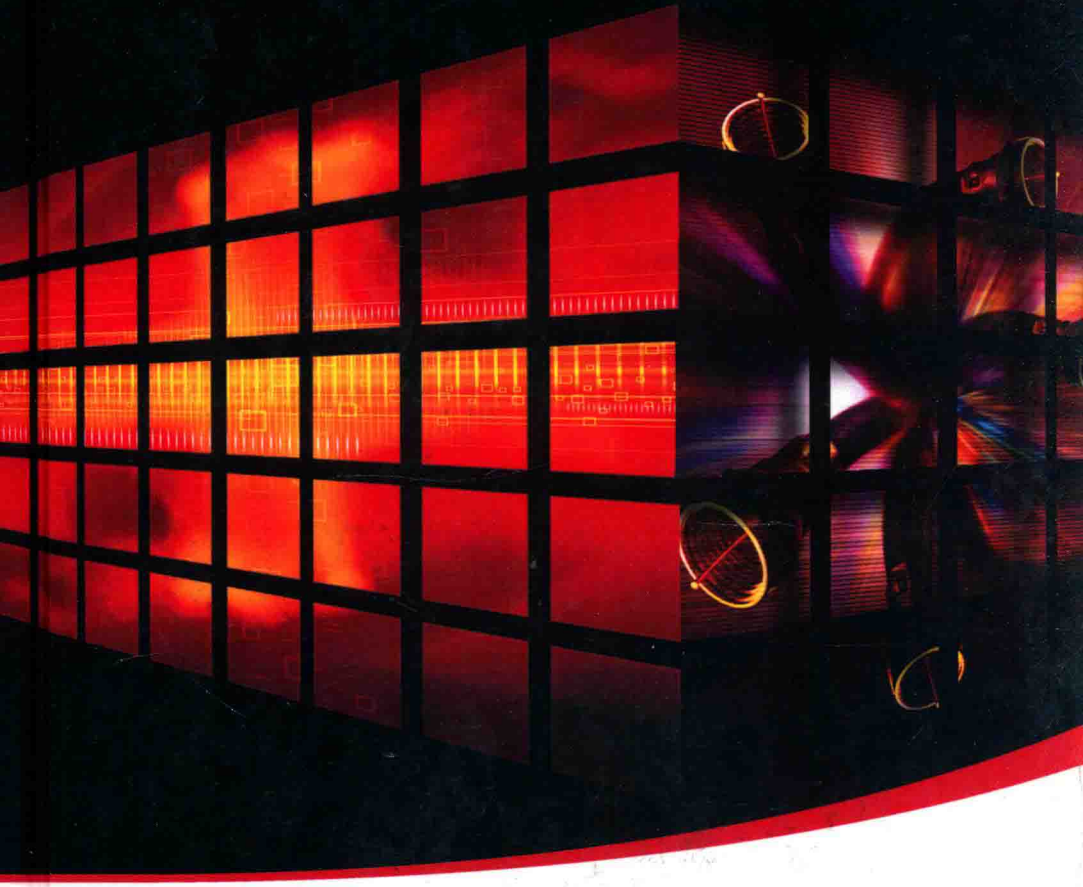
The Knowledge Grid

Second Edition

Toward Cyber-Physical Society

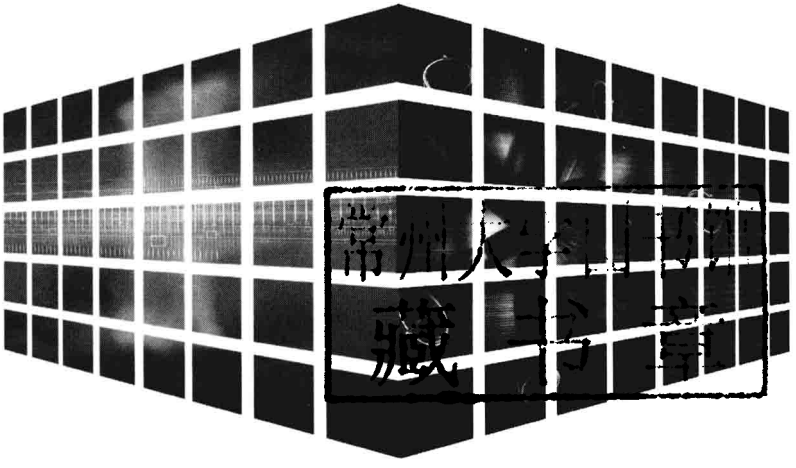
Hai Zhuge



World Scientific

THE KNOWLEDGE GRID Second Edition

Toward Cyber-Physical Society



Hai Zhuge
Nanjing University of Posts and Telecommunications, China
Chinese Academy of Sciences, China

