Web Intelligence (WI)

Research Challenges and Trends in the New Information Age

Y.Y. Yao[1], Ning Zhong[2], Jiming Liu,[3] and Setsuo Ohsuga[4]

[1] Department of Computer Science, University of Regina
Regina, Saskatchewan, Canada, S4S 0A2
E-mail: yyao@cs.uregina.ca
[2] Department of Information Engineering, Maebashi Institute of Technology
Japan
E-mail: zhong@maebashi-it.ac.jp
[3] Department of Computer Science, Hong Kong Baptist University
Hong Kong
E-mail: jiming@Comp.HKBU.Edu.HK
[4] Department of Information and Computer Science, Waseda University
Japan

Abstract. This paper is about a new research field called *Web Intelligence* (WI for short). We try to explain the needs for coining the term as a sub-discipline of computer science for systematic studies on advanced Web related theories and technologies, as well as the design and implementation of Intelligent Web Information Systems (IWIS). Background information and related topics are discussed in an attempt to demonstrate why we consider WI to be a subject worthy of study and, at the same time, to establish a starting point for the further development of WI.

1 Introduction

With the rapid growth of Internet and World Wide Web (WWW), we have now entered into a new information age. The Web provides a total new media for communication, which goes far beyond the traditional communication medias, such as radio, telephone and television. The Web has significant impacts on both academic research and ordinary daily life. It revolutionizes the way in which information is gathered, stored, processed, presented, shared, and used. The Web offers new opportunities and challenges for many areas, such as business, commerce, marketing, finance, publishing, education, research and development. For computer scientists, the Web introduces many new research topics and provides a new platform to reconsider old problems. It might be high time to create a new sub-discipline of computer science covering theories and technologies related to the Web. *Web Intelligence* is our proposal for this purpose.

The authors of this paper conceived *Web Intelligence* (WI for short) in late 1999. We felt that although a number of conferences and journals publish or

cover Web or Internet related topics, there was no conference and journal devoted to intelligence aspects in the design and implementation of Web information systems. We suspected that there exists a need for a conference devoted to Web Intelligence. At the 24th Annual International Computer Software and Applications Conference (IEEE COMPSAC) in 2000, we first introduced Web Intelligence and formally announced the new *Web Intelligence* conference in a position paper at a Panel on Data Mining and Web Information Systems [50]. We are impressed by the quick and vast responses, as well as kind support, from research community and reputable publishers.

The main objective of this paper is to formally initiate a sub-discipline of computer science by coining the term Web Intelligence, into which Web related research can be fitted. It is more a proposal and an appeal for the creation of WI on its own rights, rather than a precise definition of what is exactly WI. We are more concerned with the necessity and benefits of WI, as well as research topics of WI. It is our intention to create further discussion and critical examination of WI among researchers working on Web related topics.

The rest of the paper is organized as follows. In Section 2, we provide a definition of Web Intelligence. In Section 3, we argue that it is necessary and beneficiary to have a new sub-discipline of computer science labelled by WI. In Section 4, we present an overview of Artificial Intelligence and show its relevance to WI. In Section 5, we provide a list of topics of WI. In Section 6, we discuss trends and challenges of WI related research and development. Section 7 is devoted to intelligent Web Agents (WA). Finally, Section 8 introduces the Web Intelligence conference, and Section 9 gives conclusion, respectively.

2 What is Web Intelligence?

At this very early stage, we are not sure if a formal definition of Web Intelligence is useful or desirable. Nevertheless, we suggest the following definition:

> "Web Intelligence (WI) exploits Artificial Intelligence (AI) and advanced Information Technology (IT) on the Web and Internet."

This definition has the following implications. The basis of WI is AI and IT. The "I" happens to be shared by both "AI" and "IT", although with different meanings in them, and "W" defines the platform on which WI research is carried out. The goal of WI is the joint goals of AI and IT on the new platform of the Web. That is, WI applies AI and IT for the design and implementation of Intelligent Web Information Systems (IWIS). An IWIS should be able to perform functions normally associated with human intelligence, such as reasoning, learning, and self improvement.

There perhaps might not be a standard and non-controversial definition of WI, as the case that there is no standard definition of AI. One may argued that our definition of WI focuses more on the software aspects of the Web. It is not our intention to exclude any research topic using the proposed definition. The term, Web Intelligence, should be considered as an umbrella or a label of a new

branch of research centered on the Web. Our definition simply states the scopes and goals of WI. This allows us to include any theories and technologies that either fall in the scopes or aim at the same goals. To complement the formal definition, we try to make the picture clearer by listing topics to be covered by WI.

WI will be an ever-changing research branch. It will be evolving with development of the Web as new media for information gathering, storage, processing, delivery and utilization. It is our expectation that WI will be evolved into an inseparable research branch of computer science. Although no one can predict the future in detail and without uncertainty, it is clear that WI would have huge impacts on the application of computers, which in turn will effect our everyday lives.

3 Motivations and Justifications for WI

The introduction of Web Intelligence (WI) can be motivated and justified from both academic and industrial perspectives.

