Book review

H. Zhuge, *The Knowledge Grid*, World Scientific Publishing Co., Singapore, 2004

The exponential growth of the data on the Web has not only enabled hundreds of millions of users to access and share information at anytime anywhere but also resulted in a vast amount of data at our fingertips. Transformation of information into knowledge is essential for human to enhance our ability to live and work effectively. Unfortunately the lack of structure and machine-readable semantics of today's online information has made the transformation of industrial, social, and scientific society from information-based to knowledge-based a hard problem. The knowledge Grid presents a new frontier in the Internet and grid computing technology by pushing the traditional computational grid and the present data and information grid beyond the current abstraction of cyberspace (Web and Internet). A computational grid provides a supercomputing platform by sharing of distributed resources for large computational tasks. A data grid provides an infrastructure to allow homogeneous access to heterogeneous and geographically distributed information sources. The Knowledge Grid, however, provides the next generation interconnection environment that enables people and machines to effectively capture, publish, share and manage knowledge resources. The goal of the Knowledge Grid is to unleash the full technological and supercomputing capabilities while providing more effective and efficient management and utilization of heterogeneous computing resources on the Internet. Its potential benefits for the development of human civilization may well be beyond imagination. This book, by professor Hai Zhuge, the founder of the China Knowledge Grid research group, is the first monograph in this area, uncovering most recent development of the Knowledge Grid, from the methodology, theories, models, to applications and future trends on the Knowledge Grid.

1. What is this book about?

Although this book has featured several threads of the recent research results produced by professor Zhuge and his Knowledge Grid research group, this is not just a book documenting on the author's knowledge grid research. Rather, the author uses his own research experience and result as a support for describing concepts and techniques of the Knowledge Grid. The focus is on the methodology and fundamental components of the Knowledge Grid. In the following, we provide an overview and an analysis of each chapter of the book. A global analysis is given at the end of the review.

Chapter 1 is an introduction to the Knowledge Grid. The Knowledge Grid is envisioned as a virtual, social, adaptive and semantic interconnection environment where people can capture, publish, understand, share and manage knowledge resources through their virtual roles, and knowledge resources are given well-defined meaning to better enable people and machines to work in cooperation. A Knowledge Grid typically incorporates epistemology and ontology to reflect human cognitive characteristics, and exploits social, ecological and economic principles and strategies, such as integrity, uniformity, self-organization, optimization, and openness to provide appropriate on-demand services to support cooperative teamwork, problem solving, and decision making. This chapter ends with a discussion on the technological challenges and opportunities of the Knowledge Grid as a promising solution towards the future interconnection environment.

In this first chapter, the objective of the author is clear: to explain in an intuitive way the concept of the Knowledge Grid and elaborate on various aspects of the Knowledge Grid methodology. The following chapters give a detailed description of how to build a Knowledge Grid, including methods, models, strategies and applications.

Chapter 2 introduces the Semantic Link Network (SLN) that is towards enabling people and machine to corporate effectively in accomplishing computing tasks. As a natural extension of the hyperlink network, the SLN endows both nodes and links with certain semantic meanings to share, manage, publish and search heterogeneous knowledge resources in a machine-understandable manner. To facilitate operations on an SLN, this chapter provides a series of normal forms, criteria, constraints for SLN semantic completeness and consistency. It also provides applications to illustrate the concepts and techniques of the SLN and why it is an elegant solution towards a semantically rich Knowledge Grid.

The subject is presented in a very methodic way; concepts are defined and then applied to the

proposed—SLN. Design choices are strongly motivated, illustrated by examples, and various aspects of the SLN are uncovered and evaluated. This allows the reader to get a holistic view of the SLN and how it plays a role in the Knowledge Grid construction.

The next two chapters take a further step towards how to build the Knowledge Grid. *Chapter 3* describes the Resource Space Model (RSM), a semantic data model that can effectively and efficiently organize heterogeneous resources into a uniform and normal form. This chapter presents a systematic description of the RSM framework for constructing the Knowledge Grid and how versatile resources with rich semantics can be well organized by norms and rules. The clear definition and strict normal forms guarantee the correctness and consistency of operations on RSM. In addition, the comparison and extension of the proposed RSM are nicely presented through formalization, normal forms, and topological properties. To provide users with a single semantic entry point to access, share and manage heterogeneous resource in the RSM and SLN, *Chapter 4* proposes the SSeI (Single Semantic Image) mechanism that integrates several techniques such as the SSeI approach, the SSeI query language, SSeIQL, the semantic browser and the semantic view.

Aiming at constructing a semantically rich layer for the Knowledge Grid, the author proposes a P2P Semantic Link Network to combine RSM with SLN, which takes the advantages of both classification and reasoning methods for better management and utilization of heterogeneous resources. In a P2PSLN, the SSeI plays an important role in that it can integrate heterogeneous data from different peers to provide users with a single data usage mode.

In *Chapter 5*, the knowledge flow is introduced as a unique characteristic of the Knowledge Grid. The author gives a definition of the knowledge flow and describes a number of important and interesting issues on the knowledge flow, such as the peer-to-peer knowledge sharing, the knowledge intensity, the knowledge spiral model and the principles and formation of a knowledge flow network.

