THE KNOWLEDGE GRID Second Edition

Toward Cyber-Physical Society

The Knowledge Grid is an intelligent and sustainable interaction environment that consists of autonomous individuals, self-organized semantic communities, adaptive networking mechanisms, evolving semantic link networks keeping meaningful connection between individuals, flows for dynamic resource sharing, and mechanisms supporting effective resource management and providing appropriate knowledge services for learning, innovation, teamwork, problem solving, and decision making. This book presents its methodology, theory, models and applications systematically for the first time.

Its second edition fulfils the ideal of the Knowledge Grid by including many new contents, including: (1) The insight of cyber-physical society; (2) The systematic method of semantic link network that supports uncertainty management, discovery of semantic links and semantic communities, autonomous semantic data model, and cyber-physical-socio semantic link network; (3) Semantic peer-to-peer infrastructures for efficient knowledge sharing; and, (4) A new centrality measure of network. This new edition will undoubtedly provide inspiring materials for researchers, academics, practitioners and students.

World Scientific
www.worldscientific.com
7567 hc

ISBN-13 978-981-4291-77-4
ISBN-10 981-4291-77-3
9 789814 291774

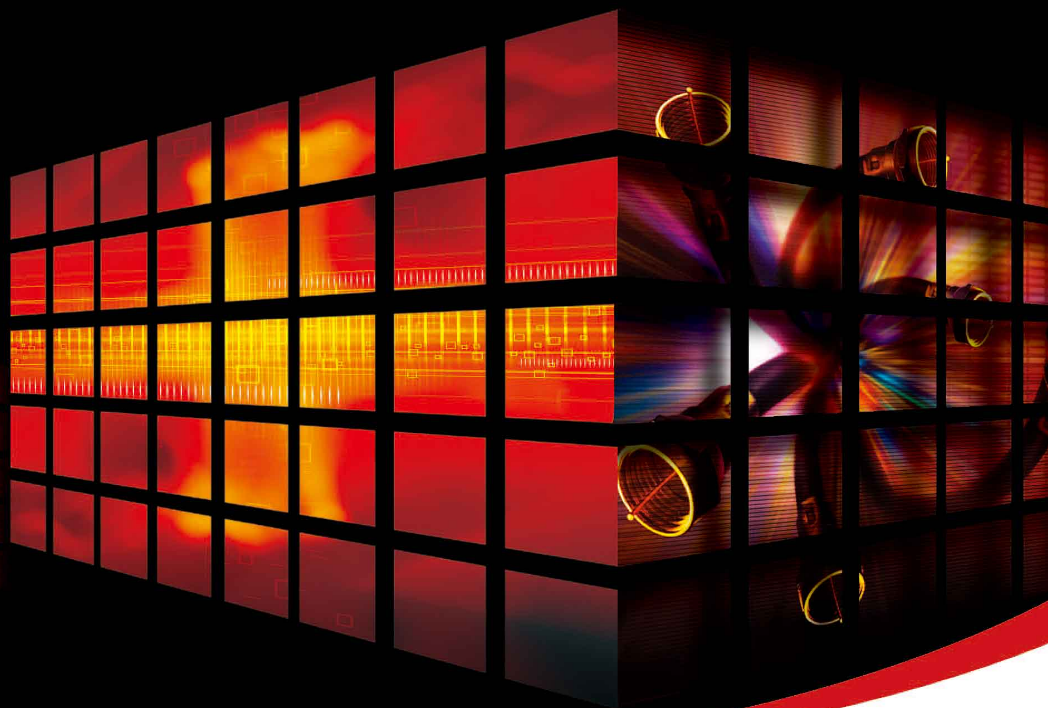
THE KNOWLEDGE GRID Second Edition

Zhuge

THE KNOWLEDGE GRID Second Edition

Toward Cyber-Physical Society

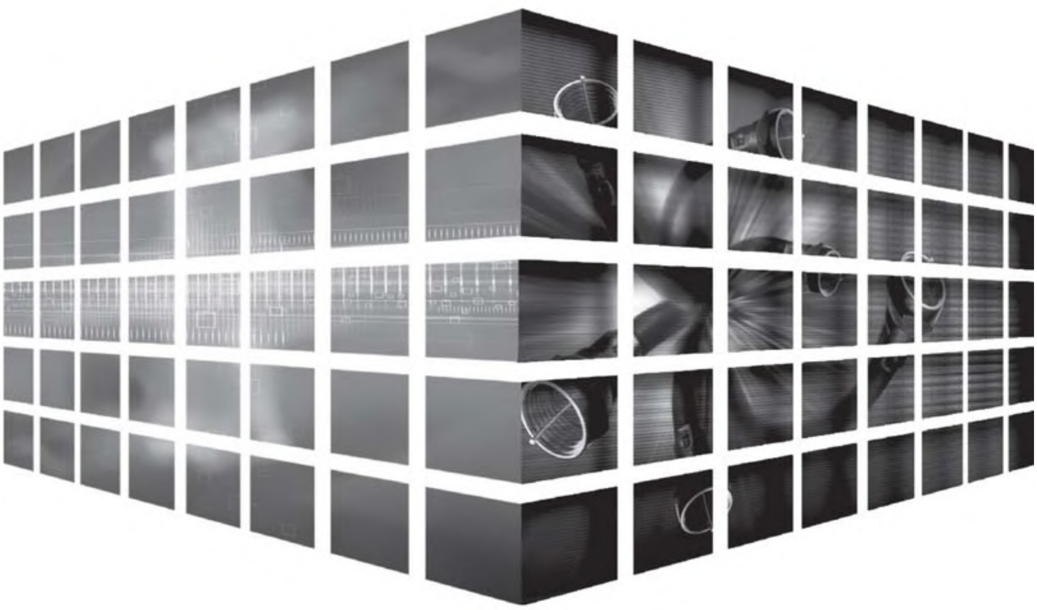
Hai Zhuge



World Scientific

The Knowledge Grid
Second Edition

Toward Cyber-Physical Society



Hai Zhuge
Chinese Academy of Sciences, China

World Scienti ic

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.

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

Chapter 2

The Semantic Link Network

Humans live with networks of versatile relations. What are the essential characteristics of relational networking? What are the laws, methods and models for relational networking? How do the relational networks emerge semantics? How to make use of the relational networks to support basic intelligence? One of the major tasks of science is to unveil various relations.

2.1 The Idea of Mapping

The idea of geographical mapping helps in exploring the Knowledge Grid environment. A geographical map is a pictorial abstraction of a region's physical or social properties. People can use certain features of such a map to accurately and efficiently locate or identify places or regions.

(1) *Coordinate location*. Lines are drawn on maps to link points of equal latitude (distance from the equator) and points of equal longitude (angle from a standard straight line joining the two poles, usually the Greenwich meridian). These enable people to locate a surface feature on a map if they know its coordinates—the latitude and longitude.

(2) *Referential location*. People can locate a surface feature on a map even if they don't know its coordinates, provided they know its distance and direction from a known point or feature on the map. For example, a town about twenty kilometers to the west of the center of Beijing can be found indirectly on a map by

referring to the location of the center of Beijing.

(3) *Regional partition*. Maps often have lines drawn on them to mark the boundaries of different social or climatic regions. Where there is a hierarchy of regions, different kinds of lines are used to mark the boundaries at different levels of the hierarchy. Knowing what region, such as country or district, a surface feature is in can make it easier to find on a map.

