gathered from Toyota (2009), General Motors (2008) and Volkswagen (2008).The objective is to determine the largest company (the alternatives) based upon revenues from different regions (the criteria). Sales in regional currencies are presented in Table 1. We can look within the columns of different currencies to observe that each company is dominant in its home region.

Table 1. Regional revenues of Major Car Manufacturers

Company	Japan & Asia	North America	Europe	Other	Total
Toyota	9,922,328	58,797	18,955	28,415	10,028,495
General Motors	1,290,847	88,365	22,115	28,486	1,429,813
Volkswagen	1,288,660	18,693	81,149	30,727	1,419,229
Unit of measure	**million ¥**	**million $**	**million €**	**million R$**	

Regional currencies: Yen (¥), US dollars ($), Euros (€), Brazilian reals (R$).

To judge which company is the largest worldwide, it would be nonsensical to rely on the total column. Before we sum across columns, we must have the measures in a common unit. That conversion to Japanese Yen is shown in Table 2. Notice that Toyota is the largest company and that the relative ratios are 0.385, 0.290 and 0.325. These true relative ratios will be used to see if different aggregation techniques can give the correct results.

Table 2. Regional revenues in Japanese Yen

Company	Japan & Asia	North America	Europe	Other	Total	Relative Ratios
Toyota	9,922,328	6,097,676	2,889,753	1,619,813	**20,529,570**	0.385
General Motors	1,290,847	9,164,112	3,371,427	1,623,852	**15,450,237**	0.290
Volkswagen	1,288,660	1,938,554	12,371,165	1,751,651	**17,350,030**	0.325

In using relative measurement for solving AHP problems, Saaty (1993) recommends both the distributive and ideal modes. In the distributive mode, the unit sum of relative ratio scales is distributed to the alternatives. The ideal mode uses regular ratio scales with the most preferred alternative (the ideal) is set as the unit of measure. For both methods, deriving priorities for the criteria is problematic. Unlike ANP, conventional AHP assumes that the criteria are independent of the alternatives. Therefore, questions of the following type are used to elicit ratio intensities.

1. Which Region is more important: Europe or North America?
2. By how many times more important?

These questions are generally easy to answer, but difficult to interpret. What the questions are really seeking is the degree each region contributes to the worldwide sales revenue *of the three companies*. If a fourth company is added (say Ford), then the questions would relate to the worldwide sales of four companies. Being non-specific, the questions do not help the user to realize that answers depend upon the alternatives in the choice set.

In Table 3, the alternative priorities for both modes are based upon the true regional values of Table 1. For analytical purposes, we assume the decision maker compares the regions and comes up with the following relative priorities for criteria: Japan & Asia: 0.25 North America: 0.36, Europe: 0.32, Other: 0.07. Using these criteria weights along with the correct priorities for alternatives, we get the composite results shown in Table 3.

Table 3. Distributive and Ideal AHP results with assumed criteria weights

Distributive Mode

Criteria Weights:	0.25	0.36	0.32	0.07		
Company	**Japan & Asia**	**North America**	**Europe**	**Other**	**Relative Ratios**	**True Ratios**
Toyota	0.794	0.355	0.155	0.324	**0.398**	0.385
General Motors	0.103	0.533	0.181	0.325	0.298	0.290
Volkswagen	0.103	0.113	0.664	0.351	**0.303**	0.325

Ideal Mode

Criteria Weights:	0.25	0.36	0.32	0.07			
Company	**Japan & Asia**	**North America**	**Europe**	**Other**	**Ideal Ratios**	**Relative Ratios**	**True Ratios**
Toyota	**1**	0.665	0.234	0.925	0.629	**0.376**	0.385
General Motors	0.130	**1**	0.273	0.927	**0.545**	**0.326**	0.290
Volkswagen	0.130	0.212	**1**	**1**	0.499	**0.298**	0.325

Notice that for both modes, the calculated composite ratios are fairly close to the true priorities. Such a result would be expected, given that the relative scales for alternatives are perfectly accurate. But notice that the distributive and ideal modes do not produce the true ratios and they give different aggregate results. Which set of priorities, distributive or ideal, should be accepted? The fact that two acceptable modes give different incorrect answers for the same problem should be evidence that perhaps something is wrong with conventional aggregation.

4 Techniques That Address Commensurability

When we derived relative (distributive) and regular (ideal) scales in Table 3, we created different units of measures for each region. Before we can sum those priority values, we must convert them to commensurate units of measure. It is the weights of higher level criteria that re-scale the priorities to commensurable values. This means that the criteria weights must be derived in a very careful manner (Choo et al, 1999).

How to interpret criteria weights was first raised by Watson and Freeling (1982) when they questioned AHP's meaning of "relative importance". Saaty et al (1983) replied with an example where the DM had to refer to the total magnitude of the objects under consideration. In reference to the matrix of paired comparisons of different types of cost criteria, they wrote:

> Each element (i,j) of the matrix gives the ratio of the average (or total) contribution to cost of attribute i to the average (or total) contribution to cost of attribute j.

Watson and Freeling (1983) agreed that such referencing would be correct, but they doubted whether such restrictions are implemented. Their fear is well founded, because conventional AHP has continued to use the un-referenced wording. Saaty accepts referenced measurement for tangibles, but not in other situations (Saaty, 1987).

At about the same time, Belton and Gear (1983) claimed that conventional AHP has a major flaw. They demonstrated that adding or deleting a copy of an alternative could cause a reversal in rank of previous alternatives. This landmark study became the basis for numerous other articles, some contending that reversals occur and are natural (Harker & Vargas, 1987, 1990; Saaty, 1987, 1990) and others discovering that unexplained reversals occur in many situations (Dyer, 1990). In response, the ideal mode was put forward as an AHP method that was immune to reversals (Saaty, 1993). Later, it was discovered that reversals can occur with the ideal mode if the ideal alternative changes (Saaty, 2006).

4.1 Referenced AHP

Schoner and Wedley (1989) adopted the term "Referenced AHP" to describe Saaty's (1983) referenced questioning procedure. Essentially, Referenced AHP is simply the distributive mode re-formulated so that criteria weights are dependent upon the alternatives in the choice set. Schoner and Wedley (1989) posed Referenced AHP questions like the following:

1. Considering the combined total revenue of all alternatives, which Region has more total revenue: Europe or North America?
2. By how many times?

Note that this type of questioning is equivalent to comparing the column totals of Table 2. If done correctly, the criteria weights would be Japan & Asia: 0.234, North America: 0.323, Europe: 0.349, Other: 0.094. Applied to distributed priorities of Table 3, the true composite results are generated.

4.2 Linking Pin AHP

The concepts of "Linking Pin AHP" stem from the original work of Belton and Gear (1983). They modified the criteria questions for the ideal mode so that the DM compares the relative values of the ideal alternatives. Since those ideal alternatives have been established as the unit of measure for each criterion (see Table 3), comparing their relative magnitudes determines criteria priorities that are commensurate for those alternatives across criteria. Other non-ideal alternatives get their correct proportion of the commensurate values during aggregation.