World Scientific
NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

Published by

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

USA office: 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

UK office: 57 Shelton Street, Covent Garden, London WC2H 9HE

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library.

THE KNOWLEDGE GRID
Toward Cyber-Physical Society
(2nd Edition)

ISBN 978-981-4291-77-4

Printed in Singapore by World Scientific Printers.

THE KNOWLEDGE GRID Second Edition

Toward Cyber-Physical Society

Foreword

The Knowledge Grid — Toward Cyber-Physical Society advances the vision of human-machine-nature symbiosis. Covering methodology as well as technologies, this innovative book aims to spur innovation.

First, the book puts forward a complex semantic space model that targets the effective management of diverse resources. This is enabled by an integration of a semantic link network model and a resource space model into a unified semantic framework that supports the fundamental concepts of generalization and specialization through multi-dimensional classifications, and also supports important aspects such as semantic self-organization, community discovery, and complex reasoning. The resulting framework offers a semantic foundation for the management of resources. Second, it advances a self-organizing, semantic environment for the sharing and management of knowledge that features a knowledge flow model, social networking methods and principles, and semantic peer-to-peer networking mechanisms.

While the first edition of this book introduced the methodology, models, and technologies, the present edition extends these within the context of the cyber-physical society, which concerns not only cyber-space, but also the physical and social spaces. The book reports on a range of disciplinary innovations that relate in different ways to semantics, knowledge, and intelligence. As a result, the book can inspire novel research in different fields, which may in turn result in new technologies that benefit humanity. I recommend the book to anybody with a deep interest in human-machine-nature symbiosis and its advancement.

Christian S. Jensen
ACM Fellow, IEEE Fellow, Aarhus University, Denmark

Preface

Exploring the source and essence of knowledge and intelligence, promoting knowledge generation and sharing, facilitating knowledge management, and extending human intelligence are the grand scientific challenges.

The Internet connects computers all around the world to support data transmission. The Web makes informative pages conveniently available to Internet users everywhere. The initial aim of the Knowledge Grid is to facilitate knowledge sharing and evolution by meaningfully linking and efficiently managing globally distributed resources.

The Knowledge Grid is an optimized human-machine environment, and will be a large-scale man-machine-nature symbiosis environment, where people, society, artifacts, minds, and nature can productively coexist and harmoniously evolve. It stands for the ideal of a live, autonomous, humanized, efficient, systematic, optimal, harmonious, and sustainable social environment.

The Knowledge Grid bases its methodology on multi-disciplinary thinking because any single disciplinary method cannot solve the issues in this complex environment. Recognizing the essence, source and principle of knowledge and mind is essential to implementing the Knowledge Grid.

The 1st edition of this book published in 2004 actually founded a specific research area. The 2nd edition completes its theory, model and method by significantly enhancing previous contents and increasing four chapters (Chapter 7, 8, 9, and 10). Applications and philosophical discussion are added to help render the ideas.

This new edition puts the Knowledge Grid research into the Cyber-Physical Society environment consisting of cyber space, physical space, socio space and mental space. The ideal goes beyond the scope of the Web, Grid computing, cloud computing, Internet/Web of Things, cyber-physical systems, social network, smart Grid, and machine intelligence.

The research work was supported by many foundations and organizations, especially the Key Discipline Fund of National 211 Project (Southwest University: NSKD11013), the Natural Science Foundation of Chongqing (cstc2012jjB40012), the Chongqing Municipal Government, the National Science Foundation of China (61075074), National Basic Research Program of China, the Nanjing University of Posts and Telecommunications, and scholar programs of Ministry of Education of China and Chinese Academy of Sciences.