(4) *Color*. When regions are marked on a map, neighboring regions are usually filled in with contrasting colors. This makes it much easier to see the extent of any particular region of search.

(5) *Overlays*. To map information of different kinds, such as traffic, climate, and ecological data, onto the same area, transparent overlays with boundary markings are sometimes used. The distribution of properties such as traffic and population density, rainfall and temperature, and cropping, can then be conveniently compared.

(6) *Legends*. A legend is used to explain aspects of a map, such as the name, scale, and meaning of the symbols and labeling.

If we can create a *semantic map for the Web*, users will be able to use the map to effectively browse its contents. If we can create a *semantic map for the society and a knowledge map for thoughts*, it would bring us much closer to the ideal Knowledge Grid.

Location by coordinates, the most important aid in geographical mapping, can be adapted to Web mapping by using classification. This approach requires Web page providers to encode categories when adding resources. A well-defined classification scheme will help both providers and users.

Referential location of a kind has been the purpose of Web hyperlinks from their beginning, as they enable users to browse from page to page. However, they do not entail semantic relationships, and we cannot accurately locate a resource by just referring to a known resource and giving the semantic relationship.

Regional partition of the Web is provided by the names in the URL (Uniform Resource Locator) of each of its pages. Each URL identifies exactly where in the Web a resource is stored, much like a postal address.

Color can help distinguish classifications. So far, the use of color on the Web is quite arbitrary. There are no regulations for using color to convey semantic properties although national preference in selecting color has been studied. Building a semantic image as a map for the Web is a way to smart Web applications.

A map provides people the context and tool for efficiently finding a location relevant to people. Multi-layer and multi-facet information embedded in a location enable on-demand information services through the cyber space with mobile devices.

The initial motivation of creating the semantic link network model and method (in short SLN) is to reflect the complex systems as a semantic image like the map, to analyze the basic nature of the systems, and to support some basic intelligence.

The idea of map implicates an approach to modeling the functions in the cyber space, physical space, social space, and mental space.

2.2 Basic Concepts and Characteristics

Humans wave semantic link networks consciously and subconsciously in life-time and through generations, and act intelligently with knowing a part of the network.

Fig.2.2.1 depicts a semantic link network formed by the following scientific activities: reading, writing, citing, discussing, submitting, publishing, conferencing, and collaborating.

Semantic Link Network has the following characteristics:

(1) *Dynamicity*. It keeps evolving with continuous interactions and reasoning. It cannot be simply regarded as a static graph.
(2) *Rules*. It contains social rules for regulating the evolution of the network and the reasoning rules for deriving links. For example,

students could not be the supervisor of student, and professor could not be the supervisor of professor.

(3) *Openness*. An individual can link to the network according to the rules. This is different from the Web, which enables any page to link to any other page.
(4) *Self-organization*. There is generally no central control on the organization and evolution of the network except the consensus of the rules.
(5) *Complex reasoning*. New semantic links may be derived from the existing network and rules through various reasoning, including relational, deductive, inductive, and analogical reasoning.
(6) *Order sensitive*. The effects of different orders of operations may be different.
(7) *Support basic intelligence*. Individuals can behave intelligently with knowing a part of the network, even though only a very small part of the network.
(8) *Locality and global influence*. The influence of operation is local in short-term, but influence may be global in long-term.
(9) *Multiple spaces*. While experiencing in the physical space and the cyber space, humans wave semantic link networks in the social space, cyber space, and their mental spaces. So, semantic links pass through the cyber space, physical space, social space, and mental space.

Previous research on complex networks and social networks neglect the networking semantics (R. Albert, H. Jeong, A.-L. Barabási, Diameter of the world wide web, *Nature*, 401 (1999) 130-131; M. E. J. Newman, Coauthorship networks and patterns of scientific collaboration, *PNAS*, 101 (2004) 5200–5205; M. E. J. Newman and J. Park, Why social networks are different from other types of networks, *Phys. Rev. E*, 68(2003)036122; J. Balthrop, et al., Technological networks and the spread of computer viruses, *Science*, 304(2004)527–529; B. Karrer and M. E. J. Newman, Competing epidemics on complex networks, *Phys. Rev. E*, 84(2011)036106).

Relationship is a key component of the semantic link network.

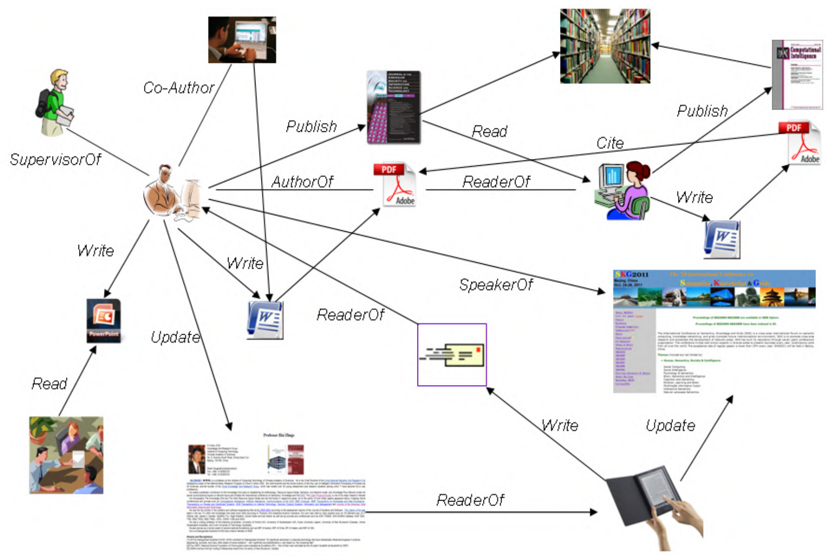


Fig. 2.2.1 Humans have been waving Semantic Link Networks in life-time and through generations.

The following issues come from the semantic link network:

(1) How to model the dynamic network?
(2) What are its motion principles?
(3) How to discover the implicit relations？
(4) What are its networking effects？
(5) What is the semantics emerging with the motion of the network?
(6) How does knowledge flow through the network?
(7) What are the principles and methods for effectively sharing knowledge through the network?

Symbiosis is a relation between two individuals or species that benefit each other. It is a basic relation for organizing individuals or species in the biological world.

Blood relation links one person to another by birth rather than by marriage. It is a basic relation of forming a society.

Kinship is a relationship between persons that share a genealogical origin. In anthropology, it includes persons related both by descent and marriage. The kinship relation through marriage is commonly called *affinity* in contrast to *descent*. Kinship is one of the most basic relations for organizing individuals.

From the network shown in Fig.2.2.2, humans know that it is about family relationships and can infer some characteristics of nodes (e.g., gender), but machines are not able to understand it like humans. Why? One reason is that humans share a semantic space in minds. If a *semantic space* can be assigned to machines to regulate the semantics of the network, they should know more about the network.