Schoner et al (1993) generalized Belton and Gear's modified approach to any other alternative as the unit link between criteria. With those other units as links between criteria, similar criteria comparisons of the unit alternatives would establish commensurate values for them that could then be passed in proportion to the other alternatives. They called this procedure linking pin AHP to signify that a link alternative acts as a pin that both spans the criteria and distributes commensurate values to alternatives.

Table 4 illustrates linking pin AHP with different alternatives as links. The appropriate criteria comparisons for such links would be between the link alternatives' revenue in each region: 9,922,238 million¥, 1,938,554 million¥, 3,371,427 million¥ and 1,623,852 million¥. Since in most AHP problems we do not know these figures beforehand, the questions would be of the following form.

1. Which company has more sales: Toyota in Japan & Asia or Volkswagen in North America?
2. By how many times?

Because the referent company can change from region to region, it is good policy to select just one company as the link across all regions. With such a common link, the questioning is equivalent to the dependent questions used in ANP and dominant AHP.

Table 4. Linking Pin AHP results with correct criteria weights for links

Criteria Weights:	0.589	0.115	0.200	0.096			
Company	**Japan & Asia**	**North America**	**Europe**	**Other**	**Linked Ratios**	**Relative Ratios**	**True Ratios**
Toyota	**1**	3.145	0.857	0.998	**1.218**	**0.385**	0.385
General Motors	0.130	4.727	**1**	**1**	**0.917**	**0.290**	0.290
Volkswagen	0.130	**1**	3.669	1.079	**1.029**	**0.325**	0.325

4.3 Benchmark AHP

The concept behind Benchmark AHP (Wedley et al, 1996) is that there can be a template hierarchy with correct composite priorities that become an invariant standard for evaluating an unlimited number of other alternatives. For example, Referenced AHP results can be considered the benchmark set and we can add the worldwide sales of Ford Motor Company to it. In 2008, Ford's worldwide sales were $146,277 million. Converted to million¥, the regional sales of Ford are 641,949, 7,362,194, 5,881,977 and 1,283,899 for the four respective regions.

With Ford added, the comparison matrix of each region would be comprised of the three benchmark alternatives (Toyota, General Motors and Volkswagen) and the non-benchmark alternative (Ford). Priorities are calculated for each region, except that the benchmark set would be normalized to sum to one (they are the invariant standard) and we would leave the criteria weights unchanged (they are the weights for the benchmarks). Table 5 gives the results with Ford judged in relation to the benchmarks.

Notice that the benchmarks sum to unity, Ford is a portion of that unity and Ford's composite weight is correctly derived. For example, General Motor's worldwide sales in are 148,979/146,277 = 1.02 greater than Ford. This is the same ratio from composite priorities: 0.290/0.285 = 1.02.

Table 5. Benchmark AHP with Ford added

Criteria Weights:	0.234	0.323	0.349	0.094		
Company	**Japan & Asia**	**North America**	**Europe**	**Other**	**Relative Ratios**	**True Ratios**
Toyota	0.794	0.355	0.155	0.324	**0.385**	0.385
General Motors	0.103	0.533	0.181	0.325	**0.290**	0.290
Volkswagen	0.103	0.113	0.664	0.351	**0.325**	0.325
Ford	**0.051**	**0.428**	**0.316**	**0.257**	**0.285**	

Benchmark measurement can accommodate an unlimited number of alternatives. To avoid comparison matrices from becoming too large, the comparison of each new alternative can be limited to just the benchmark set. If we wanted even fewer comparisons for new alternatives, the benchmark process also works with linking pins. For example, we could use the criteria weights and link units in Table 4 as the benchmarks and express Ford and other non-linked alternatives as proportions of the links.

4.4 Dominant AHP

Dominant AHP and the concurrent convergence method were developed by Kinoshita and Nakanishi (1999) and given a mathematical structure by Kinoshita, Sekitani & Shi (2002). A top down process is used whereby a single regulating (link) alternative is chosen for evaluating criteria and distributing weights to dependent alternatives. The regulating alternative is dominant only in the sense that it controls the establishment of criteria importance and the determination of each alternative's overall priority.

Dominant AHP is similar to a linking pin or benchmark process with a single alternative linking across criteria. Using Toyota as the dominant alternative, the results would be as shown in Table 6. Criteria weights are relative to Toyota as the link. Thus, the questioning procedure is similar to ANP or linking pin AHP with criteria derived for the dominant alternative.

Table 6. Dominant AHP results with Toyota as the regulating alternative

Criteria Weights:	0.483	0.297	0.141	0.079			
Company	**Japan & Asia**	**North America**	**Europe**	**Other**	**Linked Ratios**	**Relative Ratios**	**True Ratios**
Toyota	**1**	**1**	**1**	**1**	**1**	**0.385**	0.385
General Motors	0.130	1.503	1.167	1.002	**0.753**	**0.290**	0.290
Volkswagen	0.130	0.318	4.281	1.081	**0.845**	**0.325**	0.325

Two interesting aspects of dominant AHP relate to a convergence procedure. Given the results in Table 6, it is possible to impute the value of the criteria weights if a different alternative is chosen as the dominant element. Since the criteria weights 0.483, 0.297, 0.141, and 0.079 relate to Toyota that has local priorities of unity, the relative criteria weights for General Motors as dominant should be 0.130*0.483 = 0.063, 1.503*0.297 = 0.446, 1.167*0.141 = 0.164, and 1.002*0.079 = 0.079. Renormalized to unity or used directly, these criteria weights yield the same true ratios.

The second aspect about convergence occurs when multiple dominant alternatives are used concurrently Different regulating alternatives give different composite results and different imputed values of criteria. Kinoshita and Nakanishi (1999) use averages of criteria weights (both estimated and imputed) to recursively determine a stabilized set. This type of convergence can be used on multiple criteria estimates or conflicting alternative estimates. As well, concurrent convergence can be extended to network structures and group situations (Kinoshita, et al (2002).

4.5 Iterative AHP

Zahir (2007) analyzed the rank reversal problem and realized that the norms (units of measure) change when an alternative is added or deleted. He demonstrated that ignoring the unit of measure across the criteria is the root cause for rank reversals, because "…the unity of normalization does not bear the same level of meaning - neither across the criteria nor across the decision space as we add or delete an alternative." His solution was to add a variable to the aggregation equation that captures the relative ratios of norms for the criteria. If an alternative is added or deleted, concurrent information from local priorities can be used to determine new norms.

Zahir's approach is iterative in the sense that alternatives can be added or deleted without changes in the ranks or ratios of alternatives. Norms across criteria only have to be captured once and results thereafter can be calculated from the relative priorities of alternatives. Zahir's very efficient aggregation formula are a product of the alternative priorities, the norms and criteria weights. While his procedure is innovative, he is not clear about the purpose that norms and criteria weights play in the same aggregation equation. In applying his equation to the car manufacturer example, only equal criteria weights allowed the true ratios to be emulated. This is because the norm weights capture the commensurability across columns, not the criteria. This issue needs to be cleared up, because Zahir's method with differential criteria weights may be preserving the wrong ratios.

