

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 objective world and themselves more profoundly. But we still have much to find out, especially about machine-enabled knowledge creation, evolution, inheritance and sharing.

Knowledge in nature is a product of society. It evolves and endures throughout the life of a culture rather than that of an individual as mentioned in “As We May Think” (V. Bush, *The Atlantic Monthly*, 1945, vol.176, no.1, pp.101-108). Modern communication facilities like the Internet provide people with unprecedented social opportunities for knowledge generation and sharing. However, our increasing computing power and communication bandwidth does not of itself improve this knowledge generation and sharing. To do this, the semantic ability of the facilities that transmit and store knowledge must be improved. Improving our social interaction in this way would help enrich knowledge in our society by supporting social activities at different levels (both the physical and the mental level) and in different environmental spaces (entity space, semantic space and knowledge space).

The ideal of the Knowledge Grid is to foster worldwide knowledge creation, evolution, inheritance, and sharing in a world of humans, roles and machines as stated in “China’s e-Science Knowledge Grid Environment” (H. Zhuge, *IEEE Intelligent Systems*, 2004, vol.19, no.1, pp.13-17; the Knowledge Grid forum, center and community portal <http://www.knowledgetgrid.net>).

The capturing and expressing of semantics involves complex psychological and cognitive processes. The exploitation of psychology, cognitive science and philosophy plays an important role in studying semantics—the basis for knowledge sharing.

The Knowledge Grid methodology is a multi-disciplinary system methodology for establishing a global knowledge world that obeys the principles and laws of economics, nature, society, psychology and information technology.

Implementation of the Knowledge Grid will speed up the development of human civilization.

1.1 Towards The Next-Generation Web

The Internet and the World Wide Web are milestones of information technology. People have become increasingly reliant on them for supporting modern work and life. For example, scientists can communicate with each other using net forums and email, share their experimental data and research results by posting them in Web pages on personal or corporate websites, and retrieve technical reports and academic papers of interest to them from online digital libraries or from less formal websites using general-purpose search engines.

But the exponential growth and intrinsic characteristics of the Web and its pages prevent people from effectively and efficiently sharing information. Much effort has been put into solving this problem with but limited success. In any case it is hard for the Web to provide intelligent services because the representation used by the current Web does not support the inclusion of semantic information.

With the development of communication facilities and Web applications, computing is struggling to extend its support from individual to group and social behavior, from closed to open systems, from simple and centralized to complex and distributed computing, and from static computing to dynamic and mobile information, computing and knowledge services.

To overcome the deficiencies of the current Web, scientists and developers are working towards a next-generation Web. These efforts lie

in three main categories: the first includes the *Semantic Web*, *Web Service* and *Web Intelligence*, which aim to improve the current Web, the second is the *Grid* (<http://www.gridforum.org>), which aims at a new application platform of the Internet, and the third is *Peer-to-Peer Computing*, which enables resource sharing in an egalitarian, large-scale and dynamic network as discussed in “Peer-to-peer Prospects” (D. Schoder and K. Fischbach, *Communications of the ACM*, 2003, vol.46, no.2, pp.27-29). Recent reports indicate that all are indeed moving closer to their targets.

The aim of the *Semantic Web* (<http://www.semanticweb.org>) is to support cooperation between Web resources by establishing ontological and logical mechanisms by using standard markup languages like XML (eXtensible Markup Language, <http://www.w3.org/XML>), RDF (Resource Description Framework, <http://www.w3.org/RDF>), OIL (Ontology Interchange Language) and DAML (DARPA Agent Markup Language) to replace HTML (HyperText Markup Language, <http://www.w3.org/MarkUp>) and to allow Web pages to hold descriptions of their content.

The aim of *Web Service* is to provide an open platform for the development, deployment, interaction, and management of globally distributed e-services based on Web standards like UDDI (Universal Definition Discovery and Integration) and WSDL (Web Service Description Language, www.w3.org/TR/wsdl). It enables the integration of services residing and running in different places. Using intelligent agents is an important technique that can be used to implement active Web Service.

The aim of *Web Intelligence* is to improve the current Web by using artificial intelligence (especially distributed intelligence) and information processing technologies such as symbolic reasoning, text mining, information extraction and information retrieval.

The aim of the global *Grid* is to share, manage, coordinate, schedule and control distributed computing resources, which could be machines, networks, data, and any types of devices. The ideal of the Grid is that any compatible device could be plugged in anywhere on the Grid and be guaranteed the required services regardless of their locations, just like the electrical power grid. Grid computing does not use Web technologies.

The *Semantic Grid* (<http://www.semanticgrid.org>) attempts to incorporate the advantages of the Grid, Semantic Web and Web Service approaches. The Grid architecture has become the service-oriented Open Grid Services Architecture (OGSA) (I.Foster, C.Kesselman, J.M.Nick, and S.Tuecke, "Grid Services for Distributed System Integration", *Computer*, 2002, vol.35, no.6, pp.37-46), in which some features of Web Service can be plainly seen.

By defining standard mechanisms for creating, naming, and locating services, the Semantic Grid can incorporate peer-to-peer technology under OGSA, and so enable autonomous computing objects to cooperate in a network of equals and with scalability.

Peer-to-peer networking should work not only at the computing level but also at the semantic level. How to automatically map from a semantically rich space into a peer-to-peer network is an important research problem that must be solved before the gap between the peer-to-peer network and high-level intelligent applications can be bridged.

1.2 Challenges and Opportunities

The development of operating systems, advanced languages, and database systems were crucial events in computing history, and were also very important to the success of personal computing. The Internet and its applications are now moving computing towards an open and socialized service environment. But it is still difficult for the current Internet application platform (based on client/server architecture) to support complex and intelligent applications, especially under large-scale and dynamic conditions.

Demanding application requirements provide researchers with many challenges and opportunities. Complex and intelligent applications require the support of completely new databases, programming languages, and operating systems to support distributed sharing, coordination, deployment, execution and management of services and resources, and the implementation of applications that are beyond the ability of the current Web. Fig. 1.1 depicts the evolutionary trend of computing.

The Knowledge Grid is a newly proposed computing platform, which supports new resource organization models, new computing models, and new networks such as sensor, mobile and wireless networks.

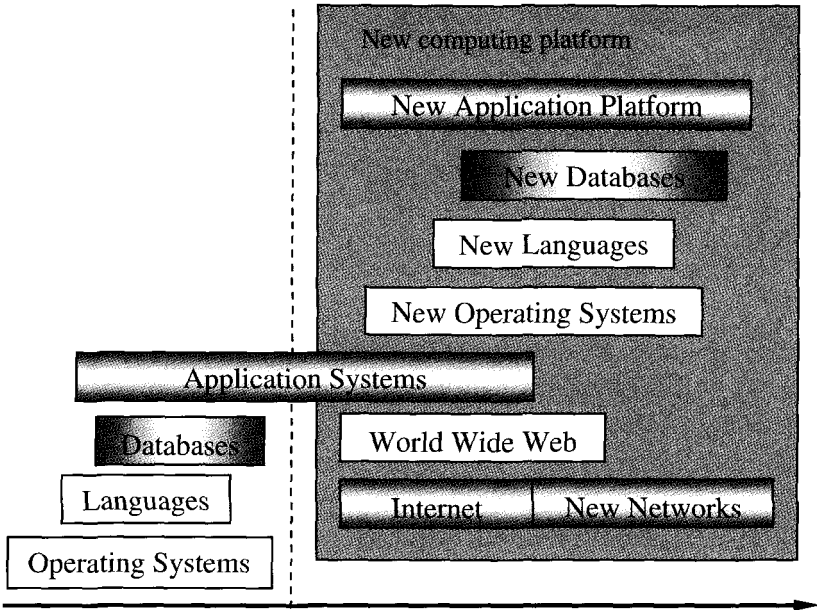


Fig. 1.1 The evolution of computing environments.