Two features of the Web make it a useful and unique platform for computer applications and research, the size and complexity. The Web contains a huge amount of interconnected Web documents known as Web pages. For example, the popular search engine Google claims that it can search 1,346,966,000 pages as of February 2001. The sheer size of the Web leads to difficulties in the storage, management, and efficient and effective retrieval of Web documents. The complexity of the Web, in terms of connectivity and diversity of Web documents, forces us to reconsider many existing information systems, as well as theories, methodologies and technologies underlying those systems. One has to deal with a heterogeneous collection of structured, unstructured, semi-structured, interrelated, and distributed Web documents consisting of texts, images and sounds, instead of homogeneous collection of structured and unrelated objects. The latter is the subject of study of many conventional information systems, such as databases, information retrieval, and multi-media systems. To accommodate the needs of the Web, one needs to study issues on the design and implementation of the Web-based information systems by combining and extending results from existing intelligent information systems. Existing theories and technologies need to be modified or enhanced to deal with complexity of the Web. Although individual Web-based information systems are constantly being deployed, advanced issues and techniques for developing and for benefiting from the Web remain to be systematically studied. The challenges brought by the Web to computer scientists may justify the creation of the new sub-discipline, WI, for carrying out Web-related research.

The Web increases the availability and accessibility of information to a much larger community than any other computer applications. The introduction of Personal Computers (PCs) brought the computational power to ordinary people. It is the Web that delivers more effectively information to everyone at finger tips. The Web, no doubt, offers a new means for sharing and transmitting in-

formation unmatchable by other media. The revolution started by the Web is just beginning. New business opportunities, such as e-commerce, e-banking, and e-publication, will increase with the maturity of the Web. It can hardly over-emphasize more impacts of the Web on the business and industrial world. The creation of a new sub-discipline devoted to Web related research and applications might has a significant value in the future.

The needs for WI may be further illustrated by the current fast growing research and industrial activities centered on it. We searched the Web by using the keyword "Web Intelligence" through several search engines in February 2001. The results are summarized in Table 1.

Table 1. A Statistics on WI

Search Engine	Number of hits
Lycos (http://search.lycos.com/)	1,102,279
Google (http://www.google.com/)	1,080,000
Excite (http://www.excite.com)	223,825
AltaVista (http://www.AltaVista.com/)	1,271
Netscape (http://Netscape.com/)	77
Yahoo (http://www.yahoo.com/)	74
LookSmart (http://www.looksmart.com/)	62

There are some interesting observations from the search results. The Web pages returned by most search engines contain both keywords "Web" and "Intelligence", although they may not appear as a phrase in many pages. The co-occurrences of the two keywords show their strong association. This provides a piece of convincing empirical evidence supporting WI. The identification of this association may lead to the recognition of the importance of WI. We also used advanced search option of Google to search for the exact phrase "Web Intelligence". We obtained 3,660 hits. We found that many companies concentrate on WI to provide intelligent solutions to business in the new Web-based information age. In fact, the majority of the top 40 pages returned by Google is industry related. For comparison, we search ResearchIndex (the NECI Scientific Literature Digital Library, http://citeseer.nj.nec.com/cs) containing an extreme large collection of scientific papers on-line. We found only one paper contains the phrase "Web Intelligence". A further search of "Web" and "Intelligence" within two words results in 12 documents. They deal with topics such as Web browser intelligence, artificial intelligence for Web search, and Internet marketing intelligence through Web log mining. We also used "Web Intelligence" to query Ask Jeeves (http://www.ask.com/) and obtained related topics, such as intelligent Web systems, Web artificial intelligence, Web business intelligence, intelligent Web agents, intelligent Web robots, intelligent user interfaces, and Web user interfaces. Those topics clearly fit the proposed research areas of WI. From the

search results, we also noticed that the Call for Papers of this conference (WI'01) was either archived on, or linked by many Web sites.

In summary, we can conclude that there is an interest and a need for WI. It seems that academic research needs to speed up to be in pace with the industrial demands. The introduction of WI sub-discipline may be helpful in bridging the gap between industry demands and academic research.

4 Perspectives of WI

As a new branch of research, Web Intelligence exploits Artificial Intelligence (AI) and Information Technology (IT) on the Web. On the one hand, it may be viewed as applying results from these existing disciplines to a totally new domain. On the other hand, WI may also introduce new problems and challenges to the established disciplines. WI may also be viewed as an enhancement or an extension of AI and IT. It remains to be seen if WI would become a sub-area of AI and IT or a child of a successful marriage of AI and IT. However, no matter what happens, studies on WI can benefit a great deal from the results, experience, success and lessons of AI and IT.

In their very popular textbook, Russell and Norvig [39] examined different definitions of artificial intelligence from eight other textbooks, in order to decide what is exactly AI. They observed that the definitions vary along the two dimensions. One dimension deals with the functionality and ability of an AI system, ranging from thought processes and reasoning ability of the systems to the behavior of the systems. The other dimension deals with the designing philosophy of AI systems, ranging from intimating human problem solving to making rational decision. The combination of the two dimensions results in four categories of AI systems adopted from Russell and Norvig [39]:

Systems that think like humans.	Systems that think rationally.
Systems that act like humans.	Systems that act rationally.