5 Benefit Cost Analysis

In AHP/ANP, Benefit/Cost priorities (Bp/Cp) replace the usual monetary values (B$/C$) of benefit cost analysis. When this is done, there is a potential problem. While priorities are measured relative to each other, it is not clear that the two sets of priorities are in the same unit of measure. It is quite possible that the actual costs far exceed the actual benefits, yet the Bp/Cp ratio appears positive (Wedley et al, 2001). To address this problem, Bernhard and Canada (1990) proposed a cut-off ratio based upon values measured in dollars. Saaty (1994) has suggested that the DM should ask questions such as “do the benefits justify the costs?” and whether they are sufficiently commensurable (Saaty, 1996). Wedley et al (2001) suggested a magnitude adjustment procedure that converts the benefits and costs to a common unit. That way, commensurable values are compared.

Despite such concerns, AHP/ANP continues to approve relative ratio scales for benefits, opportunities, costs and risks. Wijnmalen (2007) studied the conventional BOCR equations and concluded that none of them gives reliable measures of profitability. Attributing this deficiency to the failure to make the BOCR components commensurable, he demonstrated that revised equations with adjustment weights would give correct indications of profitability. As well, he verified Millet and Schoner’s (2005) contention that taking reciprocals of cost and risk priorities is an impermissible transformation. It is now generally acknowledged that reciprocals should be avoided and negative values allowed (Millet & Schoner, 2005, Saaty & Ozdemir, 2003).

6 Analytic Network Process (ANP)

The ANP (Saaty, 1996) is the general framework for both analytic hierarchies and networks. It utilizes the concept of a supermatrix where the relative ratio scales associated with a node are represented in a row and column. Each vector is placed in a column representing the destination (influenced) node and rows the source of the influence. Several priority scales can be placed in a column, signifying that several factors influence that destination node. In such cases the columns are renormalized to sum to one, yielding a columnar stochastic matrix. This weighted supermatrix raised to sufficient powers stabilizes with values that incorporate all the influences in the network.

Table 7 presents the initial ANP supermatrix and the resulting stabilized values.ANP adopts a network structure in which the criteria weights are influenced by the alternatives and vice versa. Local alternative priorities given a criterion are in the lower left quadrant and criteria weights given a link alternative are in the top right quadrant. These criteria priorities are derived in exactly the same as in dominant AHP. Notice that the criteria weights for Toyota are exactly the same as dominant AHP in Table 6.

Table 7. ANP Supermatrix and stabilized results

	Japan & Asia	North America	Europe	Other	Toyota	General Motors	Volks-wagen	Stabilized priorities
Japan & Asia	0	0	0	0	0.483	0.084	0.074	**0.234**
North America	0	0	0	0	0.297	0.593	0.112	**0.323**
Europe	0	0	0	0	0.141	0.218	0.713	**0.349**
Other	0	0	0	0	0.079	0.105	0.101	**0.094**
Toyota	0.794	0.355	0.155	0.324	0	0	0	**0.385**
General Motors	0.103	0.533	0.181	0.325	0	0	0	**0.290**
Volkswagen	0.103	0.113	0.664	0.351	0	0	0	**0.325**

Observe that the stabilized ANP results give the correct overall priorities: 0.385, 0.290, 0.325. As well, the stabilized criteria weights represent the overall criteria priorities that we should strive to achieve if we were using hierarchical distribution. They are the same as Reference AHP.

ANP is considered to be the general form, because it will also give the same answer as an AHP formulation. If we entered distributive weights in the criteria columns, the stabilized priorities would be the same as distributive mode shown in Table 3. We note however that unreferenced distributive questions fail to produce the true priorities. Had we inserted referenced AHP weights in the supermatrix, then the stabilized results for alternatives would have been the true overall priorities. The implication is that ANP gives correct results only when dependencies are considered.

7 Discussion

AHP/ANP makes use of both regular ratio scales (unit is defined) and relative ratio scales (no specific object identified as the unit). Most problems in AHP/ANP applications arise because the unit of measure is overlooked or ignored. Triantaphyllou (2001) for example, decomposed a hierarchy into several sub-problems of 2 alternatives without realizing that the renormalization of alternative priorities changes the unit used for criteria comparisons. Similarly, Finan and Hurley (2002) did not realize that the removal of non-discriminating (wash) criteria affects the link between normalization and associated alternatives. Had these authors recognized units of measure, commensurability and the need to re-adjust criteria weights, they may reached more correct conclusions about how AHP/ANP avoids unwarranted rank reversals (Wijnmalen & Wedley, 2008a, 2008b).

Table 8 illustrates what happens when another alternative (Ford) is added to the ANP structure. As compared to Table 7, notice that the values of the local priorities in the bottom left quadrant have changed but their relative ratios have not. In both tables, Toyota/GM = 7.7 in Japan and Asia. By adding an alternative, we have made the relative ratio scales for alternatives relative to n+1=4 companies rather than the former 3 companies. The different priority values for the same alternatives in Tables 7 and 8 signify that a different unit of measure is being used in each table.

Table 8. ANP Supermatrix and stabilized results with Ford Included

	Japan & Asia	North America	Europe	Other	Toyota	General Motors	Volks-wagen	Ford	Stabilized priorities
Japan & Asia	0	0	0	0	0.483	0.084	0.074	0.042	0.192
North America	0	0	0	0	0.297	0.593	0.112	0.485	0.359
Europe	0	0	0	0	0.141	0.218	0.713	0.388	0.358
Other	0	0	0	0	0.079	0.105	0.101	0.085	0.092
Toyota	0.755	0.248	0.118	0.258	0	0	0	0	0.300
General Motors	0.098	0.373	0.138	0.259	0	0	0	0	0.226
Volkswagen	0.098	0.079	0.505	0.279	0	0	0	0	0.253
Ford	0.049	0.300	0.240	0.204	0	0	0	0	0.221

With Ford added, the overall stabilized priorities for alternatives also maintain their former ratios (Toyota is still 1.328 times larger than GM). The stabilized criteria weights, however, have changed ratios. With Ford added, the importance of North America for judging worldwide sales is 0.359/0.192 = 1.87 times the importance for Japan & Asia whereas it was formerly 0.323/0.234 = 1.38. If we were analyzing the problem from an AHP perspective, these stabilized criteria weights are what we would have to use in the distributive mode, referenced AHP or the n criteria columns of the supermatrix to generate the correct results. The implication is that with an alternative added or deleted, the criteria weights must be adjusted to reflect the different totality represented by the alternatives.

It is difficult for the distributive and ideal modes to capture correct criteria weights. With the assumption of independence, those methods use questions that are unrelated to the alternatives and norms that are being considered. Accordingly, it is unlikely that the correct weights are attained. Since those two methods express alternatives in different units, their different norms require different criteria weights to capture the true ratios. This need for criteria weights to match the norms is shown in Tables 3 to 6 where different units are used.