1.3 Towards the Knowledge Grid

Acquiring knowledge is the basis for intelligent services. Fran Berman pointed out the challenge of extracting knowledge from terabytes of data in her essay "From TeraGrid to Knowledge Grid" (*Communications of the ACM*, 2001, vol.44, no.11, pp.27-28). She commented that the "Knowledge Grid" has to synthesize knowledge from data by means of mining and reference, and to enable search engines to make references, answer questions, and draw conclusions from masses of data. Recent

understanding and development of the Knowledge Grid has gone beyond the scope of Fran Berman's proposal.

1.3.1 *Virtual characteristics*

Many computer scientists are exploring ideal computing and resource organization models for the next-generation Web. The editorial of the first special issue on Semantic Grid and Knowledge Grid for the *Future Generation Computer Systems* journal (2004, vol.20, no.1, pp.1-5) described the following scenario: The future interconnection environment will be a platform-irrelevant *Virtual Grid* consisting of requirements, roles and resources.

With machine-understandable semantics, a resource can actively and dynamically cluster and fuse relevant resources to provide on-demand services by understanding requirements and functions and relating them to each other. Versatile resources are encapsulated to provide services by way of a *single semantic image* within the uniform resource model. A resource can intelligently assist people to accomplish complex tasks and solve problems by employing versatile resource flow cycles through virtual roles to use appropriate knowledge, information, and computing resources.

With the development of society and science, people have a more profound understanding of nature, society, and themselves than ever before. So appropriate rules and principles of nature, society and economics should be adopted when we develop the next-generation Web.

1.3.2 *Social characteristics*

The Knowledge Grid has social characteristics. In the real world, people live and work in a *social grid* obeying social and economic rules and laws. The Knowledge Grid is a *virtual social grid*, where people enjoy and provide services through versatile flow cycles like control flows, material flows, energy flows, information flows and knowledge flows.

People can communicate and gain knowledge from each other through mutually understandable semantics.

An artificial interconnection environment can only be effective when it works harmoniously with its social grids. For example, an effective e-business environment requires harmonious cooperation between information flows, knowledge flows, material flows, e-services and social services. The e-services belong to e-business platforms. The material flows and social services belong to society. The information flows and knowledge flows belong to both society and the platforms. The proper semantic representation supports mutual understanding between the social grid and the artificial interconnection environment. In future, different artificial interconnection environments will co-exist and compete with each other for survival, rights and reputation, and will harmoniously evolve with the social grid (H.Zhuge and X.Shi, "Eco-grid: A Harmoniously Evolved Interconnection Environment", *Communications of the ACM*, 2004, vol.47, no.9, pp79-83).

1.3.3 Adaptive characteristics

"On-demand services" is a fashionable catchphrase in the context of the future Web. But there is no limit to demand. So, to provide all participants with services on demand is impossible and unreasonable as long as service generation and service provision carry a significant cost and services themselves differ in quality.

Economics is concerned with three kinds of entity: participants, markets and economic systems. The market is an important mechanism for automatically and reasonably adjusting the decisions and behaviors of market participants, for example, agents and soft-devices (H.Zhuge, "Clustering Soft-Devices in Semantic Grid", *IEEE Computing in Science and Engineering*, 2002, vol.4, no.6, pp.60-63). Besides the influence of the market, participants' behaviors and decisions can be adjusted by negotiation. Governments, organizations and social rights also play important roles in influencing market participants' behavior and decision making. Market participants, producers and consumers, look for satisfactory rather than optimal exchanges through agreement (the

evaluation of “satisfactory” involves psychological factors). Being based on simple principles, the market mechanism adapts by avoiding complex computation.

The natural ecological system establishes a balance among natural species through energy flow, material flow and information flow. These flows in their turn influence the social system (H. Zhuge and X. Shi, “Fighting Epidemics in the Information and Knowledge Age”, *Computer*, 2003, vol.36, no.10, pp.114-116). Different species evolve together as parts of the entire ecological system.

The Knowledge Grid also supports three major roles: producers, consumers and a market mechanism for adapting to the behavior of different participants. It should adopt economic and ecological principles to balance the interests of knowledge producers and knowledge consumers, and adapt to knowledge evolution and expansion.

1.3.4 Semantic characteristics

Research on semantic information processing has a long history in the computing field (M.L. Minsky, ed., *Semantic Information Processing*, MIT Press, 1968). Knowledge representation approaches such as frame theory (Minsky, 1975), the Knowledge Representation Language KRL (D.G. Bobrow, 1979) and the Semantic Network (M.R. Quillian; H.A. Simon, 1970) are approaches to expressing semantics. Before the emergence of the Internet interchange standard XML and the Resource Description Framework RDF (<http://www.w3.org/RDF/>), the Knowledge Interchange Format KIF (<http://www.logic.stanford.edu/kif>) and Open Knowledge Base Connectivity OKBC (<http://www.ai.sri.com/~okbc>) were two standards for knowledge sharing.

Knowledge acquisition is the bottleneck of knowledge engineering. Data mining approaches help a bit by automatically discovering knowledge (association rules) in large-scale databases. These approaches can also be used to discover semantic relationships within and between texts.

Why were the symbolic approaches, especially the KIF and OKBC of AI, and ODBC in the database area, not widely adopted in the Internet age?

One cause is the success of HTML, which is easy to use both for a writer and, in cooperation with a browser, for a reader. Its main advantage is that “anything can link to anything” (T. Berners-Lee, J. Hendler and O. Lassila, Semantic Web, *Scientific American*, vol.284, no.5, 2001, pp.34-43).

A second cause is that traditional AI’s knowledge representation approaches try to explicate human knowledge, while the Web focuses on structuring Web resources and the relationship between resources, that is, it is more concerned with semantics.

A third cause is that cooperation between machines (applications) has become the dominant aim in realizing intelligent Web applications, while traditional knowledge engineering focuses on cooperation between human and machine.

A fourth cause is the cross-platform requirement. Consequently XML has been adopted as the information exchange standard of the Web.

What are the semantic problems of the Internet age?

The first is the acquisition problem — to automatically acquire semantic relationships within and between resources.

The second is the representation problem — but the focus is on expressing the semantic relationships between Web resources (just as in the ER model in the relational data model, the focus is on expressing semantic relationships between entities), on establishing an appropriate semantic computing model, and on seeking an approach that synthesizes the semantics expressed within different semantic spaces. Prior research into expressing the internal semantics of resources helps solve this problem.

The third is the normal organization problem — to properly organize semantic information under semantic normal forms and integrity constraints so that the correctness and efficiency of semantic operations is guaranteed. If we can solve this problem, Web resources can be correctly used in light of their semantics.

The fourth is the problem of processing semantics — to refine, abstract and synthesize large-scale semantic information to provide appropriate and succinct semantic information.

The fifth is the maintenance problem — to maintain correct semantics in large-scale and dynamic semantic resource spaces.