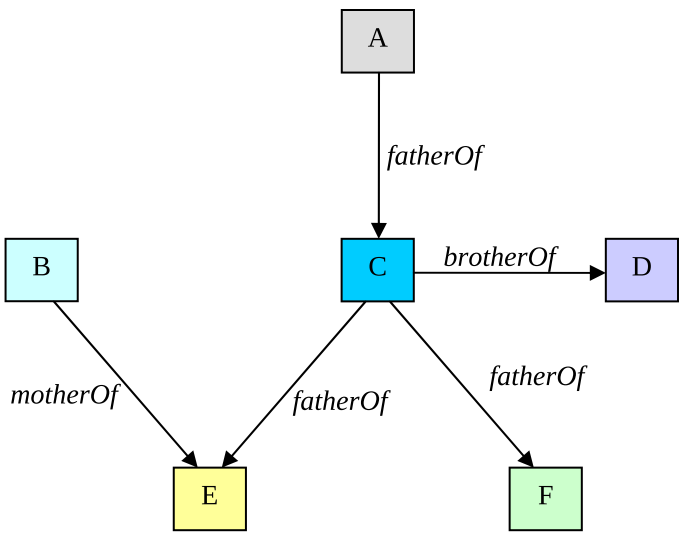


Fig. 2.2.2 A simple network of family relationship.

Definition 2.2.1 A *semantic link network* (SLN) is a relational network consisting of the following main parts: a set of *semantic nodes*, a set of *semantic links* between the nodes, and a *semantic space*. Semantic nodes can be anything. The semantic link between nodes is regulated by the attributes of nodes or generated by interactions between nodes. The semantic space includes a classification hierarchy of concepts and a set of *rules* for reasoning and inferring semantic links, for influence nodes and links, for networking, and for evolving the network.

A semantic link network can be denoted as $SLN=<N, L, \Im, f>$, where $\Im$ *is the semantic space* consisting of a concept hierarchy $\wp$ and a rule set $\Re$, and f is a mapping from $\{N, L\}$ into $\Im$. Sometimes, we simply use $SLN=<N, L, \Re>$ to denote a semantic link network in a given domain.

The following are further explanations:

(1) A semantic node can be anything, for example, text and image, individual (human or agent), event, concept, class, or even an SLN. A semantic node has attributes that reflect features in one or more spaces (e.g., the physical space and physiological space). The attributes of a node render its class.

(2) A semantic link indicates the relation between semantic nodes. It can be indicated by a relation indicator (a certain form of symbol) or a combination of relational indicator according to predefined *lightweight grammar*. The semantics of the indicator is regulated by the semantic space. Some semantic links are determined by the attributes of semantic nodes, while others are determined by direct or indirect interaction between semantic nodes.

(3) The rules regulate semantic nodes and semantic links. Implicit semantic links may be derived from the rules and semantic links. For example, *brotherOf* relation between nodes E and F can be derived from the *fatherOf* relation between nodes C and E and the *fatherOf* relation between nodes C and F. Semantic nodes may also influence each other through the rules.

(4) The classification hierarchy describes the subclass and super-class relations between concepts. For example, *fatherOf*, *brotherOf* and *motherOf* are subclases of *family relationship*.

The type and number of semantic nodes, semantic links and rules vary with applications. With the rule set varies between small and large, reasoning changes between light and heavy. The rule set can be deteriorated to some restricts on network, e.g., the influence between nodes' attributes.

The co-occurrence relationship is the most basic relationship of social events. In many cases, the relations between two co-occur events are unknown, we say that they are linked by co-occurrence relation. The co-occurrence relation can be specialized into many relations with deepening the understanding of the two events.

The hyperlink Web can be regarded as the degradation of the SLN as hyperlinks are only connections between Web pages to allow for Web browsing. The hyperlink can also be regarded as a co-occurrence relation because the two linked pages are often browsed one after another.

Adding semantic space and semantic indicators to the hyperlink network may gain the following advantages:

(1) *Semantics-based resource organization and retrieval*. Simple hyperlinking as in the current Web basically supports ad hoc retrieval, and provide the ranking basis for various search engines. Neither approach is well suited to semantic retrieval. By adding semantic qualification, the SLN can make retrieval more effective.
(2) *Semantics-based reasoning and browsing*. The semantic relationships encoded in the links of SLNs support coarse relational reasoning, which enables intelligent browsing and retrieval.
(3) *Semantic overlaying*. An SLN can be a semantic overlay for the Web, which provides the context for intelligent applications.

Fig. 2.2.3 is an example of using semantic links to cluster images. Tags are words attached to images by Web users. An image can have multiple tags given by different users. The semantics of tags can be indicated by clustering tags and statistic analysis on the usages of tags (P.A. Chirita, et al., P-TAG: large scale automatic generation of personalized annotation tags for the web, *WWW07*).

Images can be clustered according to the relations between the tags. One way to find the relations between tags is to find the co-occurrence relations between the words in commonly interested Web pages. A distance between words can be defined according to some criteria. For example, the distance between two words occurred in one sentence can be defined as shorter than the distance between the words occurred in different sentences. If two words often co-occur in the same sentence, a co-occurrence link is possibly between them. For example, a co-occurrence link is possibly between "boat" and "river". To reflect the uncertainty, a probabilistic semantic link is necessary.

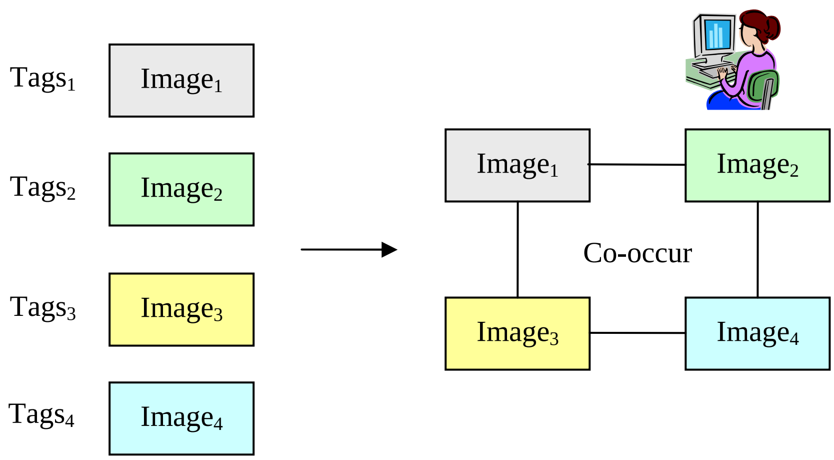


Fig. 2.2.3. Using semantic links to cluster online images.

Arbitrarily annotating semantic relations will make the network hard to be understood, but formal computational semantics is usually unacceptable by users and actually not computational (Y. Wilks, Computational Semantics Requires Computation, *FLAIRS Conference* 2011). Statistical approaches are suitable for machine processing, but abstraction plays a more important role in forming and understanding semantics (J.B. Tenenbaum, et al., How to grow a mind: Statistics, structure, and abstraction, *Science*, 331 (6022)(2001) 1279-1285; J. Heinz and W. Idsardi, Sentence and Word Complexity, *Science*, 15(2011)295-297).

One factor of the success of the Web is its simplicity in linkage and usage mode.

How to appropriately indicate the semantics of relations?

Because of the domination of natural language used in communication, humans have to rely on words as semantic indicators because the semantics of words have been regulated in commonsense. For machine understanding, a light-weight language is needed to compose the indicators.

The light-weight language consists of a set of keywords, a simple grammar regulating the relations between keywords, and the relations between keywords, instances and classes. As an abstract concept, a class includes attributes and functions.

In applications, *a single or the ordered words is often used to help users compose semantic indicators* in such ways as "*causeEffect*", "*Cite*", "*IsPartOf*" and "*IsA*" to links. Of course, operations can be added on words if needed.