To help decide whether criteria are really independent of alternatives, Harker and Vargas (1990) suggested the following test. Conduct ANP type questioning for criteria with two or more alternatives as the referent object. If the criteria weights are invariant to the selected referent alternative, then it can be concluded that independence exists. Note that in this test there is no automatic assumption that invariance and independence exists -- it must be proved by questioning the decision maker.

When we consider this type of dependent questioning, it would be very rare for all the criteria columns of the supermatrix to be the same. In some situations such as establishing weights before the alternatives are known, it is possible that the criteria are truly independent of the alternatives (Wedley & Choo, 2008). That circumstance, however, is very rare. The different imputed criteria weights of dominant AHP prove this fact. For AHP problems, a more cautious rule is to assume dependence and let independence emerge if it exists. That is what happens when we use the general case, ANP. And, as Harker and Vargas (1990) recognized, if independence exists, it will emerge from dependent questions.

If we assume dependence in AHP, then the distributive mode should adopt Referenced AHP questions – criteria weights should be established in reference to the totality of the criterion possessed by the alternatives. As well, criteria weights should be re-evaluated if there is an addition or deletion. If criteria weights remain unchanged upon addition or deletion, then perhaps independence exists. That, however, is based upon the behavior of the DM and not the assumption that independence is the norm.

As Kinoshita and Sugiura (2008) have shown, referenced AHP, linking pin AHP, dominant AHP and a simple supermatrix with ANP all yield the same solution. The common trait between them is that they all assume criteria weights to be dependent upon the alternatives. The tendency in AHP/ANP evolution is to consider more complex feedback models (i. e. ANP) that capture both dependent and independent relationships. While that process works, we should remember that the dependent methods with AHP require less effort (fewer comparisons) and a structure that is easier to understand. It is probably prudent, therefore, to keep AHP procedures for simpler structures and use ANP for only the most complex problems.

Like ANP, dominant AHP can converge to a stabilized result. An advantage of dominant AHP is that the convergence can occur while maintaining a better understanding of the unit of measure. In ANP, the nature of the unit becomes lost when columns are made stochastic and the resulting super matrix is raised to powers. Although ANP has good face validation (Whitaker, 2007), the process is like a black box within which we do not understand what is going on.

Dominant AHP is similar to linking pin AHP in the use of regular and relative ratio scales. It goes well beyond linking pins, however, by the choice of different regulating alternatives and a concurrent convergence method that emulates ANP. Whether dominant AHP can handle more complex models like ANP is an area that warrants further investigation. Perhaps one way to approach this possibility is to attempt to analyze more complex ANP examples with dominant AHP procedures.

Another area of investigation is whether dominant AHP or ANP can be adapted to become a benchmark mode that accommodates numerous alternatives. Benchmark AHP can already handle many alternatives and linking pin and iterative AHP can do the same with very little modification. If such procedures can be developed for relative measurement methods, then they can become competing and perhaps more accurate substitutes for absolute measurement. In current AHP/ANP applications, absolute measurement is often assumed for situations that are really relative.

8 Conclusion

An important distinction has been made between regular ratio scales and relative ratio scales. All ratio scales are relative to some “thing”. In regular ratio scales, that “thing” is a specific object. In relative ratio scales, that “thing” is the group of n objects under consideration. They are measured relative to each other. In effect, the “thing” that is the unit of measure for a relative ratio scale is an abstract object that has the totality of the attribute possessed by the n objects. Knowing just the relative values of the n objects is partial information. To provide complete information, we need to know both the relative values and the “thing” they are related to.

The advantage of a relative ratio scale is that it is unique to the n objects. The disadvantage is that the unit of measure is abstract, obtuse and obscure. So long as only one property is being considered, relative ratio scales are not a problem. Adding or deleting an object creates a new uniqueness, but it does not upset pre-existing ratios. It is commonly acknowledged that rank reversal never occurs in a single criterion problem.

The situation is different, however, when aggregating several properties associated with different criteria. Adding or deleting an object affects each property in a different manner – a new abstract object represents the unit of measure for each criterion. With new units of measure, the former aggregation across criteria will no longer be appropriate. We will have to reassess criteria weights to re-capture commensurability or leave the former units unchanged so the existing commensurability is not upset.

All techniques except the distributive and ideal modes endeavor to achieve commensurability. Referenced AHP does it by reconsidering criteria weights if the structure is altered. Linking pin and benchmark AHP does it by keeping a fixed structure and a specific unit. Dominant AHP with concurrent convergence does it by sampling with multiple regulating alternatives. Iterative AHP does it by calculating the change in norms. ANP does it by exhaustive linkages in a network structure. All these dependent methods are appropriate for MCDM with ratio measurement.

Acknowledgment. The Natural Sciences and Engineering Research Council of Canada provided the financial support necessary for this project (RGPIN 123621).

References

Belton, V., Gear, A.E.: On a shortcoming of Saaty's method of analytic hierarchies. Omega 11, 228–230 (1983)

Bernhard, R.H., Canada, J.R.: Some problems in using benefit/cost ratios with the Analytic Hierarchy Process. The Engineering Economist 36(1), 56–65 (1990)

Choo, E.U., Schoner, B., Wedley, W.C.: Interpretation of criteria weights in multi-criteria decision making. Computers and Industrial Engineering Journal 37, 527–541 (1999)

Dyer, J.S.: Remarks on the analytic hierarchy process. Management Science 36(3), 249–258 (1990)

Finan, J.S., Hurley, W.J.: The analytic hierarchy process: can wash criteria be ignored? Computers & Operations Research 29, 1025–1030 (2002)

Ford Motor Company, Annual Report (2008)

General Motors Corp., Form 10-K, March 5 (2009)

Harker, P.T., Vargas, L.G.: The theory of ratio scale estimation: Saaty's analytic hierarchy process. Management Science 33(11), 1383–1403 (1987)

Harker, P.T., Vargas, L.G.: Reply to Remarks on the analytic hierarchy process. J. S. Dyer, Management Science 36(3), 269–275 (1990)

Kinoshita, E., Nakanishi, M.: Proposal of new AHP model in light of dominant relationship among alternatives. Journal of the Operations Research Society of Japan 42(2), 180–197 (1999)

Kinoshita, E., Sekitani, K., Shi, J.: Mathematical properties of dominant AHP and concurrent convergence method. Journal of the Operations Research Society of Japan 45(2), 198–213 (2002)

Kinoshita, E., Sugiura, S.: A comparison of dominant AHP and similar dominant models. Journal of Research Institute of Meijo University 7, 113–116 (2008)

Michell, J.: Quantitative science and the definition of measurement in psychology. British Journal of Psychology 88, 355–383 (1997)

Michell, J.: Measurement in Psychology. Cambridge University Press, Cambridge (1999)

Millet, I., Schoner, B.: Incorporating negative values into the Analytic Hierarchy Process. Computers and Operations Research 32, 3163–3173 (2005)