This classification provides a basis for the studies of various views and approaches for AI. It also clearly defines goals in the design of AI systems. According to Russell and Norvig [39], they correspond to four approaches, the cognitive modeling approach (thinking humanly), the Turing test approach (acting humanly), the the laws of thought approach (thinking rationally), and the rational agent approach (acting rationally).

The two rows for separating AI systems in terms of thinking and acting may not be a most suitable classification. Action is normally the final result of a thinking process. One may argue that the class of systems acting humanly is a super set of the class of system thinking humanly. In contrast, the separation of human-centered approach and rationality-centered approach may have a significant implications in the studies of AI. While earlier research on AI was focus more on human-centered approach, rationality-centered approach received more attention recently [39].

The first column is centered around humans and leads to the treatment of AI as an empirical science involving hypothesis and experimental confirmation. A human-centered approach represents the descriptive view of AI. Under this view, a system is designed by intimating the human problem solving. This implies that a system should have the usual human capabilities such as knowledge representation, natural language processing, reasoning, planning and learning. The performance of an AI system is measured or evaluated through the Turing test. An system is said to be intelligent if it provides human level performance. Such a descriptive view dominates the majority of earlier studies of expert systems, a special type of AI systems.

The second column represents the prescriptive or normative view of AI. It deals with theoretical principles and laws that an AI system must follow, instead of intimating humans. That is, a rationalist approach deals with an ideal concept of intelligence, which may be independent of human problem solving. An AI system is rational if it does the right thing and makes the right decision. The normative view of AI based on the well established disciplines such as mathematics, logic, and engineering.

The descriptive and normative views also reflect the experimental and theoretical aspects of AI research. The experimental study represents the descriptive view. It covers theories and models for the explanation of the workings of the human mind, and applications of AI to solving problems that normally require human intelligence. The theoretic study aims at the development of theories of rationality, and focuses on the foundations of AI. The two views are complementary to each other. Studies in one direction may provide valuable insights into the other.

Web Intelligence concerns the design and development of intelligent Web information systems. The previous framework for the study of AI can be immediately applied to that of Web Intelligence. More specifically, we can cluster research in WI into the prescriptive approach and the normative approach, and cluster Web information systems in terms of thinking and acting. Various research topics can be identified and grouped accordingly.

Like AI, a foundation of WI can be established by drawing results from the following many related disciplines:

Mathematics:
computation, logic, probability.

Applied Mathematics and Statistics:
algorithms, non-classical logics, decision theory, information theory, measurement theory, utility theory, theories of uncertainty, approximate reasoning.

Psychology:
cognitive psychology, cognitive science, human-machine interaction, user interface.

Linguistics:
computational linguistics, natural language processing, machine translation.

Information Technology:
information science, databases, information retrieval systems, knowledge dis-

covery and data mining, expert systems, knowledge-based systems, decision support systems, intelligent information agents.

The topics under each entry are only intended as examples. They do not form an exhausted list.

In the development of AI, we have witnessed the formulation of many of its new sub-branches, such as knowledge-based systems, artificial neural networks, genetic algorithms, and intelligent agents. Recently, non-classical AI topics have received much attentions under the name of computational intelligence. Computational intelligence focuses on the computational aspect of intelligent systems [7, 53]. The application of AI in other disciplines also leads to new techniques in the corresponding fields. For instance, Business Intelligence (BI) is a result of applying artificial intelligence to the business domain. Artificial Intelligence in Medicine also proved to be a successful application. When viewing WI in such settings, we can identify at least two of its roles. WI may be interpreted "Web based Artificial Intelligence" as the study of particular aspects of AI in the context of the Web, in parallel to the study of computational intelligence. WI may also be interpreted as "Artificial Intelligence on the Web" which regards it as a new application of AI.

A more practical goal of WI is the design and implementation of intelligent Web information systems (IWIS). It should be realized that an IWIS is an integrated system containing many sub-systems. To design such a system, it is necessary to apply a variety of theories and technologies.

In his work on vision, Marr [37] convincingly made the point that a full understanding of an intelligent system involves explanations at various levels. The same argument is applicable to the development of an IWIS. We can identify at least two levels, the conceptual formulation and physical implementation. The conceptual formulation deals with foundations of IWIS, while physical implementation concerns with construction of an IWIS. The former depends on mathematics and logic, and the latter depends on algorithms and programming. Each level may be further divided into more sub-levels. Research in WI should include any topics at different levels.

5 Topics Covered by WI

In order to study advanced Web technology systematically, and develop advanced Web-based intelligent information systems, we list several major subtopics in each topic below.