This chapter offers a fresh view of the knowledge flow concept and its important characteristics. In particular, the concept of knowledge intensity and peer-to-peer knowledge sharing are interesting and novel. More discussions and examples for illustrating how to model and utilize knowledge flow, and how to relate knowledge flow with workflow can be found in author's previous publications listed in the references.

Chapter 6 is the last chapter and it describes the random graph theory and the small world theory and reviews the topology of several popular networks used today, such as the Internet, the World Wide Web, and the citation network for scientific papers. By extending the current network model, this chapter also presents an abstract live scale-free network as a possible high-level model for the Knowledge Grid development.

A nice chapter! It reviews the recent advances in the field of complex networks, focusing on the basic models and theories that are important for building scale-free knowledge networks. The modeling measures of the proposed live network are clearly motivated and compared to possible alternatives. This allows the reader to compare his/her own ideas with the enumerated alternatives, and stimulate the reader by making him/her feel like participating in the discussion.

2. Reviewers' comments

The Knowledge Grid is a promising new frontier that extends the computational grid and the information grid research and development to the knowledge world of human-centered computing and the future interconnection environment. This book is the first monograph on the Knowledge Grid and it presents the methodology, theory, models and applications of the Knowledge Grid systematically with three distinct features. First, this book synthesizes a broad knowledge on nature and social aspects of a knowledge network, and addresses the challenges of constructing and utilizing the Knowledge Grid from both computing perspective and cross-disciplines such as sociology, ecology and economics. Second, this book describes what the Knowledge Grid is and how to build it through a methodical approach and a series of reasoning and computing rules. The author envisions the Knowledge Grid as a promising solution for the future human and computer interconnection environment where effective knowledge discovery and sharing dominate the interactions between human and machines. Last but not the least, the theory and practice of the Knowledge Grid are closely combined in this book. On one hand, the book presents many interesting and innovative ideas and provides a fresh review of the Knowledge Grid research and development. On the other hand, this book also describes several applications that have successfully used the Knowledge Grid approach.

This book will be most useful for three categories of readers: the researchers and students in such fields as computer and information systems, interested in working on Semantic Web, Web Intelligence, Knowledge Networks, and Service-Oriented Computing; the Web developers and Web administrator supporting or developing systems and applications for future computing environments; and the IT professionals, engineers at companies providing Web services. This book might not answer all questions readers might have about the Knowledge Grid. For example, the author did not quantify the performance impact of adding the Knowledge Grid layer on top of current computing infrastruc-

ture. The book did not discuss how the vast amount of data mining and knowledge discovery research today be incorporated into the envisioned Knowledge Grid.

3. Conclusion

The book is well organized and easy to read. Written by one of the few researchers who are the pioneers in the Knowledge Grid field, the information contained in the book is comprehensive and informative. It is particularly beneficial that the book discusses many aspects and numerous alternative methods for the development of the Knowledge Grid, leading to a deeper understanding of both concepts and applications. This is an excellent book on the Knowledge Grid for students, knowledge and Web engineers and researchers.

Ling Liu is a professor at the College of Computing at Georgia Tech and a director of the Distributed Data Intensive Systems (DiSL) laboratory, working on various research issues and technical challenges of large scale distributed data intensive systems, ranging from decentralized overlay networks, exemplified by peer-to-peer computing and grid computing, to mobile computing systems and location based services, sensor network systems, and enterprise computing technology. She has published over 150 international journal and conference articles. She and her students have produced a number of open source software systems, among which the most popular ones are WebCQ and XWRAPElite. She is a recipient of best paper award from IEEE ICDCS 2003 for a joint work with her student on PeerCQ, and a best paper award from International Conference on World Wide Web (WWW 2004) for their work on cachingdynamic Web content, a recipient of IBM Faculty Partnership Award (2003), and a recipient of ACM service award (2001, 2004) and IEEE service award (2004). Dr. Ling Liu is on the editorial board of several top international journals, including IEEE Transactions on Knowledge and Engineering (TKDE), International Journal of Very Large Databases (VLDBJ), International Journal of Web Services Research (JWSR), and International Journals of Grid and Utility Computing. She has chaired multiple international conferences on various themes of data intensive distributed computing systems, including IEEE International Conference on Data Engineering (General chair of ICDE 2007, PC co-chair of ICDE 2006), vice PC Chair of ICDE04), a PC chair of IEEE International Conference on Collaborative Computing (2005), IEEE International Conference on Web Services (ICWS 2004).

Xue Chen is a research fellow at Institute of Computing Technology, Chinese Academy of Sciences. Her research interest is the semantic peer-to-peer overlay and scale-free network. Her research results appeared in leading international journals like IEEE Transactions on Knowledge and Data Engineering.

Ling Liu
College of Computing
Georgia Institute of Technology
USA
E-mail address: lingliu@cc.gatech.edu

Xue Chen
Institute of Computing Technology
Chinese Academy of Sciences
China
Tel.: +8601062562703; fax: +861062562703
E-mail address: chenxue@kg.ict.ac.cn

Available online 19 August 2005