Humans have established much commonsense in using natural languages. The Web provides an open platform for humans to contribute and share commonsense. So far, many commonly used relations are available online, for example, in Wikipedia, and the relations evolve with the enrichment of these online resources. The indicator set can help users to annotate semantic links. The online sources like Wikipedia and ODP (Open Directory Project) provide context for explaining the relations.

It is feasible to construct an SLN model that can support light-weight semantics-aware retrieval. The Resource Description Framework (RDF, www.w3.org/RDF/) plays the similar role in technique. RDF can be one implementation solution of applying SLN to Web applications. But, SLN research concerns dynamicity and generality, and it is at more abstraction level and methodology level.

Why are SLNs relevant to the Knowledge Grid?

Semantics is the basis of understanding and sharing knowledge. An SLN embodies a kind of coarse semantics and it has some of the characteristics of the Web. Given a formal structure and a well developed method and theory, an SLN can be used as a semantic interconnection overlay of the Knowledge Grid. However, the Knowledge Grid should have multiple semantic overlays for different semantic scales.

With the development of society and information technology, humans become increasingly relying on various explicit or implicit relationships to live and work. The key issue of many applications, from general Web search to specific domain applications like geographical information retrieval, is how to accurately locate the necessary resources and relationships in large and complex SLNs.

A semantic link can be represented as X—α→Y in simple, and a relation "—α→" is called an α-link or semantic link. The co-occurrence can be denoted as X——Y in simple.

The following are some concepts on semantic links:

(1) If there is a semantic link chain from X to Y, we say that Y is *semantically reachable* from X.
(2) If X—α→$Y \Rightarrow X$—β→Y, then we say that α implies β, denoted by $f(\alpha)\subseteq f(\beta)$, or $\alpha\subseteq\beta$ in simple, i.e., the semantic image of α is semantically included by (or, is a subclass of) the semantic image of β. For example, X—*SonOf*→$Y \Rightarrow X$—*OffspringOf*→Y if *SonOf* is the subclass of *OffspringOf* in a class hierarchy.

(3) If we have X—α→Y, Y—α→$Z \Rightarrow X$—α→Z (where $\Rightarrow$ stands for implication), we say that α or the α-link is transitive.
(4) We say that X is semantically equivalent to Y, denoted by X—*equal*—Y, or $X = Y$, if they can substitute for each other wherever they occur.

The following is a set of general semantic links:

(1) The *cause-effect link*, denoted by *ce* as in r—*ce*→r', for which the predecessor is a cause of its successor, and the successor is an effect of its predecessor. The cause-effect link is transitive, that is, r—*ce*→r', r'—*ce*→$r'' \Rightarrow r$—*ce*→r''. Cause-effect reasoning can chain along cause-effect links because of this transitivity.
(2) The *implication link*, denoted by *imp* as in r—*imp*→r', for which the semantics of the predecessor implies that of its successor. The implication link is transitive, that is, r—*imp*→r', r'—*imp*→$r'' \Rightarrow$ r—*imp*→r''. Implication links can help a reasoning process find the relationships between documents.
(3) The *subtype link*, denoted by *stOf* as in r—*stOf*→r', for which the successor reserves all the features of its predecessor. The subtype link is transitive, like set inclusion, that is, r—*stOf*→r', r'—*stOf*→$r'' \Rightarrow r$—*stOf*→r''.
(4) The *similar link*, denoted by (sim, sd) as in r—(sim, sd)→r', for which the semantics of the successor is similar to that of its predecessor, and where *sd* is the degree of the similarity between r and r'. Like the partial-inheritance relationship (H. Zhuge, Inheritance rules for flexible model retrieval, *Decision Support Systems*, 22(4)(1998)379-390), the similar link is intransitive.
(5) The *instance link*, denoted by *insOf* as in r—*insOf*→r', for which the successor is an instance of the predecessor.
(6) The *sequential link*, denoted by *seq* as in r—*seq*→r', which requires that r be browsed before r'. In other words, the content of r' is a successor of the content of r. The sequential link is transitive, that is, r—*seq*→r', r'—*seq*→$r'' \Rightarrow r$—*seq*→r''. The transitivity enables relevant links to be connected in a sequential chain.

(7) The *reference link*, denoted by *ref* as in r—ref→r', for which r' is a further explanation of r. The reference link is transitive, that is, r—ref→r', r'—ref→$r'' \Rightarrow r$—ref→r''. However, partial reference like the citation link is intransitive.

(8) The *equal link*, denoted by e as in r—e→r', for which r and r' are identical in meaning. Clearly, any resource is equal to itself.

(9) The *empty link*, denoted by ϕ as in r—ϕ→r', for which r and r' are completely irrelevant to each other.

(10) The *null or unknown link*, denoted by *Null* or N as in r—N→r', for which the relation between two resources is unknown or uncertain. The *Null* relation means that there might be a relationship, but we do not yet know what it is. A null link can be replaced by a provider or by a reasoning process.

(11) The *non-α relation*, denoted by $Non(\alpha)$ or α^N as in r—α^N→r', for which there is no α relationship between r and r'. It is sometimes useful in the reasoning process to know that a particular relationship between two resources is absent.

(12) The *reverse relation operation*, denoted by $Reverse(\alpha)$ or α^R as in r—α^R→r'. If there is a semantic relation α from r to r', then there is a reverse relation from r' to r, that is, r'—α→$r \Rightarrow r$—α^R→r'. A semantic relation and its reverse declare the same thing, but the reverse relation is useful in reasoning sometimes.

Is there any primitive set of semantic indicators?

It is ideal that there exists a set of semantic indicators (properties or factors) Ω in a given domain, such that any semantic relationship between two resources can be described by an indicator or combination of indicators in Ω.

For example, *overlap*, *include*, *disjoin*, *neighbor*, and *equal* are semantic indicators for geographical relations. The *overlap* and *disjoin* are not primitive indicators in describing geographical relations since *overlap* and *disjoin* relations can be defined by the *include* relation.

Different domains have different semantic link primitives. In some domains, primitives are small and simple. In social network, *spouseOf* and *childOf* are primitives in family relations as other relations can be derived from the primitives and the attributes of semantic nodes.

For semantic relations α, β and $\gamma \in \Omega$, we have: $\alpha \subseteq \beta \subseteq \gamma \Rightarrow \alpha \subseteq \gamma$.

Let X, Y and Z be different semantic nodes, and α and β be two semantic indicators. We say that α is semantically orthogonal to β, denoted by $\alpha \perp \beta$, if and only if X—α→Z and Y—β→Z can uniquely determine Z, that is, if there exists Z' such that X—α→Z' and Y—β→Z', then Z is semantically equivalent to Z'.

If two semantic link chains X_1—α_1→X_2—α_2→ … X_{n-1}—α_{n-1}→X_n and Y_1—β_1→Y_2—β_2→ … Y_{m-1}—β_{m-1}→Y_m are both transitive, and $X_n = Y_m$ can be uniquely determined, the two chains are orthogonal in SLN.

Orthogonal semantic relationships play an important role in some applications especially those that are relevant to layout or geographical positioning.

Like latitudes and longitudes, orthogonal semantic links help people accurately locate a node. For example, if we want to find the destination *X* and know that *A*—*SouthOf*→*X* and *B*—*EastOf*→*X*, then *X* can be largely located at the meeting point of the two links.