Solving the above problems is beyond the scope of current research on the Semantic Web. The RDF would be a less than ideal approach to expressing the semantics of Web resources because it does not directly address the semantic problems of the Internet age.

At a high level, the Knowledge Grid is a “world” of requirements, roles and services. Services are provided by resources that are implemented on the basis of a uniform resource model. Services can actively find and advertise requirements. People can play the roles of services, and enjoy services provided by others. Requirements and services are organized into conglomerations that belong to different communities. Some services can play the broker role and be responsible for dynamically integrating services to meet varying requirements.

At a low level, a Semantic Resource Space organizes and uses resources by way of a *single semantic image*, that is, various resources are mapped onto a single semantic space to expose their commonality. The single semantic image can be realized by semantic relationship models like orthogonal classification and semantic linking. Normalization theory guarantees the correctness and effectiveness of resource operations (H.Zhuge, “Semantics, Resources and Grid”, *Future Generation Computer Systems*, 2003, vol.20, no.1, pp.1-5). The Semantic Resource Space requires a semantic browser that enables not only people but also services to exploit the semantics of a resource being browsed, to extract from the resource and reason from the extract, to explain the display, and to anticipate subsequent browsing.

A huge gap exists between low-level semantics and high-level semantics as shown in Fig. 1.2. Current Web application software technologies and tools only use low-level semantics such as keywords and surface content (texture, color, and so on). Users have to search an enormous name space to find what they need within huge and expanding Web resources.

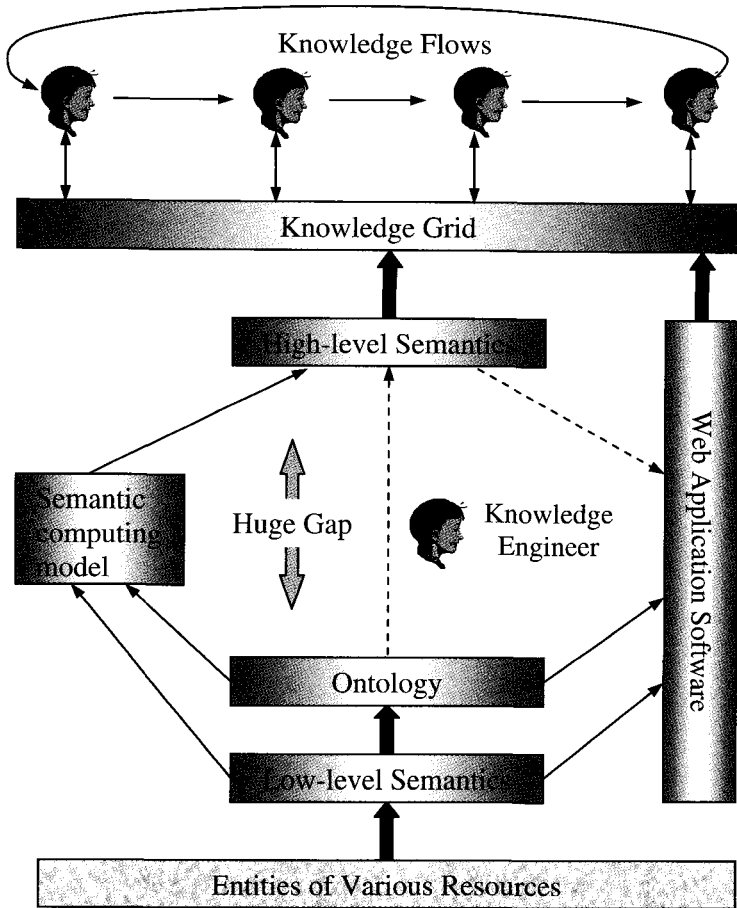


Fig. 1.2 Gap between low-level semantics and high-level semantics.

Complex cognitive and psychological processes are involved in this huge gap. Some computer scientists seek to obtain high-level semantics by establishing domain ontology mechanisms for explaining the semantics of resources and their ties. Building tools and finding methods

for creating ontologies are becoming popular in the Semantic Web area. Ontology can reflect people's consensus on semantics in name spaces or symbolic spaces to a certain extent. But it is hard to cope with the complexity of human cognitive processes.

A *semantic computing model* that could at least partly bridge this huge gap must be found. The model should go beyond the scope of traditional formal semantics, which has been extensively investigated in computer science in the past with success that is significant in theory but limited in practice. The model should balance the formal and informal and reflect human cognitive characteristics.

Epistemology, which has been neglected in previous efforts towards the next-generation Web, plays the key role in human cognitive processes.

1.4 Epistemology

Epistemology is a branch of philosophy. It concerns the nature, scope and source of knowledge. Computer scientists could improve their understanding of computing by studying the history and variety of epistemology.

Empiricism sees knowledge as the product of sensory perception. Knowledge results from a kind of mapping or reflection of external objects, from our sensory organs, possibly assisted by some observational instruments, into our brain to be used by our mind.

Rationalism considers knowledge to be the product of rational reflection. Knowledge results from the organization of perceptual data on the basis of cognitive structures called *categories*. Categories include *space, time, objects* and *causality*.

Pragmatism holds that knowledge consists of mental models that simplify our perception of the real world. It is assumed that a model only reflects the main characteristics of the real world. Otherwise, it would be too complicated to be of any practical use. Thus different or even seemingly contradictory models for solving the same problem could co-exist. Problem solving is a process of developing and selecting useful

models. Mathematical modeling plays an important role in scientific problem solving.

Individual constructivism sees the individual as trying to build coherence between different pieces of knowledge. In the mental construction process, knowledge that succeeded in integrating previously incoherent pieces of knowledge will be kept, and knowledge that is inconsistent with the bulk of other knowledge that the individual has will tend to be rejected.

Social constructivism regards consensus on different subjects as the ultimate criterion for judging knowledge. Truth or reality will be accorded only to those constructions on which most people of a social group agree. Karl Marx's theory of ideology can be regarded as a type of social epistemology. An ideology is a set of beliefs, a world view, or a form of consciousness that is in some fashion false or delusive. The cause of these beliefs and their delusiveness is the social situation and interests of the believers. The theory of ideology is concerned with the truth and falsity of beliefs, so it is a kind of *classical social epistemology*. *Feminist epistemology and philosophy of science* studies the ways in which gender influences, and ought to influence, our conceptions of knowledge, the knowing subject, and practices of inquiry and justification.

Evolutionary epistemology assumes that knowledge is constructed by a subject or group of subjects in order to adapt to their environment in the broad sense. That construction is a process going on at different levels—at biological as well as psychological or social levels. Construction happens through blind variation of existing pieces of knowledge, and the selective retention of those new combinations that somehow contribute more to the survival and reproduction of the subjects within their given environment. Knowledge is regarded as an instrument questing for survival. Evolutionary epistemology emphasizes the importance of natural selection. Selection is the generator and maintainer of the reliability of our sensory and cognitive mechanisms, as well as of the *fit* between those mechanisms and the world. Trial and error learning and the development of scientific theories can be explained as evolutionary selection processes.