- Web Information System Environment and Foundations:
  - competitive dynamics of Web sites,
  - emerging Web technology,
  - network community formation and support,
  - new Web information description and query languages,
  - the semantic Web,
  - theories of small world Web,

  - Web information system development tools,
  - Web protocols.
- Web Human-Media Engineering:
  - the art of Web page design,
  - multimedia information representation,
  - multimedia information processing,
  - visualization of Web information,
  - Web-based human computer interface.
- Web Information Management:
  - data quality management,
  - information transformation,
  - Internet and Web-based data management,
  - multi-dimensional Web databases,
  - OLAP (on-line analytical processing),
  - multimedia information management,
  - new data models for the Web,
  - object oriented Web information management,
  - personalized information management,
  - semi-structured data management,
  - use and management of metadata,
  - Web knowledge management,
  - Web page automatic generation and updating,
  - Web security, integrity, privacy and trust.
- Web Information Retrieval:
  - approximate retrieval,
  - conceptual information extraction,
  - image retrieval,
  - multi-linguistic information retrieval,
  - multimedia retrieval,
  - new retrieval models,
  - ontology-based information retrieval,
  - automatic Web content cataloguing and indexing.
- Web Agents:
  - dynamics of information sources,
  - e-mail filtering,
  - e-mail semi-automatic reply,
  - global information collecting,
  - information filtering,
  - navigation guides,
  - recommender systems,
  - remembrance agents,
  - reputation mechanisms,
  - resource intermediary and coordination mechanisms,
  - Web-based cooperative problem solving.
- Web Mining and Farming:

- data mining and knowledge discovery,
- hypertext analysis and transformation,
- learning user profiles,
- multimedia data mining,
- regularities in Web surfing and Internet congestion,
- text mining,
- Web-based ontology engineering,
- Web-based reverse engineering,
- Web farming,
- Web-log mining,
- Web warehousing.

– Web-Based Applications:
  - business intelligence,
  - computational societies and markets,
  - conversational systems,
  - customer relationship management (CRM),
  - direct marketing,
  - electronic commerce and electronic business,
  - electronic library,
  - information markets,
  - price dynamics and pricing algorithms,
  - measuring and analyzing Web merchandising,
  - Web-based decision support systems,
  - Web-based distributed information systems,
  - Web-based electronic data interchange (EDI),
  - Web-based learning systems,
  - Web marketing,
  - Web publishing.

It should be pointed out that WI researches are not limited to the topics listed above. We expect that new topics will be added, and existing topic will be regrouped or redefined.

In summary, we can observe two ways in which WI research can be characterized. The first one is by adding "Web" as a prefix to an existing topic. For example, from "digital library", "information retrieval", and "agents", we can obtain "Web digital library", "Web information retrieval", and "Web agents". On the other hand, we can add "on the Web" as a postfix. For example, we can obtain "digital library on the Web", "information retrieval on the Web", and "agent on the Web". Our list of research topics is given by the prefix method. However, we must avoid mistakes of seductive semantics as discussed by Bezdek [5]. That is, "words or phrases which convey, by being interpreted in their ordinary (non-scientific) usage, a far more profound and substantial meaning about an algorithm or computational architecture than can be readily ascertained from the available theoretical and/or empirical evidence." For a healthy development of Web Intelligence, we have to be more realistic about our goals and try to avoid over-selling of the subject.

6 Trends and Challenges of WI Related Research and Development

Web Intelligence presents excellent opportunities and challenges for the research and development of new generation Web-based information processing technology, as well as for exploiting business intelligence. With the rapid growth of the Web, research and development on WI have received much attention. We expect that more attention will be focused on WI in the coming years. Many specific applications and systems have been proposed and studied. Several dominant trends can be observed and are briefly reviewed in this section.

E-commerce is one of the most important applications of WI. The e-commerce activity that involves the end user is undergoing a significant revolution [42]. The ability to track users' browsing behavior down to individual mouse clicks has brought the vendor and end customer closer than ever before. It is now possible for a vendor to personalize his product message for individual customers at a massive scale. This is called *targeted marketing* or direct marketing [25]. Web mining and Web usage analysis play an important role in e-commerce for customer relationship management (CRM) and targeted marketing. Web mining is the use of data mining techniques to automatically discover and extract information from Web documents and services [23, 42, 48]. Zhong *et al.* proposed a way of mining peculiar data and peculiarity rules that can be used for Web-log mining [52]. They also proposed ways for targeted marketing by mining classification rules and market value functions [44, 49]. A challenge is to explore the connection between Web mining and the related agent paradigm such as Web farming that is the systematic refining of information resources on the Web for business intelligence [14].

Text analysis, retrieval, and Web based digital library is another fruitful research area in WI. Topics in this area include semantics model of the Web, text ming, automatic construction of citation. Abiteboul *et al.* systematically investigated the data on the Web and the features of semistructured data [1]. Zhong *et al.* studied text mining on the Web including automatic construction of ontology, e-mail filtering system, and Web-based e-business systems [47, 51].

Web based intelligent agents are aimed at improving a Web site or providing help to a user. Liu *et al.* worked on e-commerce agents [29]. Liu and Zhong worked on Web agents and KDDA (Knowledge Discovery and Data Mining Agents) [31, 32]. We believe that Web agents will be a very important issue. It is therefore not surprising that we decide to hold the WI conference in parallel to the Intelligent Agents conference. In the next section, we provide a more detailed description of intelligent Web agents.