Let SLN_1 and SLN_2 be two SLNs sharing the same rule set. If there is a sub-graph of SLN_2 that is semantically equivalent to SLN_1, we say that SLN_1 is semantically included by SLN_2, denoted by $SLN_1 \subseteq SLN_2$.

The semantic inclusion relationship is transitive, that is, $SLN_1 \subseteq SLN_2$, $SLN_2 \subseteq SLN_3 \Rightarrow SLN_1 \subseteq SLN_3$.

SLN_1 implies SLN_2 if there is a mapping φ: $SLN_1 \rightarrow SLN_2$ such that:

(1) *for any node n in SLN_1, we have $n = \varphi(n)$ or the semantic image of n is included by the semantic image of $\varphi(n)$*;
(2) *for any semantic link l in SLN_1, we have $l = \varphi(l)$ or the semantic image of l is included by the semantic image of $\varphi(l)$*; *and,*
(3) *for any rule in SLN_1, we have rule = $\varphi(rule)$ or the semantic image of rule is included by the semantic image of $\varphi(rule)$.*

Two SLNs are *semantically equivalent* to each other if there is an isomorphism between them such that corresponding nodes are the same or semantically equivalent to each other, the semantic indicators of the corresponding links imply each other, and the corresponding rule sets are the same.

The notion of the *minimal cover* of an SLN introduced in section 2.4 provides an effective approach for determining the relationship between SLNs. Section 2.8 will discuss the implementation of the operations on SLNs.

2.3 Relational Reasoning and the Semantic Space

Relational reasoning rules are rules for chaining related semantic links to obtain a reasoned result. Reasoning acts through rules, as for example r—ce→r', r'—ce→r'' $\Rightarrow$ r—ce→r''. A rule for reasoning can also be represented as $\alpha \cdot \beta \Rightarrow \gamma$, where α, β and γ are semantic indicators, and the above rule can be represented as $ce \cdot ce \Rightarrow ce$.

A simple case of reasoning is where all the semantic links have the same type, that is, single-type reasoning. For transitive semantic links, we have the following reasoning rule: r_1—α→r_2, r_2—α→r_3, …, r_{n-1}—α→r_n $\Rightarrow$ r_1—α→r_n. In general, *ce*, *imp*, *st*, *ref*, *and e* are transitive in many fields.

Some semantic links are general, while others are specific to domain. Some rules in one domain may not be valid in the other domains. Some general heuristic reasoning rules are given in (H.Zhuge, Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering*, 21(6)(2009)785-799; H.Zhuge, Retrieve Images by Understanding Semantic Links and Clustering Image Fragments, *Journal of Systems and Software*, 73(3)(2004)455-466).

In some applications, an order relationship "≤" could be defined on the semantic link set. To obtain a well reasoned result, the reasoning mechanism should find the strongest link among its candidates.

Semantic links can also be inexact. An inexact semantic link represents an uncertainty for its relationship, and is denoted by $r—(\alpha, cd)\rightarrow r'$, where α is a semantic indicator and cd is the degree of certainty. Inexact single-type reasoning is of the following form:

$$r_1—(\alpha, cd_1)\rightarrow r_2,\ r_2—(\alpha, cd_2)\rightarrow r_3,\ \ldots\ldots,\ r_n—(\alpha, cd_n)\rightarrow r_{n+1}$$
$$\Rightarrow r_1—(\alpha, cd)\rightarrow r_{n+1},\ \text{ where } cd = min\ (cd_1, \ldots, cd_n).$$

Different types of inexact semantic links can be also chained according to the rules, for example:

$$r—(ce, cd_1)\rightarrow r',\ r'—(imp, cd_2)\rightarrow r'' \Rightarrow r—(ce, min\ (cd_1, cd_2))\rightarrow r''.$$

Another kind of inexactness is associated with the similar link. For example, connecting the *cause-effect* link to the similar link can give the following inexact reasoning rules:

$$r—ce\rightarrow r',\ r'—(sim, sd)\rightarrow r'' \Rightarrow r—(ce, cd)\rightarrow r'',$$
where cd is derived from sd ($cd = sd$ is a simple choice).

With relational reasoning rules, some implicit semantic links can be derived out as indicated by the dotted lines in Fig.2.3.1. Only with a *semantic space* (e.g., constructed by rules and classification hierarchy), a labeled network can be called a semantic link network.

The *semantic link network* will be enhanced in the cyber-physical society:

(1) *The cyber space*, which mainly consists of classification hierarchy and rules. It can be centralized or distributed in the cyber space, social space, and mental space. The rules include not only the reasoning rules but also the rules in the cyber space, physical space and social space.
(2) *The physical space*, which provides the Euclidean space and resources for individuals to live, work, study and interaction. The local spaces where the individuals reside and the distances between individuals indicate local semantics around individuals in the

physical space. For example, home, office, and airplane are local spaces that offer space-specific semantics.

(3) *The social space*, which provides commonsense, social rules and economic principles for indicating and understanding semantics.

(4) *The global semantic link network*, which provides the context for the local semantic link networks. Individuals usually know a small part of the network, but they have the ability to emerge appropriate local semantic link network, and to use it as the context to render the meaning through interaction, although indication may be incomplete. For example, individuals can emerge appropriate local semantic link networks in mind during conversation. This is why incomplete sentences (even incorrect in grammar) do not disturb understanding in conversation.

(5) *Language*, which provides the ground for semantic indicators. It has evolved a symbol space that obeys specific rules that are different from the physical space and social space.

The cyber-physical society provides rich semantics for semantic link network. The semantic link $X—\alpha\rightarrow Y$ or $X—\alpha—Y$ will be mapped from the representation in the symbol space into the enriched semantic space $\Im$:

$$f: \{X—\alpha\rightarrow Y\}\rightarrow \Im.$$

Where f is a mapping from the symbol space representing X, Y and α into the semantic space, or from one symbol space into another well-defined symbol space, in which all symbol expressions are commonsense.

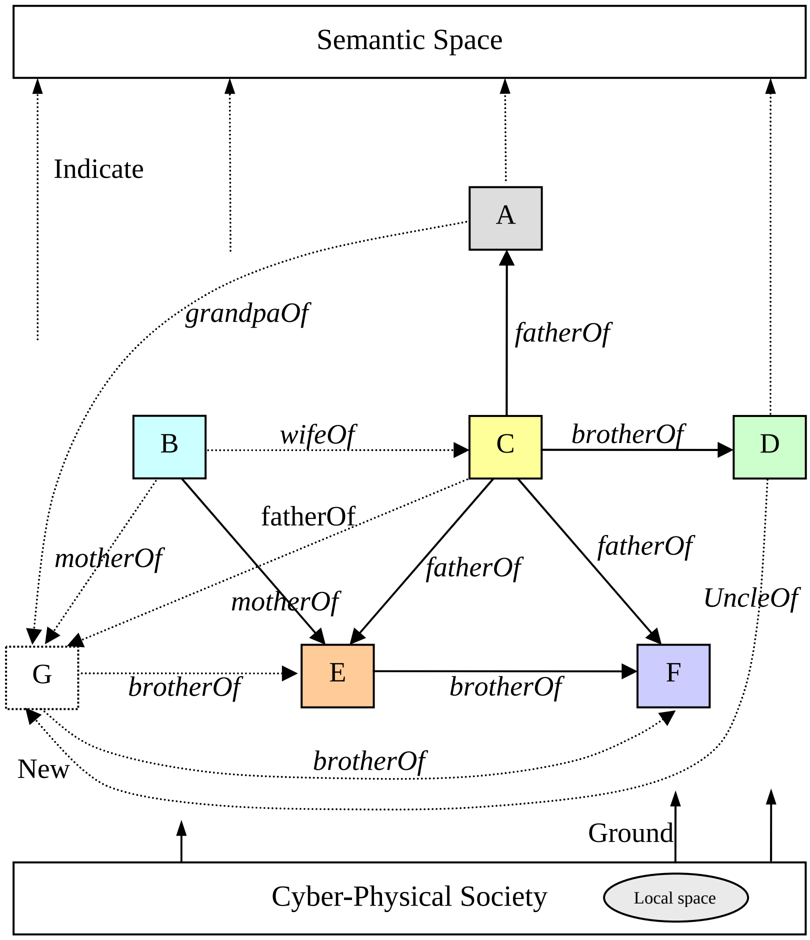


Fig. 2.3.1 Reasoning on the family semantic link network.