A recent evolutionary view is that knowledge can actively pursue goals of its own. It notes that knowledge can be transmitted from one individual to another, and thereby lose its dependence on any particular individual. A piece of knowledge may be successful even though its predictions may be totally wrong, as long as it is sufficiently convincing to new carriers. In this theory, the individual having knowledge has lost his primacy, and knowledge becomes a force in its own right with proper goals and ways of developing itself. This emphasizes communication and social processes in the development of knowledge, but instead of regarding knowledge as the object constructed by the social system, it rather views social systems as constructed by knowledge processes. Indeed, a social group can be seen as organized by members sharing the same types of knowledge. *To keep evolution sustainable, knowledge should have the characteristic of diversity.*

The following references can help readers learn more about epistemology:

- (1) F. Heylighen (1993), "Epistemology Introduction", in: F. Heylighen, C. Joslyn and V. Turchin (editors): *Principia Cybernetica Web*, <http://pespmc1.vub.ac.be/epistemology.html>.
- (2) P.D. Klein, (1998). "Epistemology". In E. Craig (ed.), *Routledge Encyclopedia of Philosophy*. London: Routledge. <http://www.rep.routledge.com/article/P059>.
- (3) E. Anderson, "Feminist Epistemology and Philosophy of Science", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/feminism-epistemology/>.
- (4) M. Brady and W. Harms, "Evolutionary Epistemology", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/epistemology-evolutionary/>.
- (5) R. Feldman, *Stanford Encyclopedia of Philosophy* article, "Naturalized Epistemology".
- (6) A. Goldman, "Social Epistemology", *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/entries/epistemology-social/>.

1.5 Ontology

Ontology — another branch of philosophy — is the science of what is, and of the kinds and structures of the objects, properties and relations in every area of reality. Ontology in this sense is often used in such a way as to be synonymous with metaphysics. In simple terms, it seeks to classify entities. Each scientific field has its own preferred ontology, defined by the field's vocabulary and by the canonical formulations of its theories.

Traditional ontologists tend to model scientific ontologies by producing theories, organizing them and clarifying their foundations. Ontologists are concerned not only with the world as studied by sciences, but also with the domains of practical activities such as law, medicine, engineering, and commerce. They seek to apply the tools of ontology to solving problems that arise in these domains.

In the field of information processing, different groups of data-gatherers have their own idiosyncratic terms and concepts that guide how they represent the information they receive. When an attempt is made to put information together from different groups, methods must be found to resolve terminological and conceptual incompatibilities. At first such incompatibilities were resolved case by case. Then people gradually came to realize that providing once and for all a common backbone taxonomy of entities relevant to an application domain would have significant advantages over resolving incompatibilities case by case. This common backbone taxonomy is called an *ontology mechanism* by information scientists.

In the context of knowledge sharing and reuse, an ontology mechanism establishes a terminology for members of a community of interest. These members can be humans, application software, or automated agents. An ontology can be represented as a formal vocabulary organized in taxonomic hierarchies of classes, whose semantics is independent of both user and context. Readers can obtain more information about ontology from the Semantic Web website (<http://www.semanticweb.org>).

1.6 System Methodology

Web and Grid computing have not yet made use of the principles and methods of system methodology.

Darwin's theory of evolution holds time to be an "arrow" of evolution. The subjects of the evolutionary process evolve as time progresses, so the overall process is irreversible, just like time. Life evolves from the simple to the complex, from a single-celled ameba to a multi-celled human being.

The second law of thermodynamics (R.J.E. Clausius and L. Boltzmann) tells us of a degenerate arrow: all processes manifest a tendency toward decay and disintegration, with a net increase in what is called the *entropy*, or state of randomness or disorder, of the overall system.

The theory of dissipative structure was created against the background of the collision between the two arrows.

1.6.1 *The theory of dissipative structure*

A system with dissipative structure is an open system that exists far from thermodynamic equilibrium, efficiently dissipates the heat generated to sustain it, and has the capacity to change to higher levels of orderliness.

According to Prigogine's theory, systems contain subsystems that continually fluctuate. At times a single fluctuation or a combination of them may become so magnified by positive feedback that it shatters the existing organization. At such revolutionary moments, it is impossible to determine in advance whether the system will disintegrate into chaos or leap to a new, more differentiated, higher level of order. The latter case is called a dissipative structure, so termed because it needs more energy to sustain itself than the simpler structure it replaces and is limited in growth by the amount of heat it is able to dissipate.

According to the theory of dissipative structure, the exponential growth of Web resources tends to disorder. The current efforts towards the future Web are trying to establish a new kind of order — the order of diverse resources. But how such an order can be prevented from

becoming disordered again is a critical issue that needs to be considered as we work towards the future Web.

Here is a very interesting question: *Can we design a dissipative structure for the future Web?*

1.6.2 Synergetic theory

Synergetics is a theory of pattern formation in complex systems. It tries to explain structures that develop spontaneously in nature. Readers can obtain more information from H. Haken's works (*Synergetics, An Introduction: Nonequilibrium Phase-Transitions and Self-Organization in Physics, Chemistry and Biology*, Springer, 1977; *Synergetics of Cognition*, Springer-Verlag, 1990 (with M. Stadler); *Principles of Brain Functioning: A Synergetic Approach to Brain Activity, Behavior, and Cognition*, Springer-Verlag, 1995).

The purpose of introducing the relevant concepts here is to provoke constructive thought about their possible influence on the future Web.

How order emerges out of chaos is not well defined, so synergetics employs the ideas of probability (to describe uncertainty) and information (to describe approximation). Entropy is a central concept relating physics to information theory. Synergetics concerns the following three key concepts: compression of the degrees of freedom of a complex system into dynamic patterns that can be expressed as a collective variable; behavioral attractors of changing stabilities; and the appearance of new forms as nonequilibrium phase transitions.

Systems at instability points are driven by a *slaving principle*: long-lasting quantities can enslave short-lasting quantities (that is, they can act as order parameters). Close to instability, stable motions (or "modes") are enslaved by unstable modes and can be ignored, thereby reducing the degrees of freedom of the system. The macroscopic behavior of the system is determined by the unstable modes. The dynamic equations of the system reflect the interplay between stochastic forces ("chance") and deterministic forces ("necessity").

Synergetics deals with self-organization, how collections of parts can produce structures. Synergetics applies to systems driven far from equilibrium, where the classic concepts of thermodynamics are no longer adequate. Order can arise from chaos and can be maintained by flows of energy or matter.

Synergetics has wide applications in physics, chemistry, sociology and biology (population dynamics, evolution, and morphogenesis). Completely different systems exhibit surprising analogies as they pass through an instability. Biological systems are unique in that they exhibit an interplay between structure and function that is embodied in structure and latent in form.

The ideas introduced above imply that synergetics can help us explore the intrinsic self-organization principle of the future Web and its resources (for example, how the components such as services can form a well behaved structure to provide an appropriate service), and find better approaches to solve existing problems in information processing.

1.6.3 *The hypercycle — a principle of natural self-organization*

A living system has three features: *self-reproduction*, *metabolism*, and *evolution*. A *hypercycle* is a system that consists of self-reproducing macro-molecular species that are linked cyclically by catalysis. It is interesting to investigate pre-biotic evolution since it might explain how molecular species having a small number of molecules could evolve into entities with a great amount of genetic information. The idea of the hypercycle, introduced by Eigen in 1971, has been experimentally and theoretically verified by Gebinoga in 1995. The purpose of introducing here the concept of the hypercycle is to provoke some rethinking about work towards the future Web.

The following example explains the concept of the hypercycle. Living cells contain both nucleic acids and proteins, and molecules of the two classes interact. Genetic information controls the production of polypeptide chains, that is, proteins. Data encoded in nucleic acids ensure that certain proteins can be produced. Information about proteins helps the replication of nucleic acids and enables information transmission. A system of nucleic acids and proteins helping to replicate

each other is an important basis for evolution as evolution can occur only when the state information can be obtained, maintained and extended.