The Web itself has been studied from two aspects, the structure of the Web as a *graph* and the *semantics* of the Web. Studies on Web structures investigate several structural properties of graphs arising from the Web, including the graph of hyperlinks, and the graph induced by connections between distributed search servants. The study of the Web as a graph is not only fascinating in its own right, but also yields valuable insight into Web algorithms for crawling,

searching and community discovery, and the sociological phenomena which characterize its evolution [6]. Studies of the *semantics* of the Web were initiated by Tim Berners-Lee, the creator of the World Wide Web [4]. The Web is referred to as the "semantic Web", where information will be machine-processible in ways that support intelligent network services such as information brokers and search agents [11, 12]. The semantic Web requires interoperability standards that address not only the syntactic form of documents but also the semantic content. A semantic Web also lets agents utilize all the data on all Web pages, allowing it to gain knowledge from one site and apply it to logical mappings on other sites for ontology-based Web retrieval and e-business intelligence. Ontologies and agent technology can play a crucial role in enabling such Web-based knowledge processing, sharing, and reuse between applications. A new DARPA program called DAML (DARPA Agent Markup Languages) is a step toward a "semantic Web" where agents, search engines and other programs can read DAML mark-up to decipher meaning rather than just the content on a Web site [16].

7 Intelligent Web Agents

Intelligent agents are computational entities that are capable of making decisions on behalf of their users and self-improving their performance in dynamically changing and unpredictable task environments [26, 31–34]. In [27], Liu provided a comprehensive overview of related research work in the field of autonomous agents and multi-agent systems, with an emphasis on its theoretical and computational foundations as well as in-depth discussions on the useful techniques for developing various embodiments of agent-based systems, such as autonomous robots, collective vision and motion, autonomous animation, and search and segmentation agents. The core of those techniques is the notion of synthetic or emergent autonomy based on behavioral self-organization.

Intelligent Web Agents (WA) are software programs that primarily serve two important roles: a). autonomous entities for exploring and exploiting Web-based services, and b). prototype entities for exhibiting and explaining Web-generated regularities. These two roles are summarized below.

7.1 From WA to Web-Based Services

The first role for WA can be readily described and appreciated by examining the following typical scenarios in which various tasks and objectives are achieved [9, 28, 29, 45, 46].

1. **Personalized Multimodal Interface** WA can provide users with a user-friendly style of presentation that personalizes both the interaction with users and the content presentation. This activity involves the creation of various cognitive aids, including tables, charts, executive summaries, indices, and personalized visual assistants (*e.g.,* graphically animated personas and virtual-reality avatars). WA as interfaces must offer the ease of using electronic services. The provided cognitive aids must be concise (*i.e.,* accessible

with as fewer manipulations as possible and as less memorization as possible) and consistent (*i.e.,* understandable based on users' previously customized cognitive styles).

2. **Push and Pull** WA can play an important role in dynamically creating pull-and-push advertising. Here, by pull-and-push advertising we mean that a user expresses his or her favorites during the interaction with the agents (pull advertising) and in return the agents search and deliver the information about the favorite items dynamically to the user (push advertising). Such agents can also increase the positive externality of products, that is, the better people are informed about certain products, the more likely the products will be sold.
3. **Pattern Discovery and Self-Organization** WA will enable to detect what users' buying patterns are forming and how they are structured, and hence effectively manage the online commerce. Collaborative recommendation agents can help individual users aggregate into groups, which can in turn form a dynamical marketplace (for example, see [13]).
4. **Information Gateway** WA can provide users with immediate access to the most relevant information. This support encompasses a wide spectrum of information filtering and delivery activities by manipulating various heterogeneous Web sources including databases, data warehouses, newswire, financial reports, newsletters, newsgroups, outbound emails, electronic bulletin boards, and hypermedia documents, and based on users' profiles, tailoring and delivering the retrieved information to the users. The provided summary information must be just-in-time (*i.e.,* delivered whenever is needed), relevant (*i.e.,* focused on whichever topics the users are concerned with), and up-to-minute (*i.e.,* refreshed whenever a new piece of information arrives). An example of applications with this type of agent support is comparison shopping that utilizes WA with mobile and filtering capabilities. Some related experiences have been reported in [24, 35, 41].
5. **Reward** WA can motivate users to enter and re-enter a certain electronic service. While an ever-greater proliferation of content continues to consume individuals' attention, *e.g.,* through push technology to sell something or to support users, WA can play a crucial role in creating a captive audience, in educating it constantly, and even in removing away users' old purchase habits. To be rewarding is to add value. The motivational rewards or incentives can be created by offering free access to certain information and utility resources (*e.g.,* free software download), opportunities to participate in multi-user information/commodity exchange activities (*e.g.,* collaborative recommendation, chat, bidding, and auction), and scheduled plans for promotional deals.
6. **Matchmaking** WA can serve as a new means for trading commodities. Since the interests of users as well as the availability of products from dealers can change dynamically from time to time, what usually happens in present day electronic commerce is: (1) a dealer sells his or her items simply because these are the *only* items that he or she has at the moment, or (2) a user buys a certain item simply because it is the *last* item that he or she can find

that partially fits his or her need. WA-based customized business attempts to change the existing online buying and selling into the following new scenarios: (1) a dealer identifies and offers what exactly users are interested in, and (2) a user finds and purchases what he or she really loves – some technical issues related to matchmaking have been addressed in [8, 22, 36].