The semantics of the indicators $f(\alpha)$ is explained by a semantic space defined by a hierarchy of classes and rules. Fig.2.3.2 shows the relations among the physical space, map, concept hierarchy, and indicator space.

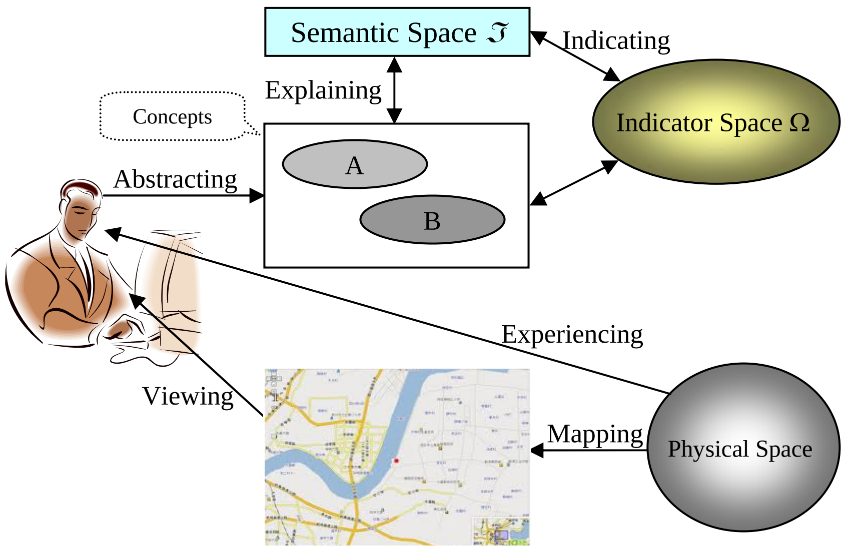


Fig. 2.3.2. The relations among the physical space, map, semantic space, and indicator space.

Section 2.19 will further discuss the issues of semantics.

2.4 An Algebraic Model of the SLN

Operations such as reversal, addition and multiplication can be defined for transforming and composing semantic links. These operations take one or two semantic indicators in and put one semantic indicator out.

Semantic indicators and the operations on them constitute an algebraic system. This section investigates such an algebraic system and its characteristics, assuming that there are no contradict semantic links in the same SLN.

By representing an SLN as a matrix of semantic indicators, reasoning can be carried out by the self-multiplication of the matrix.

If there is a semantic link with indicator α from node r_1 to node r_2, there is a reverse semantic link from r_2 to r_1 called the reversal semantic link, denoted by $Reverse\,(\alpha)$ or α^R.

For example, a cause-effect link from r_1 to r_2 signifies that r_1 is a cause of r_2 and that r_2 is an effect of r_1, that is, it implies a $Reverse\,(ce)$ or ce^R link from r_2 to r_1. A semantic relation and its reverse are equivalent, but the reverse relationship is useful in reasoning.

The following operational laws are clearly true:

(1) $e^R = e$.
(2) $N^R = N$.
(3) $\phi^R = \phi$.
(4) $sim^R = sim$.
(5) $(\alpha^R)^R = \alpha$.

Definition 2.4.1 If there exist two semantic links with indicators α and β from r_1 to r_2, then the two links can be merged into one with the semantic indicator $\alpha + \beta$. Merging is termed the *semantic addition* of α and β.

Certain laws and characteristics of addition follow from this definition.

Laws of Semantic Addition

(1) $\alpha + \alpha = \alpha$ (Idempotency).
(2) $\alpha + \beta = \beta + \alpha$ (Commutativity).
(3) $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$ (Associative Addition).
(4) $\alpha + Null = \alpha = Null + \alpha$.
(5) If $\alpha' \leqslant \alpha$, then $\alpha + \alpha' = \alpha$. In particular, $e + \alpha = e$, where α is a semantic indicator that is compatible with e.
(6) $(\alpha + \beta)^R = \alpha^R + \beta^R$.

Characteristic 2.4.1 For any two semantic indicators α and β in an SLN, we have $\alpha \leqslant \alpha + \beta$ and $\beta \leqslant \alpha + \beta$.

Characteristic 2.4.2 For any three semantic indicators α, β and γ in an SLN, if $\alpha \geqslant \beta$ and $\alpha \geqslant \gamma$, then $\alpha \geqslant \beta + \gamma$.

Definition 2.4.2 If there are two relations, α from r_1 to r_2, and β from r_2 to r_3, in an SLN, and if we can get the semantic indicators γ_1, γ_2, …, and γ_k from r_1 to r_3 by reasoning, then we call the reasoning process *semantic multiplication*, denoted as $\alpha \times \beta = \gamma$ where $\gamma = \gamma_1 + \gamma_2 + \ldots + \gamma_k$.

Laws of semantic multiplication

(1) $\alpha \times e = \alpha = e \times \alpha$
(2) $\alpha \times N = N = N \times \alpha$
(3) $\alpha \times \phi = N = \phi \times \alpha$ (note that $\phi \times \phi = N$)
(4) $(\alpha + \beta) \times \gamma = \alpha \times \gamma + \beta \times \gamma$, and $\alpha \times (\beta + \gamma) = \alpha \times \beta + \alpha \times \gamma$
(5) $(\alpha \times \beta)^R = \beta^R \times \alpha^R$

Lemma 2.4.1 For any semantic links $r_1 \text{—}\alpha\text{→} r_2$, $r_1 \text{—}\beta\text{→} r_2$, and $r_2 \text{—}\gamma\text{→} r_3$ in an SLN, if $\alpha \geqslant \beta$, then the semantic relation from r_1 to r_3 is $\alpha \times \gamma$ $(\alpha \times \gamma \geqslant \beta \times \gamma)$.

Within a certain time period, semantic link network is stable, especially, the semantic space is stable. We can define the concept of semantic closure.

Definition 2.4.3 The *semantic closure* of an SLN S = <N, L, *Rules*> is S^+ = <N, L', *Rules*> such that:

(1) $L \subseteq L'$; and,
(2) A semantic link is added to L' if it is available via rule reasoning on L.

The closure has the following characteristics:

Characteristics:
(1) $S \subseteq S^+$.
(2) *An SLN is equivalent to its closure.*
(3) *Two SLNs are equivalent if and only if their closures are the same or equivalent.*

(4) *The equivalence between closures is a reflexive, symmetric and transitive relation*.