I would like to thank joint and visiting professorships from the Southwest University in China, the University of New Brunswick in Canada, the University of Hong Kong, the University of Queensland in Australia, and the Kyoto University in Japan.

I sincerely thank the members of the Knowledge Grid Research Group at the Key Lab of Intelligent Information Processing, Institute of Computing Technology in Chinese Academy of Sciences for their cooperative work, especially my former students Xiaoping Sun, Xue Chen, Xiang Li, Liang Feng, Junsheng Zhang, Yunchuan Sun, Yunpeng Xing, Ruixiang Jia, Weiyu Guo, and Bei Xu.

I would like to take this opportunity to thank my parents, parents in law, and daughter. Special thank gives to my wife for her consistent support at every stage of my academic career.

Finally, I hope this book will help in promoting research of the Knowledge Grid and the Cyber-Physical Society.

Hai Zhuge
Spring in 2012

Contents

Foreword v

Preface vii

1. The Knowledge Grid Methodology 1
   - 1.1 The Knowledge Space — Knowledge as a Space 1
   - 1.2 The Cyber Space 7
   - 1.3 Effort Toward the Intelligent Interconnection Environment 9
   - 1.4 Challenge and Opportunity 14
   - 1.5 The Knowledge Grid Environment 18
     - 1.5.1 The notion 18
     - 1.5.2 Virtual characteristic 20
     - 1.5.3 Social characteristic 21
     - 1.5.4 Adaptive characteristic 22
     - 1.5.5 Semantic characteristic 23
   - 1.6 Epistemology and Knowledge 29
   - 1.7 Ontology 32
   - 1.8 System Methodology 33
     - 1.8.1 The theory of dissipative structure 33
     - 1.8.2 Synergetic theory 34
     - 1.8.3 The hypercycle — a principle of natural self-organization 35
     - 1.8.4 General principles and strategies 37
   - 1.9 Dynamic Knowledge Management 40
   - 1.10 The Knowledge Grid as a Research Area 43
     - 1.10.1 Research Scope 43
     - 1.10.2 Parameters 45
     - 1.10.3 Distinctive characteristics of the Knowledge Grid 46
     - 1.10.4 The Knowledge Grid's general research issues 47
     - 1.10.5 Difference between the Web and the Knowledge Grid 50
     - 1.10.6 The technological basis of the Knowledge Grid 51
     - 1.10.7 A new dream 52

1.10.8 Beyond the vision of Turing and Bush 57

2. The Semantic Link Network 65
2.1 The Idea of Mapping 65
2.2 Basic Concepts and Characteristics 67
2.3 Relational Reasoning and the Semantic Space 78
2.4 An Algebraic Model of the SLN 82
2.5 SLN Normalization 86
2.5.1 The normal forms of an SLN 86
2.5.2 Operations on SLNs 87
2.6 Constraints and Displaying 91
2.7 SLN Ranking 92
2.7.1 Hyperlink network ranking. 92
2.7.2 SLN ranking 93
2.7.3 A ranking algorithm. 94
2.8 Implementation of SLN Operations 96
2.8.1 Matching between SLNs 96
2.8.2 The union operation 99
2.8.3 SLN-level reasoning 101
2.9 SLN Analogical Reasoning 102
2.9.1 Analogical reasoning modes 103
2.9.2 Process of analogical reasoning 107
2.10 Dynamic SLN 110
2.11 SLN Abstraction 116
2.12 Application: SLN-based Image Retrieval 120
2.13 Application: Active Document Framework (ADF) 124
2.14 Application: e-Learning 127
2.15 Potential Applications, Relevant Work and Q&A 130
2.16 SLN 2.0: Autonomous Semantic Data Model 134
2.17 Probabilistic Semantic Link Network 141
2.18 Discovering Semantic Link Network 144
2.18.1 The General process. 144
2.18.2 Discover semantic links by content analysis 146
2.18.3 Enrich semantic links through evolution 148
2.19 SLN 3.0: Cyber-Physical-Socio-Mental Network 151
2.19.1 Origin of semantics 151
2.19.2 Characteristics 155
2.19.3 Intention and extension 156
2.19.4 Semantic Link Network of events SLN-E 158
2.19.5 Through minds via words 160
2.19.6 Through society, culture and thought 166
2.19.7 Principles of emerging semantics 172