M. Eigen and P. Schuster consider hypercycles to be predecessors of protocells (primitive unicellular biological organisms). As quasispecies, hypercycles have also been mathematically analyzed in detail.

The self-reproducing automaton was investigated early on by John von Neumann.

A similar system of catalytically interacting macromolecules called a *syser* is comprised of a polynucleotide matrix and several proteins. There are two obligatory proteins: the replication enzyme and the translation enzyme. A *syser* can also include some structural proteins and additional enzymes. The polynucleotide matrix encodes the composition of proteins, and the replication enzyme controls the matrix replication process. The translation enzyme controls the protein synthesis according to the data encoded in the matrix. Structural proteins and additional enzymes can provide optional functions. Different *sysers* should be inserted into different organisms for effective competition.

Compared to hypercycles, *sysers* are more like simple biological organisms. The concept of *sysers* makes it possible to analyze evolutionary stages starting from a mini-*syser*, which contains only a matrix and replication and translation enzymes. An adaptive *syser* includes a simple molecular control system, which “turns on” and “turns off” synthesis by some enzyme in response to change in the external medium.

Readers can learn more about the hypercycle from M. Eigen and P. Schuster, *The Hypercycle: A principle of natural self-organization*, Springer, Berlin, 1979.

The notion of a *soft-device*, the uniform resource model of the future Web, is introduced in “Clustering Soft-Devices in Semantic Grid” (H.Zhuge, *IEEE Computing in Science and Engineering*, 2002, vol.4, no.6, pp.60-63) and envisions that the future Web will be a world of versatile soft-devices and roles. Ideal soft-devices have the same function as *sysers*.

Hypercycle theory will give us some useful notions for when we explore the organization mode of the future Web. For example, the future Web can be imagined as a living system or environment, which

consists of resource species in the form of soft-devices and versatile flow cycles. Resources could be dynamically organized into diverse flows such as knowledge flows, information flows, and service flows to provide users or applications with on-demand services. Once a requirement is confirmed, all relevant flows could be formed automatically (H.Zhuge and X.Shi, “Eco-Grid: A Harmoniously Evolved Interconnection Environment”, *Communications of the ACM*, 2004, vol.47, no.9, pp.79-83).

The future Web can be imagined as a quasihuman body, which has knowledge and intelligence and operates with special hypercycles. It can cooperate with people in a humanized way and provide appropriate, up-to-date, on-demand and just-in-time services.

1.6.4 Principles and strategies

The Knowledge Grid methodology should adopt the principles and rules of social science, economics, psychology, biology, ecology and physics, and inherit the fundamental ideas, views, rules and principles of system science.

Principles

- (1) *Integrity and uniformity principles* — The idea of integrity requires us to resolve the issue of correctness (for example, the correctness of operations). The idea of uniformity requires us to resolve the issue of simplicity, that is, to simplify a system. The integrity theory of the relational database model is a good example of integrity and it could give us useful ideas for developing the theory and system of a Knowledge Grid, especially in resource management.
- (2) *The hierarchical principle* — H.A. Simon unveils the hierarchical principle of artificial systems in his book (*The Sciences of the Artificial*, Cambridge, MA: The MIT Press, 1969). The construction of a Knowledge Grid should follow this principle. Furthermore, different levels of a system could work in different semantic spaces. Consistency should be maintained between multiple semantic spaces,

such as between the semantic space of logic and the semantic space of algebra.

- (3) *The open principle* — This principle would keep the Knowledge Grid away from the equilibrium state. Standards are a critical criterion for open systems. The Knowledge Grid could make use of the standards of the Internet, the World Wide Web and the Grid. For example, XML could be the basis for knowledge expression and transformation.
- (4) *The self-organization principle* — Resources including systems themselves can actively collaborate with each other according to some principle (for example, the economic principle) and common regulations.
- (5) *The principle of competition and cooperation* — Resources including systems (multiple systems or environments could coexist) evolve through competition and cooperation so that competitive resources or systems could play a more important role.
- (6) *The optimization principle* — Optimization means making a system more effective. Information flow, knowledge flow and service flow can be optimized to achieve efficiency in logistic processes.
- (7) *The principle of sustainable development* — Sustainable development requires individuals and communities, the interconnection environment and its human-machine interfaces, the human-machine society, and even the natural environment to harmoniously coexist and coevolve.

Strategies

The following strategies could help develop the Knowledge Grid as a future interconnection environment.

The fusion of inheritance and innovation — the Knowledge Grid environment should absorb the advantages of the Grid, the Semantic Web, and Web Services. Current Web applications should be able to work in the new environment. Smooth development would enable the future Web to exploit research on the current Web.

The fusion of centralization and decentralization — Advantage should be taken of both centralization and decentralization. On the one hand, an ideal system should be able to dynamically cluster and fuse relevant resources to provide complete and on-demand services for applications. On the other hand, it should be able to deploy the appropriate resources into the appropriate locations to achieve optimized computing.

The fusion of abstraction and specialization — On the one hand, we need to abstract a variety of resources to investigate common rules, and on the other hand, to investigate the special rules of different resources to properly integrate and couple resources.

The fusion of mobility and correctness — On the one hand, the Knowledge Grid should support mobile applications to meet the needs of ubiquitous applications. On the other hand, we should guarantee the quality of services and the means of verification.

The fusion of symbolic and connectionist approaches — Current ontology only uses the symbolic approach, which is very similar to traditional knowledge base construction. The combination of the symbolic approach and the connectionist approach would help find better solutions for intelligent applications.

The incremental strategy — As a worldwide interconnection environment, the Knowledge Grid will undergo a development process similar to that of the World Wide Web — from simple to complex, from immature to mature, from a small community to a large-scale human-machine environment with an exponential expansion of developers, users, and demands. So the Knowledge Grid development methodology should support an incremental strategy.

Adoption of new computing models — The functions of new computing models will go beyond the abilities of the Grid and of peer-to-peer and client/server networks. It is hard for any single model to yield an ideal solution. An ideal computing model should incorporate the advantages of various models or even be a set of collaborative models.

Cross-disciplinary research — Compared to natural and social systems, the artificial interconnection environment is only at a very primitive stage. The principles and rules of natural and social systems will give us useful guidelines for establishing the ideal Knowledge Grid.

1.7 Knowledge Management

Early in 1880, the American engineer F.W. Taylor investigated workers' efficiency, and formulated a scientific management method for raising productivity by standardizing operations and work. He published his authoritative book *The Principles of Scientific Management* in 1911. With increasing industrial productivity and social development, scientific management methods nowadays pay increasing attention to production processes, social changes, psychological factors, and so on.

With the development of information technology, enterprises more and more become knowledge organizations, which leads to great changes in decision processes, management methods and working conditions. P.F. Drucker pointed out that existing knowledge organizations, such as the symphony orchestra, can inspire us to develop new management approaches to knowledge organizations (*Harvard Business Review on Knowledge Management*, Boston, MA: Harvard Business School Press, 1998). After all, a large-scale orchestra can perform very well with just one conductor. Organizational learning and knowledge innovation become the key competitive abilities of a knowledge organization (I. Nonaka, "A dynamic theory of organizational knowledge creation", *Organization Science*, 1994, vol.5, no.1, pp.14-37).