Lemma 2.4.2 For two SLNs S and T, S is equivalent to T if and only if $T \subseteq S^+$ and $S \subseteq T^+$.

The *minimal semantic cover* is obtained by removing all redundant semantic links from a semantic cover.

Definition 2.4.4 An SLN M is the *minimal semantic cover* of another SLN S, if M and S satisfy the following conditions.

(1) $M^+ = S^+$; and,
(2) no semantic link l exists in M such that $(M - l)^+ = M^+$.

The minimal semantic cover of an SLN involves in the fewest possible semantic links while its semantics indication is unchanged. Although the minimal cover of an SLN is unique in most cases, exception can be found in a circular network (B. Xu and H. Zhuge, Basic operations, completeness and dynamicity of cyber physical socio semantic link network CPSocio-SLN. *Concurrency and Computation: Practice and Experience*, 23(9)(2011) 924-939).

Matrix representation. An SLN can be represented by an adjacency matrix called the Semantic Relationship Matrix (SRM). An SLN with n nodes $r_1, r_2, \ldots, r_n$ can be represented by an SRM as follows, where α_{ij} represents the semantic indicator from r_i to r_j, $\alpha_{ii} = e$, and $\alpha_{ij} = \alpha_{ji}{}^R$.

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}$$

If there are no semantic links between r_i and r_j, then $\alpha_{ij} = \alpha_{ji} = Null$. The SRM of any SLN is unique if the allocation of nodes is fixed.

Generally, an element in SRM should be a set of semantic indicators as there may have multiple semantic links between the same pair of nodes. A three dimensional tensor can be used to represent a semantic link network (T. Franz, et al. TripleRank: Ranking Semantic Web Data by Tensor Decomposition. *In Proceedings of the International Semantic Web Conference*, 2009, pp.213-228).

A large-scale complex SLN can be separated into several networks with simpler semantic links, and block matrix can be used to decrease the scale (H. Zhuge, et al., Algebra model and experiment for semantic link network. *IJHPCN*, 3(4)(2005) 227-238).

Reasoning on an SLN derives implicit semantic links between nodes by semantic link reasoning. Suppose an SLN has n nodes: r_1, r_2, …, r_n, and its SRM is *mat*. Can we reliably derive the semantic relations of any two nodes from the SRM?

Clearly, we can get α_{ij} as the semantic link between r_i and r_j if the link is in the matrix. However, α_{ij} is sometimes *Null* even though there may in fact be a semantic link that can be computed. In such cases the reliable semantic link, denoted by $\alpha_{ij}^{\#}$, is derived by reasoning.

Theorem 2.4.1 In an SLN, a reliable semantic link can be computed as $\alpha_{ij}^{\#} = mat_{i*} \times mat^{n-2} \times mat_{*j}$, where mat_{i*} is the i^{th} row vector and mat_{*j} is the j^{th} column vector.

Corollary 2.4.1 In an SLN, we have $\alpha_{ij}^{\#} \geqslant \alpha_{ij}$.

If we compute the semantic links of all pairs of nodes in an SLN, then we can get a new SRM, called the full SRM (FSRM), denoted by mat_f. An FSRM is the SLN matrix of the closure of the original SLN. We can get the reliable semantic link between any two nodes in the SLN from the FSRM. Of course, some of these semantic links are in the original SLN, and some are derived. Any semantic reasoning can be done by self-multiplication of the SLN matrix.

The FSRM is an efficient tool for semantic reasoning because there is a reliable semantic link between every pair of nodes.

Corollary 2.4.2 For a semantic link matrix *mat* and its FSRM mat_f, we have:

(1) $mat_f = mat^{n-1}$
(2) $mat_f \times mat = mat_f$

This corollary suggests a useful way to compute the FSRM.

The following are some characteristics.

Characteristics 2.4.

(1) *A semantic link may have several minimum spanning graphs*.
(2) *Different semantic links may exist between the same pair of semantic nodes*.
(3) *The spanning graph of a semantic link l is a subgraph of the minimum semantic cover such that this subgraph's closure includes l*.
(4) *The minimum spanning graph of semantic link l is a subgraph of the minimum semantic cover such that this subgraph's closure includes l, and there is no other subgraph, whose closure also includes l*.
(5) *Removing a semantic link outside the minimum semantic cover does not influence the semantics of SLN*.
(6) *Adding a semantic link belonging to the closure of the SLN and then removing it do not influence the semantics of SLN*.
(7) *Adding a set of semantic links to SLN may change the minimum semantic cover and the closure of SLN*.

2.5 SLN Normalization

2.5.1 *The normal forms of an SLN*

SLN normal forms are used to ensure the correctness and effectiveness of an SLN's semantic indication and operations. The following definitions help in removing redundancy and inconsistency.

Two semantic nodes are equivalent in indicating semantics if they have the same attributes and belong to the same class.

Definition 2.5.1 If there is no semantically equivalent node in an SLN, then we say that the SLN is in first normal form, or 1NF.

1NF regulates the SLN by excluding redundant semantic nodes.

Definition 2.5.2 If an SLN is in 1NF and there is no inconsistent and implication semantic link between the same pair of nodes, then we say the SLN is in second normal form, or 2NF.

2NF regulates the SLN by excluding redundant and inconsistent semantic links.

Definition 2.5.3 If an SLN is in 2NF and there is no isolated node or part, then we say that the SLN is in third normal form, or 3NF.

The 3NF guarantees that we can reach all the nodes in SLN starting from any of its nodes.

Definition 2.5.4 If an SLN is in 3NF and it cannot be simplified into semantically equivalent one by removing semantic links, then we say that the SLN is in fourth normal form, or 4NF.

2.5.2 *Operations on SLNs*

Applications or user interfaces that manage the resources of, or need services such as browsing for, a large-scale SLN need operations applied to that SLN.

There are six basic operations:

(1) Add a semantic node to the SLN.
(2) Remove a semantic node from the SLN.
(3) Add a semantic link to the SLN.
(4) Remove a semantic link from the SLN.
(5) Add a rule to the SLN.
(6) Remove a rule from the SLN.

Graph-like operations can be defined as follows. Let $SLN_1 = <N_1, L_1, Rules_1>$ and $SLN_2 = <N_2, L_2, Rules_2>$ where N_1 and N_2 are node sets, and L_1 and L_2 are semantic link sets.

(1) *Intersection*: $SLN_1 \cap SLN_2 = <N_1 \cap N_2, L_1 \cap L_2, Rule_3>$, where $Rule_3$ is generated by removing the rules relevant to the semantic links in $L_1 \cup L_2 - L_1 \cap L_2$ from $Rule_1 \cup Rule_2$.
(2) *Union*: $SLN_1 \cup SLN_2 = <N_1 \cup N_2, L_1 \cup L_2, Rule_1 \cup Rule_2>$.
(3) *Inclusion*: returns *true* if $SLN_1 \subseteq SLN_2$, otherwise returns *false*.

Where $L_1 \cap L_2 = \{n \text{—}\alpha\text{→} n' \mid \alpha = min(\alpha_1, \alpha_2),\ n \text{—}\alpha_1\text{→} n' \in SLN_1, n \text{—}\alpha_2\text{→} n' \in SLN_2\}$, if α_2 is equal to or implies α_1, $min(\alpha_1, \alpha_2) = \alpha_1$ else $min(\alpha_1, \alpha_2) = Null$. $L_1 \cup L_2 = \{n \text{—}(\alpha_1 + \alpha_2)\text{→} n' \mid n \text{—}\alpha_1\text{→} n' \in SLN_1, n \text{—}\alpha_2\text{→} n' \in SLN_2\}$.