2.19.8 Discovering semantic communities — semantic localization 176
2.19.9 SLN-based relation and knowledge evolution 182
2.19.10 Building and performing semantic images 184
2.19.11 Structure and networking rules of social space 186
2.19.12 Communication rules and principles 188
2.19.13 Influence between mental space and social space 191
2.19.14 Application in cognitive-behavioral therapy 196
2.20 Principles of Mental Concepts 198
2.21 Discussion: Philosophy, Language and Semantics 203

3. A Resource Space Model 211
3.1 Examples of Using Multi-Dimensional Classifications 211
3.2 The Virtual Grid 213
3.3 The Resource Space Model (RSM) 215
3.3.1 Resource spaces 215
3.3.2 Normal forms 218
3.4 Criteria for Designing Resource Spaces 222
3.5 Designing Resource Spaces 224
3.6 Representation of Resources 225
3.7 The Resource Using Mechanism (RUM) 226
3.8 Comparisons 228
3.9 Dealing with Exponential Growth of Resources 232
3.10 Extension of the Resource Space Model 235
3.10.1 Formalizing resource space 235
3.10.2 Resource space schemas and normal forms 235
3.10.3 Topological properties of resource spaces 241
3.11 Integrity Constraint for the Resource Space Model 244
3.11.1 Entity integrity constraints 244
3.11.2 The membership integrity constraint 248
3.11.3 Referential integrity constraints 248
3.11.4 User-defined integrity constraints 252
3.12 Storage for Resource Space and Adaptability 254
3.13 Application: Faceted Navigation 257
3.14 Application: Personal Resource Space 261
3.14.1 The idea 261
3.14.2 Uploading linked resources 265
3.15 The Dimension 267

4. The Single Semantic Image 271
4.1 Obtaining Single Semantic Image from Multi-Facet Views 271
4.2 Combining SLN and RSM 273
4.3 The SSeI Mechanism 275

4.4 The Single Semantic Image Query Language 278
4.5 SSeIQL Syntax Specification 281
4.5.1 Space definition 281
4.5.2 Multiple space manipulation 283
4.5.3 Resource modification 285
4.5.4 Operating semantic link space 287
4.5.5 View definition 287
4.6 The Programming Environment 289
4.7 The Single Semantic Image Browser 292
4.8 The SSeI in a Peer-to-Peer Semantic Link Network 295
4.9 Abstraction Level, Time, Epistemology, Location and Space .. 298
4.10 A Semantic Lens 300
4.11 A Semantic Lens for Text Visualization 304
4.12 Single Semantic Image through Multiple Channels 308
4.13 Philosophical Discussion 311

5. Knowledge Flow 313
5.1 Concept 313
5.2 A Knowledge Flow Process Model 316
5.3 Peer-to-Peer Knowledge Sharing 317
5.4 Knowledge Intensity 319
5.5 Knowledge Flow Principles 320
5.6 Computational Model of Knowledge Intensity 322
5.6.1 Computing knowledge intensity in a closed environment 322
5.6.2 Computing knowledge intensity in an open environment 324
5.6.3 Knowledge intensity evaluation 325
5.7 Knowledge Spiral Model 326
5.8 Knowledge Flow Network Planning 327
5.8.1 Composition operations and principles 328
5.8.2 Knowledge flow network components 329
5.8.3 The team organization principle 331
5.9 Resource-Mediated Knowledge Flows 331
5.10 Exploring Knowledge Flows 336
5.10.1 Market mechanism 336
5.10.2 Knowledge growth 337

6. Exploring Scale-Free Network 339
6.1 Concepts 339
6.2 The Topologies of Some Real Networks 341
6.2.1 The Internet 341
6.2.2 The World Wide Web 343