7. **Decision** WA can assist Web users in making decisions. Such decision support may be in the forms of evaluations or recommendations on the various features of certain specific items, cost-benefit analysis, inference support for optimizing utility and resources with respect to functional, time, and cost requirements, and model-based trend analysis and projections concerning new patterns of demand (for example, see [19, 40]).
8. **Delegation** WA can act on behalf of Web users in online activities. The tasks that WA may delegate to achieve include matchmaking, server monitoring, negotiation, bidding, auction, transaction, transfer of goods, and follow-up support. This scenario will empower a new paradigm shift from user-centric to user-delegated electronic business. The delegations of these tasks may be carried out in either semi-autonomous (with users' intervention on decisions) or fully autonomous manners. To this end, various computational theories and models have been proposed and reported in [15, 18, 38, 43]).
9. **Collaborative Work Support** WA can offer the infrastructure support as well as the necessary function for collaboratively solving problems and managing workflow activities (for related examples, see [10, 19, 43]).

7.2 From WA to Web-Generated Regularities

The World Wide Web has evolved into a dynamic, distributed, heterogeneous, complex network, which is hard to control [2, 20]. To many people, whether Web developers or researchers who are concerned about the dynamics of complex systems [17] such as Internet, human community, and ecology, it has become imperative to truly understand and interpret (in addition to merely observe) the strong regularities emerged from the 'messy' universe of the World Wide Web. Up till now, there have been few efforts on describing different aspects of the orders in the World Wide Web [20, 21, 3]. However, as an entire system, the origin and interrelated elements of the regularities still remain unknown.

Liu and Zhang [30] have designed and validated an agent-based model that takes into account Web topology, information distribution, and user interest profile to simulate user surfing behavior and explore the origin of regularities on the World Wide Web surfing. In their experiments, they have discovered that it is the unique distribution of user interest that leads to the regularities in user surfing behavior, i.e., a power law distribution of user surfing depth. The Web topology can only influence the shape parameters of the distribution without changing the nature of the distribution. Also discovered is that the power law of link click frequency is largely due to user purposeful surfing behavior. Their work shows that the regularities in the Web are interrelated and not artifacts of a particular surfing process.

Also in their studies, they have studied three categories of users, according to their interest and familiarity with the Web: Random users who have no obvious intention in Web surfing, rational users who have certain goals to achieve but are not familiar with the Web structure, recurrent users who have certain specific intents and are very familiar with the Web structure. The ability to predict the content at the next-level nodes becomes stronger when moving from random to recurrent users. The result of simulations with respect to the three user categories unveiled that the regularities of user surfing depth on pages and domains still remain the same, while a power law of link click frequency distribution will disappear as we move from recurrent users to random users. This result shows that the order existing in link click frequency comes from user's content-prediction ability, that is whether or not a user can determine his/her next step according to his/her own interest and names of the hyperlinks.

8 Web Intelligence Conference

In order to meet the challenges of WI in the new information age, a new high-quality, high-impact international conference series, namely the Asia-Pacific Conference on Web Intelligence (WI) is initiated. WI-2001 is the first meeting in this new series (http://kis.maebashi-it.ac.jp/wi01). It is an international forum for researchers and practitioners to present the state-of-the-art in the development of Web intelligence, to examine performance characteristics of various approaches in Web-based intelligent information technology, and to cross-fertilize ideas on the development of Web-based intelligent information systems among different domains. By idea-sharing and discussions on the underlying foundations and the enabling technologies of Web intelligence, we hope to stimulate future development of new models, new methodologies, and new tools for building a variety of embodiments of Intelligent Web Information Systems. By jointly holding WI conference and Intelligent Agents conference, we expect a close interaction between the two groups.

The title "Web Intelligence" of the conference was chosen to reflect the distinct feature that the conference is focused on intelligence aspects of Web and Web information systems. The name is short enough to catch attention to this important subfield. It is also general enough to attract contributions from all Web related research.

9 Conclusion

While it may be difficult to define what is exactly Web Intelligence (WI), one can easily argue for the need and necessity of creating such a subfield of study in computer science. With the rapid growth of the Web, we foresee a fast growing interest in Web Intelligence.

Roughly speaking, we define Web Intelligence as a field that "exploits Artificial Intelligence (AI) and advanced Information Technology (IT) on the Web

and Internet." It may be viewed as a marriage of artificial intelligence and information technology in the new setting of the Web. By examining the scope and historical development of artificial intelligence, we discuss some fundamental issues of Web Intelligence in a similar manner. There is no doubt in our mind that results from AI and IT will influence the development of WI.