The multi-dimensional knowledge space has been used by M.E. Nissen to explain organizational learning ("An extended model for knowledge-flow dynamics", *Communications of the Association for Information Systems*, 2002, vol.8, pp.251-266). An epistemological dimension is used to classify knowledge into explicit knowledge and tacit knowledge. An ontological dimension is used to describe knowledge that is shared between members of an organization. Knowledge flow within dimensions is of four types: 1) social flow, in which knowledge moves from creation by an individual to acceptance by the organization (from small to large); 2) externalization flow, in which knowledge moves from tacit form to explicit; 3) combination flow, in which the knowledge of small teams is combined and coordinated to generate the knowledge of a large team; and 4) internalization flow, in which explicit knowledge of the organization becomes tacit.

A peer-to-peer knowledge flow refers to the propagation of external and explicit knowledge that can be formalized, transmitted via communication media, and stored in computing machinery as discussed in “A knowledge flow model for peer-to-peer team knowledge sharing and management” (H. Zhuge, *Expert Systems with Applications*, 2002, vol.23, no.1, pp.23-30). The planning of a knowledge flow network seeks to formalize and optimize knowledge flows.

A knowledge flow network is in itself a kind of organizational knowledge that is more relevant to the roles of the organization and less focused on individuals. It is more concerned with the content of knowledge and with effective knowledge sharing in distributed cooperative teams especially the agent-based virtual organization. Knowledge content can be layered as concepts, axioms, rules and methods, can be classified according to its application domains, and can be organized in a knowledge space of three dimensions: the knowledge category dimension, the knowledge level dimension, and the knowledge location dimension.

Epistemology plays the key role in the process of generating knowledge. In the context of knowledge sharing and the future Web, we are aiming at a kind of semantic description and generation mechanism that reflects human cognition. Different people may see a different epistemology in the same object or event. Epistemological mechanisms help humans and agents understand, generate and describe new knowledge when they share resources. An easy way to implement an epistemological mechanism in the current Web would be to develop an epistemological appendage, generated and used in conjunction with the original resources.

Some researchers are interested in mining Web usage logs. These logs record data about the Web’s users. The researchers are striving to use the results of mining Web usage logs to support personalized Web services. However, the Web logs only reflect a small portion of users’ behavior, and they are unable to capture users’ actual intentions, thinking or understanding.

A personalized epistemological appendage together with a domain ontology could improve the current Web's keyword-based approaches and ontology-only approaches. The form of the appendage could be a kind of semantic representation like the semantic link network, which will be introduced later.

1.8 Definition, Characteristics and Strategies of the Knowledge Grid

1.8.1 Definition

We first explain what is *not* the major concern of the Knowledge Grid. Traditional natural language processing, recognition of human speech and handwriting, and formal semantics are not the major concern of the Knowledge Grid. Security and scientific computing are not its key issues. The Knowledge Grid will go beyond the traditional and will look to improved information retrieval, filtering, mining and question answering techniques.

The Knowledge Grid is an intelligent and sustainable Internet application environment that enables people or virtual roles (mechanisms that facilitate interoperation among users, applications, and resources) to effectively capture, coordinate, publish, understand, share and manage knowledge resources. It provides on-demand and robust services to support innovation, cooperative teamwork, problem solving and decision making in a distributed environment, which varies in scale and stability. It incorporates epistemology and ontology to reflect human cognition characteristics; exploits social, ecological and economic principles; and adopts the techniques and standards developed during work toward the future web (H. Zhuge, "China's e-Science Knowledge Grid Environment", *IEEE Intelligent Systems*, 2004, vol.19, no.1, pp.13-17).

Efforts towards the future Web supply the Knowledge Grid with several candidate techniques and implementation platforms. *The Grid is not the only platform for realizing the Knowledge Grid, and the Knowledge Grid should absorb the ideal and some ideas of the Grid.*

The meaning of the term *grid* in the Knowledge Grid is broader than it is in Grid computing. Actually, people have a long history of using the word “grid” in drawing and mapping, geodetic surveying, and mathematics. The word was borrowed from the power grid to refer to clustered computing power when the concept of Grid computing appeared in 1995.

Fig. 1.3 depicts the relationships among the Internet, the Grid, the Semantic Web, the Semantic Grid, and the Knowledge Grid. We can see that the Semantic Grid is the direct basis of the Knowledge Grid.

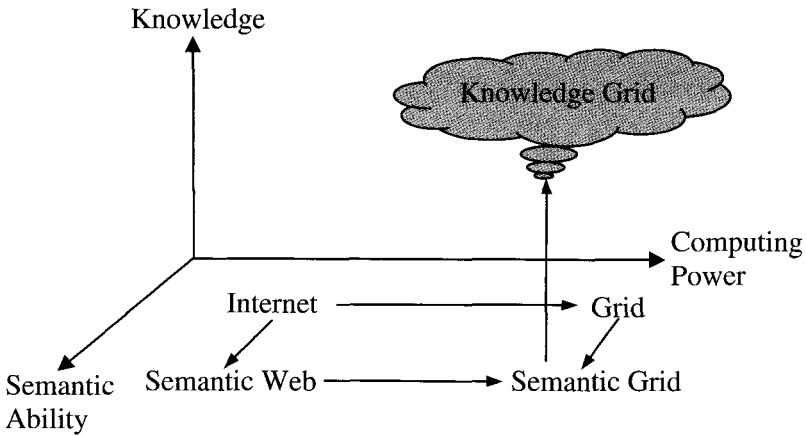


Fig.1.3 Relationships between the Knowledge Grid, the Semantic Grid, the Semantic Web, the Grid and the Internet.

In the long run, a Knowledge Grid should be established on the *future interconnection environment*, which should be a large-scale, autonomous, living, sustainable and intelligent network where society and nature can develop together, functioning and evolving cooperatively. It would collect useful resources from the environment, transform and organize them into semantically rich forms that could be used easily by both machine and human. Geographically dispersed people and resources could work together to accomplish tasks and solve problems by

using the network to actively promote the flow of material, energy, techniques, information, knowledge and services through roles and machines, improving both the natural and the artificial environment.

Intelligence, Grid, peer-to-peer and environment represent humanity's four aspirations for the future working and living environment. The intelligence reflects humanity's pursuit of recognizing themselves and the society. The Grid reflects humanity's pursuit of optimization and system. The peer-to-peer reflects humanity's pursuit of freedom and equality. The environment reflects humanity's pursuit of understanding of nature and its harmony.

1.8.2 Parameters

As a large-scale dynamic human-machine system, a Knowledge Grid environment will be characterized by five parameters:

1. *Space* — the capacity to hold a great variety of both individual and shared knowledge resources.
2. *Time* — the arrow of evolution.
3. *Structure* — the construction of the environment and resources in the environment.
4. *Relation* — relationships among parameters and among resources.
5. *Measurement* — the evaluation of the status of, and the prospects for, resources, processes and their relationships.

Einstein's general theory of relativity reveals the relationship between space and time in physical world: space and time are malleable entities. On the largest scales, space is naturally dynamic, expanding or contracting over time.

A Knowledge Grid environment will foster the growth of knowledge by supporting social activities at different levels (from the simple physical level to the complex human-machine community level) and in different disciplines. As a product of society in nature, it evolves and endures throughout the life of the race rather than the life of an individual.

Human social activities generate and develop the semantics of natural languages. Human-machine social activities will need to be based on a kind of human-machine semantics, to establish an “understanding” between inanimate resources and humans. Such semantics will be needed so that the services and knowledge of future machines can be beneficially used and protectively regulated by humans.