Saaty, T.L.: Rank generation, preservation, and reversal in the analytic hierarchy decision process. Decision Sciences 18, 157–177 (1987)

Saaty, T.L.: An exposition of the AHP in reply to the paper Remarks on the analytic hierarchy process. Management Science 36(3), 259–268 (1990)

Saaty, T.L.: Rank preservation and reversal: the AHP ideal and distributive modes. Mathematical and Computer Modelling 17(6), 1–16 (1993)

Saaty, T.L.: Fundamental of decision making and priority theory. RWS Publications, Pittsburgh (1994)

Saaty, T.L.: Decision Making with Dependence and Feedback: The Analytic Network Process. RWS Publications, Pittsburgh (1996)

Saaty, T.L.: Scales from measurement – not measurement from scales. In: Proceedings of the 17th International Conference on Multiple Criteria Decision Making, Whistler, B. C. Canada, August 6-11 (2004), `http://www.bus.sfu.ca/events/mcdm/Proceedings/MCDM2004%20SUM.htm` (Accessed May 22, 2010)

Saaty, T.L.: Rank from comparisons and from ratings in the analytic hierarchy/network processes. European Journal of Operational Research 186, 557–570 (2006)

Saaty, T.L.: Relative measurement and its generalization in decision-making – Why pairwise comparisons are central in mathematics for the measurement of intangible factors: The Analytic Hierarchy/Network Process. RACSAM, Revista de la Real Academia de Ciencias Serie A: Matemáticas 102(2), 251–318 (2008)

Saaty, T.L., Ozdemir, M.: Negative priorities in the Analytic Hierarchy Process. Mathematical and Computer Modelling 37, 1063–1075 (2003)

Saaty, T.L., Sodenkamp, M.: Making decisions in hierarchic and network systems. International Journal of Applied Decision Sciences 1(1), 24–78 (2008)

Saaty, T.L., Vargas, L.G., Wendell, R.E.: Assessing attribute weights by ratios. Omega 11(1), 9–12 (1983)

Schoner, B., Wedley, W.C.: Ambiguous Criteria Weights in AHP: Consequences and Solutions. Decision Sciences 20(3), 462–475 (1989)

Schoner, B., Wedley, W.C., Choo, E.U.: A unified approach to AHP with linking pins. European Journal of Operational Research 64, 384–392 (1993)

Toyota Motor Corporation, FY 2009 Financial Summary

Triantaphyllou, E.: Two new cases of rank reversals when the AHP and some of its additive variants are used that do not occur with the multiplicative AHP. Journal of Multi-Criteria Decision Analysis 10, 11–25 (2001)

Volkswagen Financial Services, Annual Report (2008)

Watson, S.R., Freeling, A.N.S.: Assessing attribute weights by ratios. Omega 10(6), 582–583 (1982)

Watson, S.R., Freeling, A.N.S.: Comment on: Assessing attribute weights by ratios. Omega 11, 13 (1983)

Wedley, W.C., Choo, E.U., Schoner, B.: Benchmark measurement: Between relative and absolute. In: Proceedings of the Fourth International Symposium on the Analytic Hierarchy Process, Burnaby, BC, Canada, pp. 335–345 (1996)

Wedley, W.C., Choo, E.U., Schoner, B.: Magnitude adjustment for AHP benefit/cost ratios. European Journal of Operational Research 133, 342–351 (2001)

Wedley, W.C., Choo, E.U.: Sharing the pie and distributing the reward: a case of bonus distribution among executives. Journal of Multi-Criteria Decision Analysis 15(5-6), 151–162 (2008)

Wijnmalen, D.J.D.: Analysis of benefits, opportunities, cost and risks (BOCR) with the AHP-ANP: a critical validation. Mathematical and Computer Modelling 46, 892–905 (2007)

Wijnmalen, D.J.D., Wedley, W.C.: Non-discriminating Criteria in the AHP: removal and rank reversal. Journal of Multi-Criteria Decision Analysis 15(5-6), 143–149 (2008)

Wijnmalen, D.J.D., Wedley, W.C.: Correcting illegitimate rank reversals: proper adjustments of criteria weights prevent alleged AHP intransitivity. Journal of Multi-Criteria Decision Analysis 15(5-6), 135–141 (2008)

Whitaker, R.: Validation examples of the Analytic Hierarchy Process and Analytic Network Process. Mathematical and Computer Modelling 46, 840–859 (2007)

Zahir, S.: A new approach to understanding and finding remedy for rank reversals in the additive Analytic Hierarchy Process. In: Conference Proceedings, Administrative Sciences Association of Canada, Ottawa (2007), `http://attila.acadiau.ca/library/ASAC/v28/02/index.html` (Accessed May 22, 2010)

An Application of Dominant Method: Empirical Approach to Public Sector Reform

Yuji Sato

Graduate School of Policy Science, Mie Chukyo University, 1846, Kubo, Matsusaka, Mie, 515-8511 Japan

Abstract. Under tight financial conditions, local governments in Japan are required to conduct performance-based administrative management and subsequent public sector reform. Concerning public sector reform, a joint research group proposed a framework of how to reform public sector with theoretical approach. The research concluded that present public sector should be decomposed into two parts: one is the organization specialized in decision-making and ruling with authority and legal power; the other is collective entities which coproductively cover social issues. On the other hand, by employing the dominant method of the Analytic Hierarchy Process, the author of this paper developed a rational approach to the delegation of power from the public sector to other potential sectors. Based on the results of the study, this paper considers the scheme of public sector reform. Principal component analysis concerning properties featuring projects of a local government is carried out, and new concepts representing the aspect of the projects are extracted. The results of this paper illustrate the structure of the publicness and show the scheme of public sector reform.

1 Introduction

The declination of the fiscal condition of governments has its origin in the collapse of the bubble economy in the early 90s in Japan. The total amount of government bonds increased almost threefold, from 2.14 trillion to 6.07 trillion USD (1 USD = 90 JPY) during these 15 years; the total governmental debt to GDP ratio is 172.1%, while that of the United States stays 71.1% as of the end of 2008. Those situations require governments to enhance the efficiency of the public administration in conjunction with conducting proper evaluation of governments' activities. In connection with the passage of the Government Agencies Policy Evaluation Law in 2002; every government organization in Japan is obliged to evaluate its policies and announce the results to the public annually. The law has made evaluating the achievement of policy goals a critical component of project implementation, sometimes resulting in public sector reform (Sato 2007).

An evaluation of public activities consists of a series of processes evaluating the adequacy and achievement of the policies and measures of public administration against a specific benchmark. In Japan, a prefecture first introduced a rigorous administrative evaluation system in 1996. Since then, some local governments have begun to introduce a variety of such evaluation systems. As of the end of September 2008,

G. Phillips-Wren et al. (Eds.): Advances in Intel. Decision Technologies, SIST 4, pp. 43–51.
springerlink.com

all prefectural and city governments, along with more than 90% of the local authorities, have introduced their original evaluation systems. (Ono and Tabuchi 2001). On the other hand, public sector reform encompasses a spectrum of issues including civil service reform, performance measurement, and governance regulations.