Instead of searching for a precise and non-controversial definition of WI, we list topics that might be interested by a researcher working on Web related issues. In particular, we identify some challenging issues of WI, including e-commerce, studies of Web structures and Web semantics, Web information storage and retrieval, Web mining, and intelligent Web agents.

We advocate for a new conference devoted to WI, namely, the Asia-Pacific Conference on Web Intelligence. The conference will be an international forum for researchers and practitioners to present the state-of-the-art in the development of Web intelligence, to examine performance characteristics of various approaches in Web-based intelligent information technology, and to cross-fertilize ideas on the development of Web-based intelligent information systems among different domains.

References

1. Abiteboul, S., Buneman, P., and Suciu, D. *Data on the Web*, Morgan Kaufmann, 2000.
2. Albert, R., Jeong, H. and Barabasi, A.-L., Diameter of the World-Wide Web, *Nature*, **410**, 130-131, 1999.
3. Barabasi, A.-L. and Albert, R., Emergence of scaling in random networks, *Science*, **286**, 509-512, 1999.
4. Berners-Lee, T., Hendler, J., and Lassila, O. The semantic Web, *Scientific American*, 29-37, May 2001.
5. Bezdek, J.C. What is computational intelligence? in: *Computational Intelligence: Imitating Life*, Zurada, J.M., Marks II, R.J. and Robinson, C.J. (Eds.), IEEE Press, 1-12, 1994.
6. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A. and Wiener, J.L. Graph structure in the web, *Computer Networks*, **33**, 309-320, 2000.
7. Cercone, N. and McCalla, G. (eds.) *Computational Intelligence*, **1**, iii-vi, 1985.
8. Chavez, A. and Maes, P., Kasbah: an agent marketplace for buying and selling goods, *Proceedings of the First International Conference on the Practical Application of Intelligent Agents and Multi-Agent Technology (PAAM'96)*, 1996.
9. Cheung, K. W., Li, C. H., Lam, E. C. M., and Liu, J., Customized electronic commerce with intelligent software agents, in: *Internet Commerce and Software Agents - Cases, Technologies and Opportunities*, Rahman, S.M. and Bignall, R.J. (Eds.), IDEA Group Publishing, 150-176, 2001.
10. Clearwater, S. (ed.) *Market-Based Control: A Paradigm for Distributed Resource Allocation*, World Scientific, 1995.
11. Decker, S., Melnik, S. et al. The semantic web: the roles of XML and RDF, *IEEE Internet Computing*, **4:5**, 63-74, 2000.
12. Decker, S., Mitra, P., and Melnik, S. Framework for the semantic web: an RDF tutorial, *IEEE Internet Computing*, **4:6**, 68-73, 2000.

13. Guttman, R., Moukas, A., and Maes, P., Agent-mediated electronic commerce and consumer buying behavior, *Knowledge Engineering Review Journal*, **13**, June, 1998.
14. Hackathorn, R.D. *Web Farming for the Data Warehouse*, Morgan Kaufmann, 2000.
15. Hausch, D., Multi-object auctions: sequential vs. simultaneous sales, *Management Science*, **32**, 1599-1610, 1986.
16. Hendler, J.A. Agents and the semantic Web, *IEEE Intelligent Systems*, **16:2**, 30-37, 2001.
17. Helbing, D., Huberman, B. A., and Maurer, S. M, Optimizing traffic in virtual and real space, in: *Traffic and Granular Flow '99: Social, Traffic, and Granular Dynamics*, Helbing, D., Herrmann, H. J., Schreckenberg, M., and Wolf, D. E. (Eds.), Springer-Verlag, 2000.
18. Hon-Snir, S., Monderer, D. and Sela. A., A learning approach to auctions, *Journal of Economic Theory*, **8**, 65-88, 1998.
19. Hu, J. and Wellman, M., Online learning about other agents in a dynamic multiagent system, *Proceedings of the Second International Conference on Autonomous Agents*, 1998.
20. Huberman, B. A. and Adamic, L. A., Growth dynamics of the World-Wide Web, *Nature*, **410**, 131, 1999.
21. Huberman, B. A., Pirolli, P. L. T., Pitkow, J. E., and Lukose, R. M., Strong regularities in World Wide Web surfing, *Science*, **280**, 96-97, 1997.
22. Jennings, N., Faratin, P., Johnson, M., Norman, T., O'Brien, P., and Wiegand, M., Agent-based business process management, *International Journal of Cooperative Information Systems*, **5**, 105-130, 1996.
23. Kosala, R. and Blockeel, H. Web mining research: a survey, *ACM SIGKDD Explorations Newsletter*, **2**, 1-15, 2000.
24. Kushmerick, N., Weld, D., Doorenbos, R., Wrapper induction for information extraction, *Proceedings of the 16th International Joint Conference on Artificial Intelligence (IJCAI'97)*, 1997.
25. Ling, C.X. and Li, C. Data mining for direct marketing: problems and solutions, *Proceedings of KDD'98*, 73-79, 1998.
26. Liu, J. Self-organized intelligence, in: *Agent Engineering*, Liu, J., Zhong, N., Tang, Y. Y., and Wang, P. S. P. (Eds.), World Scientific, 2001.
27. Liu, J., *Autonomous Agents and Multiagent Systems*, World Scientific, 2001.
28. Liu, J. and Ye, Y. E-commerce agents: marketplace solutions, security issues, and supply and demand, in: *E-Commerce Agents: Marketplace Solutions, Security Issues, and Supply and Demand*, Liu, J. and Ye, Y. (Eds.), Springer-Verlag, 2001.
29. Liu, J. and Ye, Y. (Eds.) *Advances in E-commerce Agents: Brokerage, Negotiation, Security, and Mobility*, Springer-Verlag, 2001.
30. Liu, J. and Zhang, S. W., Unveiling the origins of Internet use patterns, *Proceedings of INET 2001, The Internet Global Summit*, 2001.
31. Liu, J. and Zhong, N. (Eds.) *Intelligent Agent Technology: Systems, Methodologies, and Tools*, World Scientific, 1999.
32. Liu, J., Zhong, N., Tang, Y.Y., and Wang, P.S.P. (Eds.) *Agent Engineering*, World Scientific, 2001.
33. Liu, J., Zhong, N., Tang, Y. Y., and Wang, P. S. P., *Special Issue on Agent Technology, International Journal of Pattern Recognition and Artificial Intelligence*, World Scientific, 2001.
34. Liu, J., Zhong, N., Tang, Y.Y., and Wang, P.S.P., Introduction to agent engineering, in: *Agent Engineering*, Liu, J., Zhong, N., Tang, Y.Y., and Wang, P.S.P. (Eds.), World Scientific, 2000.