1.8.3 *Distinctive characteristics of the Knowledge Grid*

- (1) *Single semantic entry point access to worldwide knowledge.* In the Knowledge Grid environment, people could access knowledge distributed around the world from a single semantic access entry point without needing to know where the required knowledge is.
- (2) *Intelligently clustered, fused and distributed knowledge.* In the Knowledge Grid environment, related knowledge distributed around the world could intelligently cluster together and fuse to provide appropriate on-demand knowledge services with underlying reasoning and explanation. So knowledge providers should include meta-knowledge (knowledge about how to use knowledge), and could use a kind of uniform resource model to encapsulate the provided knowledge and meta-knowledge to realize active and clustered knowledge services.
- (3) *Single semantic image.* The Knowledge Grid environment could enable people to share knowledge and to enjoy reasoning services in a single semantic space where there are no barriers to mutual understanding and pervasive knowledge sharing.
- (4) *Worldwide complete knowledge service.* The Knowledge Grid could gather knowledge from all regions of the world and provide succinct and complete knowledge relevant to the solution of particular problems. To achieve this goal, we need to create a new knowledge organization model.
- (5) *Dynamic evolution of knowledge.* In the Knowledge Grid environment, knowledge would not be just statically stored, but would evolve to keep up-to-date.

1.8.4 The Knowledge Grid's general research issues

- (1) *Theories, models, methods and mechanisms for supporting knowledge capture and representation.* The Knowledge Grid should be able to help people or virtual roles effectively capture, and conveniently publish knowledge in a machine-processable form that could directly, or after simple transformation, be understood by humans. We should build an open set of semantic primitives to help knowledge representation. These primitives should be able to represent multi-granular knowledge. The capture of knowledge here has two meanings: one is when people learn from each other directly, or from the resources published by others, and then publish new knowledge on the Knowledge Grid; the other is when the Knowledge Grid gets knowledge from numeric, textual or image resources by mining, induction, analogy, deduction, synthesizing, and so on.
- (2) *Knowledge display and creation.* These come mainly through an intelligent user interface (for example, a semantic or knowledge browser) that enables people to share knowledge with each other in a visual way. The semantic link network and the cognitive map are two ways to depict knowledge. The interface should implement the distinctive characteristics of the Knowledge Grid and be able to inspire people's discovery of knowledge through analogy and induction.
- (3) *Propagation and management of knowledge within virtual organizations.* This could eliminate redundant communication between team members to achieve effective knowledge management in a cooperative virtual team. Knowledge flow management is a way to achieve knowledge sharing in a virtual team.
- (4) *Knowledge organization, evaluation, refinement and derivation.* Knowledge should be organized normally to obtain high retrieval efficiency and ensure the correctness of operations. The Knowledge Grid should be able to eliminate redundant knowledge and refine knowledge so that useful knowledge can be increased. It can also

derive new knowledge from existing well-represented knowledge, from case histories, and from raw knowledge material like text.

- (5) *Knowledge integration.* Integrating knowledge resources at different levels and in different domains could support cross-domain analogies, problem solving, and scientific discovery.
- (6) *Abstraction.* It is a challenge to automatically capture semantics from a variety of resources, to make abstractions, and to reason and explain in a uniform semantic space. The semantic constraints and rules of abstraction ensure the validity of resource usage at the semantic level.
- (7) *Scalable network platform.* The Knowledge Grid should enable a user, a machine or a local network to freely join in and leave without affecting its performance and services. It is a challenging task to organize and integrate knowledge within a dynamic network platform.

1.8.5 Differences between the Web and the Knowledge Grid

Here we use examples to make a brief comparison between the Web and the Knowledge Grid. With the current Web, people with an illness can use search engines to retrieve relevant medical information, and browse hospital or health websites to find suitable hospitals and doctors, depending on what URLs they can remember. As there are more than 82 million health websites, ill people are often annoyed by the large amount of useless information in a search result and by the time consumed in browsing through many websites. They may be further confused by the various opinions of different doctors. They may also worry about whether the result of their searching is based on up-to-date knowledge. Entering related symptoms, they can usually only obtain results for separate symptoms. Further, the whole searching process may overlook some experts, especially those who specialize in uncommon diseases.

Such searches will be improved by the Knowledge Grid, which can accurately and completely locate all relevant knowledge, cluster and synthesize the search result and then actively present it to ill people according to their illness profiles. The ill people can get an explanation

of the search result with underlying reasoning based on the clustered knowledge. The relationship between the symptoms of a disease will be considered during reasoning. The search results can adapt to change in illness profiles. Knowledge provided by different doctors worldwide will be refined, checked and evaluated as to usefulness, consistency and time-effectiveness.

In the Knowledge Grid, new knowledge can be derived from: existing knowledge, patients' feedback, and mining in medical textbooks, papers and other related sources. Ill people can also choose to provide symptoms of their disease through a single semantic entry point when accessing the Knowledge Grid to obtain instant consulting service. The result may include several candidate treatments selected by considering such factors as cost, waiting time, skill level, transportation, and so on.

Similar advantages of the Knowledge Grid also exist in scientific research, business, education and other application domains.

1.8.6 The technological basis of the Knowledge Grid

The Knowledge Grid is not pie in the sky. It is based on existing methods and technologies such as the Grid, the Semantic Web, Web Services, Peer-to-Peer, AI, proper data modeling, information processing technologies (for example, data and text mining, information filtering, extraction, fusion and retrieval), and system methodology as shown in Fig. 1.4.

The adoption of a new system methodology, a new organization model, a new computing model, and the principles of relevant disciplines will further challenge current software methodology. The implementation of the ideal Knowledge Grid requires a new software methodology that can cope with evolution, fuse resources, and support competition and sustainable development.

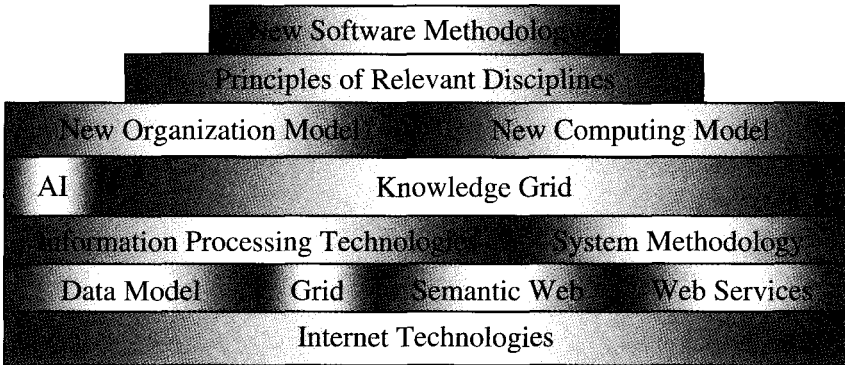


Fig. 1.4 The technologies of the Knowledge Grid.

1.8.7 *The dream and the strategy*

James Gray summed up computing history in his 1999 Turing Award Lecture: the dream of Charles Babbage (1791-1871, the father of the computer) has been largely realized, and the dream of Vannevar Bush has almost become reality, but it is still difficult for computer systems to pass the Turing Test — computing systems still do not have human intelligence although significant progress has been made. Gray extended Babbage's dream: computers should be highly secure and available, and they should be able to program, manage, and replicate themselves.