The separation of implementation of administrative activities from decision-making has been diffused as "agencification" around the world. The central government of Japan has also adopted such reform where the agencies are called "independent administrative institutions" (Yamamoto 2008). A great many projects of the local government, however, have not yet been reformed, despite the fact that local public finance is on the verge of a crisis. The main reason for the deadlock in reform is considered to be the difficulty in obtaining a rationale for delegating authority from the public sector to other potential sectors. In the context of public sector reform, assuming that the public sector would specialize in decision-making and ruling, the joint research group of prefectures in Japan clarified that the size and authority of administrative organizations should be reduced and that an agency or other potential sector would take on an important role, particularly in executing projects (2003).

On the other hand, Sato (2007) proposed a rational approach to administrative evaluation in conjunction with a prospective way to the delegation of power from the public sector to other potential sectors. Specifically, the matching process of projects of a local government with alternative sectors rationally by using the dominant method (Kinoshita and Nakanishi 1999; Kinoshita et al. 2002) of the Analytic Hierarchy Process (AHP) (Saaty 1980) was developed. The process clarified the property of each project of a local government, and the characteristic of alternative sectors including the public sector. Based on the results of this study, principal component analysis concerning properties featuring projects of a local government is carried out in this paper. New concepts representing the aspect of the projects are extracted, which illustrate the structure of publicness and show the scheme of public sector reform.

2 Summary of the Joint Research (2003)

The joint research (2003) proposed the framework of how to reform the public sector with theoretical approach, which could be evaluated as the refinement of New Public Management (NPM). Assuming that the public sector would specialize in decision-making and ruling, the research clarified that the size and the authority of administrative organizations should be reduced and that an agency or other potential sector would take on an important role, particularly in executing projects (2003). In other words, present public sector should be decomposed into two parts: one is the organization specialized in decision-making and ruling, which contributes public welfare with authority and legal power, called "refined public;" the other is collective entities which coproductively cover social issues such as agency or network sector, called "intermediate sector." Thus, the research concluded that present public-private binary social structure should change and improve to ternary structure consisting of refined public sector, intermediate sector and private sector.

Present social structure could be illustrated as the public-private binary relation, which consists of two bodies: public and private sectors divided whether they partake the property of legal power. Fig. 1 is the pattern diagram illustrating the goal of social structure deduced in the joint research. As shown in the figure, alternative sectors, such as the public sector, semi-public sector, agency, network sector and private sector, are located with reference to a marketability axis and a legal power axis. In the diagram, semi-public sector ranges in the first quadrant whose degree of marketability and that of legal power are high. In the same way, the diagram shows as follows: in the second quadrant, the public sector, agency and non-profit organization are located; in the third quadrant, network sector ranges; in the fourth quadrant, private sector is located. In the joint research, this figure was contrived as the goal of social structure which aimed to vitalize a local community by empowering potential sectors among the community.

3 Outline of the Previous Study (2007)

This section summarizes the outline of the study (2007); the matching process and the subsequent results are shown. Those will be the basis of principal component analysis in Sect. 4, which clarifies the scheme of public sector reform.

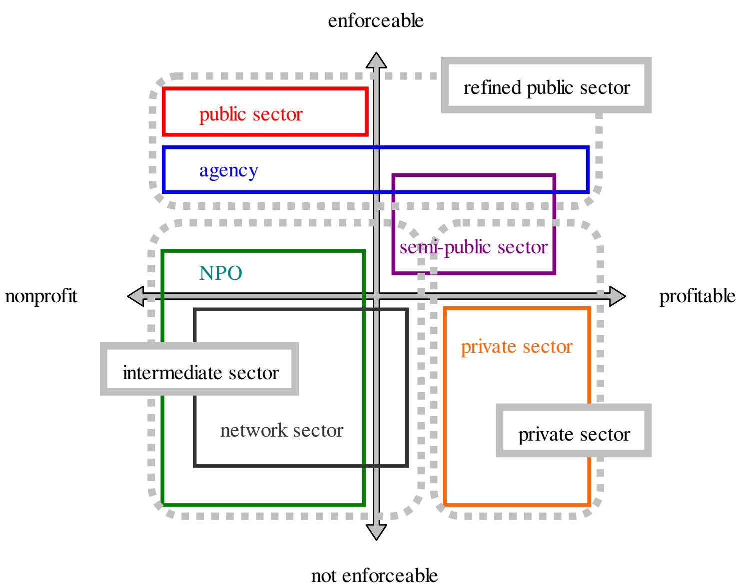

Fig. 1. The goal of improved social structure: ternary structure

A system of public administration of a local government in Japan can be stratified as a hierarchical structure, consisting of philosophies, policies, measures, projects, and operations from the top to the bottom; all these operations are administered by the public sector at present. Among these operations, the study focused on projects as the object of evaluation. Inasmuch as delegating authority of the public sector to other potential sectors would be crucial for successful reform; some sectors including the public sector were focused on as the alternatives to be delegated the power in the study. The matching process consisted of three steps shown below. In the process all the projects were evaluated based on seven properties defining the publicness listed below; these properties were listed from the viewpoint of appropriateness for the aforementioned evaluations and characterizations by interviewing the executives of a local government.

Seven properties

- *Authority*: the power or right to give orders or make decisions
- *Efficiency*: skill in avoiding wasted time and effort
- *Coproductivity*: doing with or working with others for a common purpose
- *Publicity*: benefit or utility for the people or community
- *Fairness*: ability to make judgments free from discrimination or dishonesty
- *Profitability*: the quality of affording gain or benefit or profit
- *Disciplinary*: having quality relating to a specific field

Step 1—Evaluation of projects
All (266) the projects were evaluated based on the above mentioned seven properties. In this step, 17 executives of a local government were asked to evaluate projects they were in charge of. In a practical sense, they conducted pairwise comparisons from the viewpoint of the seven properties by the dominant method for each project. In the process of pairwise comparisons, the key concept featuring the public sector, *Authority*, was selected as the dominant alternative and was compared with others. Then, the evaluation of each project was derived as a weight vector, called *e-vector* which was normalized by l_1-norm and denoted e_i={weight of property j}, (i=1, …, 266; j=*Authority*, *Efficiency*, *Coproductivity*, *Publicity*, *Fairness*, *Profitability*, *Disciplinary*).

Left out the details here, this step concluded that *Publicity* and *Coproductivity* were highly weighted, while *Profitability* and *Authority* were not in the aggregate.

Step 2—Characterization of alternative sectors
The four alternative sectors were characterized by the seven properties which were the same as those employed in Step 1. In this step, the executives of a local government were asked to characterize alternative sectors; they individually conducted pairwise comparisons of the seven properties for each sector. Based on the pairwise comparison matrices of executives, the geometric mean of each element of the matrices were calculated and one pairwise comparison matrix was generated as a consensus-based matrix. Then the characteristic of each sector was elicited as a weight vector, called *c-vector* which was normalized by l_1-norm and denoted c_j={weight of sector k}, (j=*Authority*, *Efficiency*, *Coproductivity*, *Publicity*, *Fairness*, *Profitability*, *Disciplinary*; k=public sector, semi-public sector, agency, private sector). Furthermore, the matrix (c_j^{T}) could be defined as the characteristic matrix of each sector, called *C-matrix* and denoted *C*.