35. Maes, P., Agents that reduce work and information overload, *Communications of the ACM*, **37**, 31-40, 1994.
36. Maes, P., Guttman, R., Moukas, A., Agents that buy and sell, *Communications of the ACM*, 1999.
37. Marr, D. *Vision*, Freeman, 1982.
38. Rosenschein, J. and Zlotkin, G., *Rules of Encounter: Designing Conventions for Automated Negotiation among Computers*, MIT Press, 1994.
39. Russell, S. and Norvig, P. *Artificial Intelligence, A Modern Approach*, Prentice Hall, 1995.
40. Sandholm, T., An implementation of the contract net protocol based on marginal cost calculations, *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI'93)*, 1993.
41. Shardanand, U. and Maes, P., Social information filtering: algorithms for automating 'Word of Mouth', *Proceedings of the CHI-95 Conference*, ACM Press, 1995.
42. Srivastava, J. Cooley, R., Deshpande, M. and Tan, P. Web usage mining: discovery and applications of usage patterns from web data, *SIGKDD Explorations, Newsletter of SIGKDD*, **1**, 12-23, 2000.
43. Tesauro, G. J. and Kephart, J. O., Foresight-based pricing algorithms in an economy of software agents, *Proceedings of the International Conference on Information and Computation Economies*, 1998.
44. Yao, Y.Y. and Zhong, N. Mining market value functions for targeted marketing, *Proceedings of the 25th IEEE Computer Society International Computer Software and Applications Conference (COMPSAC 2001)*, 2001.
45. Ye, Y., Liu, J., and Moukas, A., *Special Issue on Intelligent Agents in E-Commerce, Electronic Commerce Research Journal*, Baltzer Science Publishers, The Netherlands, 2001.
46. Ye, Y., Liu, J., and Moukas, A., Agents in electronic commerce, in: *Special Issue on Intelligent Agents in Electronic Commerce, Electronic Commerce Research Journal*, 2001.
47. Zhong, N. *A Study on E-mail Filtering by Uncertainty Sampling and Relation Learning*, Technical Report, Yamaguchi University, 2000.
48. Zhong, N. Knowledge discovery and data mining, *The Encyclopedia of Microcomputers*, **27**, Supplement 6, 235-285, Marcel Dekker, 2001.
49. Zhong, N., Dong, J.Z., and Ohsuga, S. Rule discovery by soft induction techniques, *Neurocomputing, An International Journal*, **36:1-4**, 171-204, Elsevier, 2000.
50. Zhong, N., Liu, J., Yao, Y.Y. and Ohsuga, S. Web Intelligence (WI), *Proceedings of the 24th IEEE Computer Society International Computer Software and Applications Conference (COMPSAC 2000)*, 469-470, 2000.
51. Zhong, N., Yao, Y.Y., and Kakemoto, Y. Automatic construction of ontology from text databases, N. Ebecken and C.A. Brebbia (Eds.) *Data Mining*, **2**, WIT Press, 173-180, 2000.
52. Zhong, N., Yao, Y.Y., and Ohsuga, S. Peculiarity oriented multi-database mining, J. Zytkow and Jan Rauch (eds.) *Principles of Data Mining and Knowledge Discovery*, LNAI 1704, Springer-Verlag, 136-146, 1999.
53. Zurada, J.M., Marks II, R.J. and Robinson, C.J. (Eds.) *Computational Intelligence: Imitating Life*, IEEE Press, 1994.