6.2.3 Networks of citations of scientific papers 345
6.2.4 Networks of collaboration 347
6.2.5 Networks of human language 349
6.2.6 Other networks 351
6.3 Random Graph Theory 352
6.4 The Small-World Theory 354
6.5 Modeling Measures for Live Scale-Free Networks 356
6.5.1 The Barabási-Albert model 357
6.5.2 Generalizations of the Barabási-Albert model 359
6.5.2.1 Link rewiring 359
6.5.2.2 Node attractiveness 360
6.5.3 The idea of random fraction for Web growth 361
6.5.4 The Krapivsky-Redner model 362
6.5.5 The Simon model 364
6.5.6 An example from software engineering 365
6.5.7 Other growth models with constraints 366
6.5.7.1 Decaying networks 366
6.5.7.2 Aging networks 366
6.5.7.3 Fitness networks 367
6.5.7.4 Age or cost constrained networks 368
6.6 Modeling Actual Scale-Free Network 369
6.6.1 An urn transfer model for a live scale-free network 371
6.6.2 A directed evolving graph for a live scale-free network 376
6.6.3 Experiments and analysis 380
6.6.4 Further consideration and comparisons 385
6.6.5 Proof of the proposition 387
6.7 Summary and Implications 390

7. Topological Centrality in Social Network 393
7.1 Principles of Influence 393
7.2 Basic Concept 394
7.3 Calculating Topological Centrality 396
7.4 Discovering Research Communities 399
7.5 Discovering Backbone in Research Network 402
7.6 Discussion 404

8. A Peer-to-Peer SLN for Decentralized Knowledge Sharing 407
8.1 Decentralized Peer-to-Peer Recommendation and Query 407
8.1.1 Peer-to-Peer networks 408
8.1.2 Towards a semantic Peer-to-Peer network 409
8.1.3 Characteristic of the Peer-to-Peer Semantic Link Network 410
8.2 Basic Idea and Architecture 411

8.3 Semantic Links for Peer-to-Peer Networking 413
8.3.1 Explicit semantic link 413
8.3.2 Implicit semantic link 413
8.3.3 Reasoning operations 417
8.4 Communities in P2P-SLN 418
8.5 Routing Strategies and Process 420
8.5.1 Routing process 420
8.5.2 Adaptive TTL 422
8.5.3 Semantics-based neighbor selection 423
8.6 Self-organization of Communities and Evolution 424
8.6.1 Peer join 424
8.6.2 Peer departure 425
8.7 Application and Experiment 425
8.7.1 Instantiation of P2P-SLN framework 426
8.7.2 Peer join and departure 427
8.7.3 Query routing process 429
8.7.4 Performance analysis 430
8.7.5 Experiment 432
8.7.5.1 Term vector simulation results 432
8.7.5.2 Query-answer simulation results 435
8.7.5.3 Semantic query experiment 437
8.8 Performance of P2P-SLN 438
8.9 Incorporating Recommendation 439
8.9.1 Recommendation process 439
8.9.2 Query routing process 440
8.9.3 Experiment 441
8.10 Toward P2P Knowledge Flow Networking 442

9. P2P Semantic Overlay Networks 447
9.1 Structure Approaches 447
9.1.1 An order-preserve structured P2P network 449
9.1.2 Distributed suffix tree 455
9.1.3 IMAGINE — A general index structure over structured DHT P2P networks 459
9.1.4 Building small-world network in multi-dimensional space 462
9.2 Unstructured Approaches 468
9.2.1 Explicit semantic Link 468
9.2.2 Gossip through semantic unstructured P2P network .. 471
9.3 Discussion 473

10. The Energy-Knowledge Grid 475
10.1 The Knowledge Grid meets the Smart Grid 475

10.2 Architecture and Characteristics 477
10.3 Multi-Dimensional Devices and Requirements..... 483
10.4 Multi-Layer Complex Networks 485

Bibliography 489

Index 503

Chapter 1

The Knowledge Grid Methodology

The development of science and technology has extended human behavior and sensation, accelerated the progress of society, and enabled people to understand the physical space and themselves more profoundly. But, we still have much to find out, especially about knowledge.

What is knowledge?

How is knowledge generated?

How does knowledge evolve?

Can knowledge be inherited?

How is knowledge shared effectively?

How is knowledge stored and retrieved?