Left out the details of *C-matrix* here, each sector has its dominant characteristics; the public sector excels at *Authority*, *Publicity* and *Fairness*; the semi-public sector does at *Coproductivity*, and so on.

Step 3—Overall judgment
An overall judgment concerning the matching of projects with alternative sectors was derived by taking the weighted average of the evaluations of each project and the characteristics of sectors. $C \cdot e_i^{\mathrm{T}}$ was calculated as the overall judgment for project i (i=1, …, 266) in this step.

As the final output of the matching process proposed in the study, Step 3 read out the overall judgment concerning which project should be administered by which sector. The study implied that 93 out of 266 projects should be devolved from the public sector to other sectors. The details are; 38 projects (14.3%) should be devolve to a semi-public sector, 11 (4.1%) to an agency and 44 (16.5%) to a private sector.

These results give quantitative outputs to administrative evaluation and rational bases for public sector reform. Based on the interview following these analyses, both the matching process and the results of this study are persuasive for administrative officials of a local government. Left out he details here, more than 70% of the respondents thought that the outputs of the matching process were true or rather true.

4 Verification of the Framework of How to Reform Public Sector

In this section, principal component analysis concerning the feature of publicness is carried out based on the results the study (2007), and the framework of how to reform public sector implied by the joint research (2003) is verified by extracted new concepts representing the aspect of the projects of a local government.

4.1 Results of Principal Component Analysis

In the analysis, the outputs of the Step 1 of the previous study, 266 eigenvectors e_i (i=1, …, 266) not-normalized by l_1-norm, are employed. Cumulative proportion of the top four components reaches 77.56%; those four components, C_m (m=1, …, 4), therefore, are employed to represent the aspect of the projects of a local government. Table 1 shows the component score of each property; the maximum of absolute value of each column are boldfaced. As shown in the table, each component has its own characteristics, which could be defined as follows; C_1: degree of involvement of markets; C_2: degree of controllability of markets; C_3: degree of freedom of new entries; C_4: degree of openness of markets.

Fig. 2 is the scatter diagram of all the projects plotted on C_1C_2-plane. As illustrated in the figure, projects are distributed all over the plane. Among those projects, some form a group in the third quadrant, for instance. Such projects as their degrees of the involvement of markets nor those of the controllability of markets are neither low, are estimated to range in the area. In fact, the previous study concluded that 27 among the 32 projects included in the circle should be transferred their operations to a semi-public sector; a project, "the promotion of activities for community welfare," is one of such projects. On the other hand, only 5 out of the 32 projects were concluded to remain under control of the public sector. These results well explains that each principal component represents the aspect of the projects.

Table 1. Component score of each property

component	Authority	Efficiency	Coproductivity	Publicity	Fairness	Profitability	Disciplinary
1	0.0082	**0.6570**	-0.5268	-0.5339	0.0817	**0.6430**	0.5764
2	**0.7158**	-0.1798	**-0.7394**	0.4376	0.3336	-0.2566	0.1633
3	0.2286	-0.3573	0.2958	**-0.6674**	0.4955	-0.2733	0.2909
4	-0.0594	0.2415	-0.0215	-0.0070	**0.7225**	0.2329	**-0.6628**

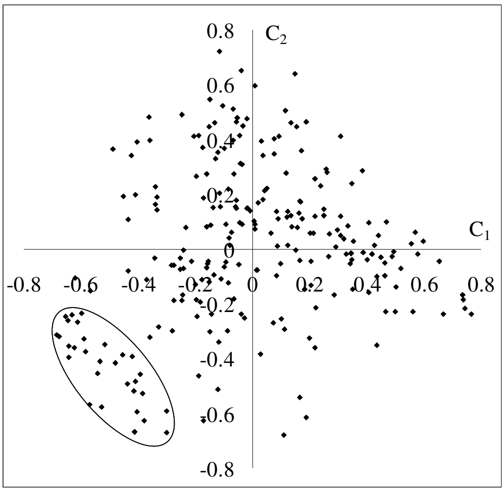

Fig. 2. Distribution of all the projects on C_1C_2-plane

Left out the details of C_1C_3-plane and C_1C_4-plane here, every quadrant in both planes has its own characteristic. Some projects form a group which partakes reasonable property in conjunction with the region where the group ranges, which coincides the result of the previous study. Some actual projects with typical properties could be found in each group

4.2 Scheme of Public Sector Reform

Based on the previous study which clarified the property of each project of a local government and the characteristic of potential sectors, this paper reveals the concept of publicness on the plane of new coordinates which is inductively derived from principal component analysis. Fig. 3 is the scatter diagram of the seven properties featuring publicness with reference to a C_1-axis and a C_2-axis. The layout of each property plotted in the figure is extracted on the basis of the administrative evaluation of projects. Since the evaluation discussed an appropriate publicness, the figure could be interpreted as the goal of social structure.

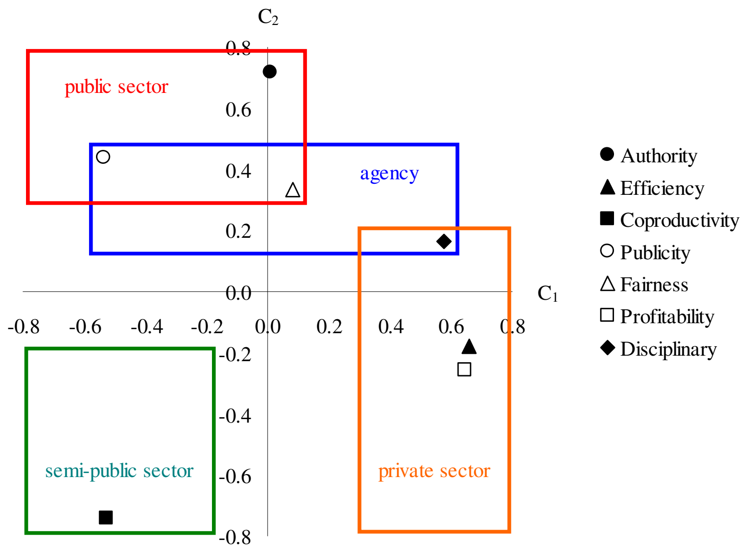

Fig. 3. Distribution of the seven properties on C_1C_2-plane

As shown in Fig. 3, the C_1C_2-plane roughly consists of the following four regions, which could be named as regions of the public sector, agency, semi-public sector and private sector with the following reasons. The region in the second quadrant is characterized by the properties of *Authority*, *Publicity* and *Fairness*, and its degree of C_1 is low and that of C_2 is high. In fact, with reference to Fig. 2, the region includes such projects as to be administrated by the public sector. Thus the area could be interpreted as the public sector region. In the same way, each regions can be interpreted as follows. The region ranging over the first and the second quadrant is featured by the properties of *Disciplinary*, *Fairness* and *Publicity*, and its degree of C_2 is medium. The region indeed, consists of such projects as to be administrated by an agency, which could be considered as the agency region. The region in the third quadrant is characterized by the property of *Coproductivity*, and its degree of C_1 nor that of C_2 is neither low. In fact, the region includes such projects as to be administrated by a semi-public sector. Thus the area could be interpreted as the semi-pubic sector region. The region in the fourth quadrant is featured by the properties of *Disciplinary*, *Efficiency* and *Profitability*, and its degree of C_1 is high and that of C_2 is low. The region indeed, consists of such projects as to be administrated by a private sector, which could be considered as the private sector region.