Scientists have made significant progress towards establishing highly secure and available systems — the goal of the Grid. But so far, we are still far from the goal of self-programming, self-managing and self-replicating. Gray extended Bush's Memex vision to an ideal that automatically organizes indexes, digests, evaluates, and summarizes information, and indeed scientists in the information processing area are making efforts towards this goal. He proposed three more Turing Tests: prosthetic hearing, speech, and vision (*Journal of the ACM*, January 2003, vol.50, no.1, pp.41-57).

What modern society needs from the future computing environment has gone far beyond the scope of the Turing Test and other automatic

machine intelligence problems such as self-programming. Computing has evolved from mainframe computers to personal computers, to locally networked computers, and to the Internet. People now primarily use computers interactively on a large scale, so that the *dynamics, evolution, cooperation, fusion, sustainability, and social effects* of computer use have become major concerns.

Fig. 1.5 shows the evolution of the computing environment. The Knowledge Grid is the platform that will support the large-scale human-computer environment.

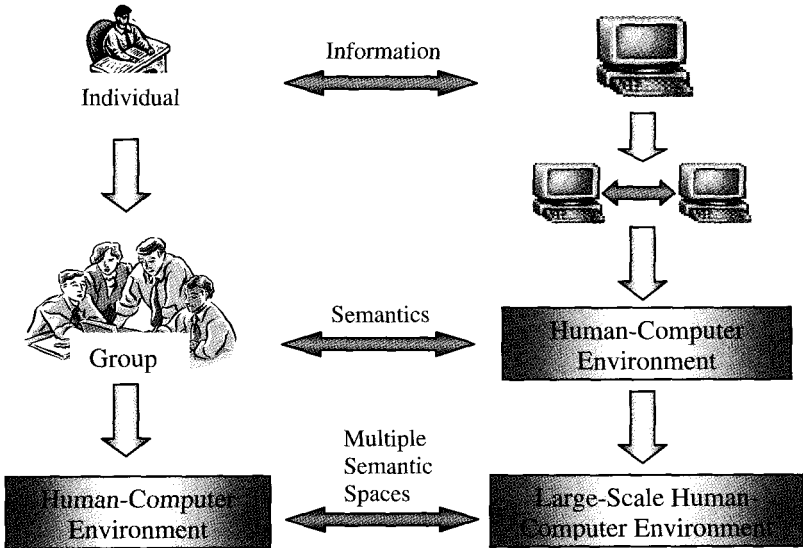


Fig. 1.5 The evolution of the computing environment.

Communication in human society is carried out in multiple semantic spaces, such as the emotional, cultural, artistic, scientific, and that of daily life, which establish the basis for mutual understanding. Loosely or tightly coupled rules could be used to coordinate between these semantic

spaces. Traditional research on natural language processing only focuses on one semantic space — the text space, where some information has inevitably been lost in the writing process. This is one reason why different people have different understandings of the same text. So it is impossible to realize the dream of automatically processing natural language if only text analysis is used.

We need an incremental strategy to develop the Knowledge Grid. A worldwide Knowledge Grid is a long-term target. A preliminary stage developing a medium-sized Knowledge Grid based on an institution's intranet would be an appropriate step in the long march towards the long-term target. It could support more effective knowledge management within institutions of various kinds. *Institutional Knowledge Grids* could then become components of the worldwide Knowledge Grid.

A *Micro Knowledge Grid* could be the basic component of a medium-sized Knowledge Grid and thus the basic component of the worldwide Knowledge Grid. It would be useful in helping individual knowledge management — managing raw knowledge (expressed in natural language text) and refined knowledge (expressed in a knowledge representation code), and transforming tacit knowledge into explicit knowledge. *However, a worldwide Knowledge Grid should be more powerful than the sum of its components.*

A Knowledge Grid should support more semantic spaces than just one text space. Knowledge sharing in a Knowledge Grid depends on a correct understanding of the semantics of its resources. But these semantics are not the same as traditional formal semantics. These should be a kind of informal computable semantics, which supports computing, reasoning, abstraction, integration and transformation between semantic spaces. The semantics of the Knowledge Grid should be easily understood by humans and readily processed by machines.

The internal structure of resources that express the same semantics could be completely different. So finding effective ways (for example, markup languages) to express the internal structure will be much more complex than finding ways to express the external semantics — the semantic relationships between resources. Since no object in the world exists in isolation, the semantics of a resource could be determined or

roughly reflected by the semantics of the resources related to it and the relationships between them.

Symbolic systems are elegant computing models. Artificial intelligence was regarded as evidenced by the behavior of working symbolic systems. However, symbolic systems have their own particular scope of ability. Non-symbolic systems also have their particular scope of ability. The Knowledge Grid should combine the approach of symbolic systems with the approach of non-symbolic systems (this still needs much work).

Billions of years of natural evolution have created a natural environment and an intelligent species that has evolved into human society.

If we draw an analogy between the future interconnection environment and the world of nature, a challenging question arises: what is the field theory of the interconnection environment? As its basic material, the various resources in the interconnection environment exist in a special field, where resources flow from higher intensity nodes to lower intensity nodes. But the duplication and generation of resources does not cause the loss of any other resource, and the flow of resources also does not mean the loss of any resource. This means that the law of energy conservation in the physical world does not hold in the abstract world of knowledge. The laws and principles in this special field will become the basic theory of the future interconnection environment.

If we draw an analogy between the future interconnection environment and human society, a challenging question arises: what is the market and what the price of services in the future interconnection environment?

If we draw an analogy between the future interconnection environment and the human body, some further challenging questions arise:

- (1) What is the circulatory system of the future interconnection environment? This question impels us to investigate and establish a flow model for information, knowledge and services.

- (2) What is the immune system of the future interconnection environment? This question impels us to investigate the principles of resource clustering and security.
- (3) What is the digestive system? This question impels us to investigate the principles of generation and understanding of resources.
- (4) What is the nervous system? This question impels us to investigate the principles of the control flow within the future interconnection environment.
- (5) What is the ecology? What are the rules of evolution? Will the environment degrade and its species diversity decrease? To answer these questions requires us to carry out research relating ecology to the future interconnection environment.
- (6) What are the sustainable development principles and the rules of the future interconnection environment that could evolve harmoniously with human society? This question impels us to carry out research relating the following areas: our society, its economy and the future interconnection environment.

So far, we can assert that both the notion and the ideal of the Knowledge Grid are understandable, useful and challenging. The methodology of the Knowledge Grid should also include the testable and incremental aspects that James Gray mentioned in his Turing Award Lecture.

The incremental aspect would make our short-term target modest. We believe that the major characteristics of the Knowledge Grid will be realized by a medium-sized Knowledge Grid in some application area, such as e-science and e-government, based on the current Web and Grid technologies.

As for the testable aspect, we can use the following basic criteria to evaluate whether it is a Knowledge Grid or a knowledge based system within the prior art:

- (1) The effectiveness of knowledge/information services, for example, the response time;

- (2) The quality of knowledge/information services, for example, users' degree of satisfaction with the use and the content of services; and
- (3) The improvement of services, for example, whether the services can be improved during use.

On average, a Knowledge Grid should perform better than other systems in 70% of tests.

The following chapters concern detailed research and practice — some preliminary attempts to fulfill the ideal of the Knowledge Grid. The methodology of the Knowledge Grid is guiding the China Knowledge Grid Research Group (<http://kg.ict.ac.cn>) in its continuing research and development.