How to enable machines to obtain knowledge so that they can act intelligently?

How to create an environment to enable, facilitate, or improve knowledge creation, evolution, inheritance, sharing and service?

These are fundamental philosophic, scientific and technologic issues relevant to the Knowledge Grid methodology.

1.1 The Knowledge Space — Knowledge as a Space

Knowledge is a multi-dimensional complex space, where dimensions emerge, evolve structures (from simple to complex, or from complex to simple at higher abstraction level), and influence each other. A

dimension can be viewed as a space, and a space can be viewed as a dimension. The dimensions include time, the physical space, the socio space, and the mental space. Each space includes individuals, structures, rules and statuses. The mental space reflects the other spaces, builds mental semantic images, and carries out reasoning while various individuals interact with each other. A point in the space is the reflection of a set of individuals, which share one set of projections on all dimensions.

According to the new notion, knowledge is not just in the mental space, it is reflected by the cyber space, physical space and social space, and it evolves with the interaction between individuals in various spaces. Minds can reflect, discover and link knowledge in these spaces through experiencing and thinking. Machines can help discover some rules in the cyber space by statistical approaches.

The knowledge space has the following characteristics.

(1) *There are multiple origin points.* The first one is the generation of the physical space and time. The generation of mind and society is the remarkable point of the development of human beings.
(2) *A dimension can be linear or nonlinear structure of coordinates.* A coordinate can be a tree structure. A dimension can be as simple as time or as complicated as a large-scale complex system. A dimension can be regarded as a space.
(3) *Projections of points on dimensions evolve with time.* Some points do not have projections on some dimensions at certain time. For example, some rules in the physical space have no projection on the dimension of the mental space. Therefore, these rules are currently unknown.
(4) *Projection on some dimensions may concern a complex process.* For example, the projection of a point onto the dimension of the mental space involves in a complex process of perception, learning, communication, association, and reasoning.
(5) *Knowledge transition between individuals also involves in the processes in and through multiple spaces.*

Fig. 1.1.1 depicts a knowledge space, where O_0 denotes the origin of generating the universe and time, and O_h denotes the origin of generating the mental space and the social space. The generated spaces evolve with time according to different rules.

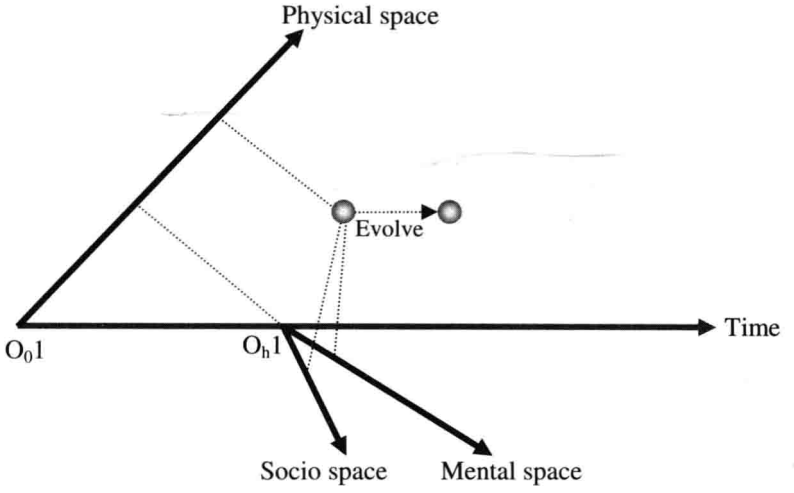


Fig. 1.1.1 A macroscopic knowledge space.

This notion of knowledge space extends the generation of knowledge not only to the origin of the mental space and the socio space but also to the origin of the physical space and time. The notion implies that knowledge is not only subjective but also objective, not only in individual mental processes but also in social processes. The links between knowledge points in and through spaces play an important role in forming and evolving the space structure.

This knowledge space notion opens new door to explore knowledge.

The generation of any space accompanies the creation of its dimensions. The evolution of spaces accompanies the expansion or shrink of dimensions. *A dimension can be viewed as a space from its coordinates. A space can be viewed as a dimension from its super-space.*