As described in Sect. 2, the joint research deduced a framework of how to reform the public sector with theoretical approach, which implied that present public sector should be decomposed into refined public and intermediate sectors as Fig. 1 in Sect. 2 illustrates. The figure concluded that present public-private binary social structure should change and improve to ternary structure consisting of refined public sector, intermediate sector and private sector. On the other hand, this paper carried out principal component analysis concerning the properties featuring the projects of a local government based on the results the study. Fig. 3 shows the structure of publicness and illustrates the scheme of public sector reform.

Compared Fig. 1 with Fig. 3, the concept of coordinate axes in the figures could be similarly interpreted; horizontal axes: profit-nonprofit in Fig. 1 and the degree of involvement of markets in Fig. 3, vertical axes: enforceable-not enforceable in Fig. 1 and the degree of controllability of markets in Fig. 3. In addition to the similarity of the concept of coordinate axes, the layout of alternative sectors in Fig. 1 and that of the four regions in Fig. 3 is almost the same. Thus, the implication of Fig. 3 could be considered almost the same as that of Fig. 1, which ensures that the result of this paper empirically verifies the joint research. The approaches taken to public sector reform were different between the joint research which is theoretical and this paper which is empirical; the implication, however, is the same with each other.

5 Concluding Remarks

In this paper, we first focused on the framework of how to reform the public sector proposed in the joint research (2003). The research concluded that present public sector should be decomposed into refined public and intermediate sectors, and the public-private binary social structure should change and improve to ternary structure consisting of refined public sector, intermediate sector and private sector. Then, we focused on the quantification of criteria employed in the study (2007), which leads to a rational approach for prospective way to delegate authority from the public sector to other potential sectors in the process of public sector reform.

Based on the results the previous study, principal component analysis concerning properties featuring the projects of a local government was carried out in this paper. The new concepts representing the aspect of the projects of a local government were extracted, which illustrated the structure of the publicness and showed the scheme of public sector reform. As a result, the analysis empirically verified the implication of the above mentioned joint research. The approaches of the research and the study were different; while the implications coincide with each other.

On the other hand, we seem to have other critical problems in the execution of public sector reform, according to interviews following this study. The results derived from this study were persuasive for administrative officials; one executive, however, said, "Evaluating administrative activities is one thing; implementing reformations is another." Indeed, almost all local governments evaluate their administrative activities to a greater or lesser degree, a great many projects of the local government have not yet been reformed. We, thus have the following open-ended questions: how to deal with so-called political maneuvers, which sometimes leave conventional projects unchanged; how to deal with the nature of organization and its personnel, clinging to vested interests that prevent reformation within local government. Further analyses based on case studies are necessary.

References

Joint Research Group of Fukui, Gifu, Mie, Shiga Prefectures: Developing New Regional Management . In: 2002 Annual Report on Policy Analysis, Shiga Prefecture (2003) (in Japanese)

Kinoshita, E., Nakanishi, M.: Proposal of New AHP Model in Light of Dominant Relationship among Alternatives. J. Oper. Res. Soc. Jpn. 42(2), 11–19 (1999)

Kinoshita, E., Nakanishi, M., Shi, J.: Mathematical Properties of Dominant AHP and Concurrent Convergence Method. J. Oper. Res. Soc. Jpn. 45(2), 198–213 (2002)

Ono, T., Tabuchi, Y.: Handbook of Public Sector Evaluation. TOYO Keizai, Tokyo (2001) (in Japanese)

Saaty, T.L.: The Analytic Hierarchy Process. McGraw-Hill, New York (1980)

Sato, Y.: Comparison Between Multiple-choice and Analytic Hierarchy Process: Measuring Human Perception. Int. Trans. Opl. Res. 11(1), 77–86 (2004)

Sato, Y.: Administrative Evaluation and Public Sector Reform: An Analytic Hierarchy Process Approach. Int. Trans. Opl. Res. 14(5), 445–453 (2007)

Yamamoto, K.: Has Agencification Succeeded in Public Sector Reform? Realities and Rhetoric in the Case of Japan. Asian J. Political Sci. 16(1), 24–40 (2008)

General Application of a Decision Support Framework for Software Testing Using Artificial Intelligence Techniques

Deane Larkman[1], Masoud Mohammadian[1], Bala Balachandran[1], and Ric Jentzsch[2]

[1] Faculty of Information Science and Engineering, University of Canberra, ACT, Australia
u950777@uni.canberra.edu.au,
Masoud.Mohammadian@canberra.edu.au,
Bala.Balachandran@canberra.edu.au

[2] Business Planning Associates Pty Ltd, ACT, Australia
Bus.Planning.Assoc@gmail.com

Abstract. The use of artificial intelligent (AI) techniques for testing software applications has been investigated for over a decade. This paper proposes a framework to assist test managers to evaluate the use of AI techniques as a potential tool to test software. The framework is designed to facilitate decision making and provoke the decision maker into a better understanding of the use of AI techniques as a testing tool. We provide an overview of the framework and its components. Fuzzy Cognitive Maps (FCMs) are employed to evaluate the framework and make decision analysis easier, and therefore help the decision making process about the use of AI techniques to test software. What-if analysis is used to explore and illustrate the general application of the framework.

1 Introduction

Software systems have penetrated much of the infrastructure of modern society, and the spread of software throughout society continues to proceed rapidly. Software is a key element of the systems and devices that support many of the activities that are an accepted part of our modern lifestyle. There is a consequent reliance on the correct behaviour of software, and an expectation that it will not fail. However, software is an increasingly complex product, remains labour intensive and is an error prone activity [1]. The consequences of software failure range from the trivial (such as the need to restart a computer program), through the very inconvenient (such as the malfunction of traffic signals), to the catastrophic, where life and property may be affected. To minimise software failure, and its various impacts, high quality software must be created. Testing is a major quality assurance technique to evaluate the quality of software throughout the development cycle [1, 2].

This paper proposes a framework that a test manager can consider in making decisions about the use of AI techniques for software testing. The framework offers test

G. Phillips-Wren et al. (Eds.): Advances in Intel. Decision Technologies, SIST 4, pp. 53–63.
springerlink.com