In the following, we focus on three operations *join*, *split*, and *view*.

Definition 2.5.5. Two SLNs can be joined in one of the following three ways.
(1) If they have one common node, then the join operation (called *join by node*) merges the common node.
(2) If they have one common semantic link or chain, then the join operation (called *join by semantic link*) merges the common chain.

(3) If at least one semantically distinct link can be added between two SLNs, then the join operation (called *join by link addition*) adds such a link.

Lemma 2.5.1. The join operation conserves 1NF, 2NF, and 3NF characteristics.

Proof. The join operation (either *by node* or *by link addition*) does not add any semantically equivalent node or link, so join conserves 1NF and 2NF. For the case of 3NF, we consider the following two aspects:

1) *Join by node* or *by a chain (of semantic links)*. Let SLN_1 and SLN_2 be two SLNs, and let SLN be the join of SLN_1 and SLN_2 by merging the common node n (or chain C) of SLN_1 and SLN_2. If both SLN_1 and SLN_2 are 3NF, then any node in SLN_1 is accessible from any node in SLN_2 through n (or any node of C or chain C), and vice versa. So SLN is 3NF.
2) *Join by link addition*. Let l be the semantic link added between node n in SLN_1 and node n' in SLN_2. If both SLN_1 and SLN_2 are 3NF, then any node in SLN_1 is accessible from n and any node in SLN_2 is accessible from n', so any node in SLN_1 is accessible from any node in SLN_2 through l, vice versa. Hence SLN is 3NF.

Thus, we conclude that the join operation conserves 1NF, 2NF and 3NF characteristics. □

By duplicating nodes or links, or by deleting nodes and links, a single SLN can be split into two SLNs.

The split operation does not increase semantically equivalent nodes and links in either of the split SLNs, so it conserves 1NF and 2NF. The split operation may break accessibility within either of the split SLNs, so it does not necessarily conserve 3NF. But, we can add some conditions to enable it to conserve both 3NF and 4NF characteristics. Fig. 2.5.2.1 shows different split operations.

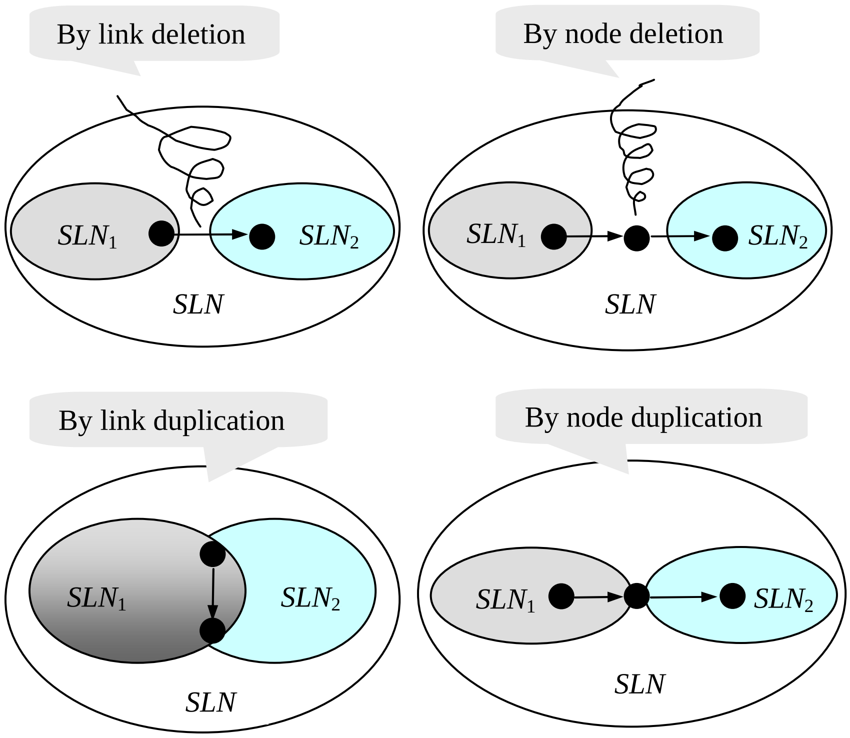


Fig.2.5.2.1 Split operations.

Lemma 2.5.2 The split operation conserves 1NF and 2NF.

Lemma 2.5.3 Let *SLN* be 3NF and let it be split into SLN_1 and SLN_2. If $SLN_1 \cup SLN_2 = SLN$, then both SLN_1 and SLN_2 are 3NF.

Proof. The condition $SLN_1 \cup SLN_2 = SLN$ ensures that the split operation does not remove links or nodes, that is, it allows only splitting by duplication. Hence the connectivity of both SLN_1 and SLN_2 is the same as that of *SLN*. Otherwise, we assume that node n in SLN_i ($i = 1, 2$) is not accessible from p in *SLN*. Consequently, n is also not accessible from *p* because no nodes or links are lost during the split operation. Hence both SLN_1 and SLN_2 are 3NF. □

Lemma 2.5.4 Let *SLN* be 4NF and let it be split into SLN_1 and SLN_2. If $SLN_1 \cup SLN_2 = SLN$, then both SLN_1 and SLN_2 are 4NF.

Proof. From the previous lemma, both SLN_1 and SLN_2 are 3NF. Since the split operation does not add nodes or links, both SLN_1 and SLN_2 are minimal. Hence, both SLN_1 and SLN_2 are 4NF. □

The join and split operations help a user to form a view that reflects the user's interests. A set of view operations can be designed to generate the views of an SLN according to the interest nodes and semantic indicators. It enables an SLN to be adapted to the needs of users.

A view of an SLN consists of a sub-graph and the rules that apply to the sub-graph.

When users do not know which nodes and links are of interest, they can express their interests as a topic, a set of keywords, or a graph consisting of keywords and the relationships between them, from which existing information retrieval (IR) technology can determine a set of nodes and links of interest. Adding logical operations to query make the query more powerful (Q.Zeng, X.Jiang and H.Zhuge, Adding logical operators to tree pattern queries on graph-structured data, *VLDB* 2012).

Therefore, users can obtain a relevant view by specifying their interests in a simple way. Just as overlays applied to a geographical map can meet a variety of needs, views of an SLN can help users browse only potentially interested parts and so can make browsing more effective. Since the view operation does not add nodes or links, it conserves 1NF and 2NF characteristics.

Lemma 2.5.5 SLN views conserve 1NF and 2NF characteristics.

Lemma 2.5.6 If two views have the same normal forms, then their join produces a view that conserves these normal forms.

2.6 Constraints and Displaying

The normal forms provide guidance towards precise SLNs. The following criteria are for building an appropriate SLN.

(1) Any semantic node or semantic link of an SLN should have mapping image (semantic image) in the semantic space.
(2) Any operation of adding a node to an SLN should accompany an operation of adding a link to connect that node to an existing node. This criterion ensures that no node of an SLN is isolated.
(3) Use of the resource represented by any node in an SLN should obey all restrictions and protocols for its use or execution environment laid down when the node was created.

An SLN can be used for the purposes such as browsing, reasoning, reusing and controlling. In some applications, constraints are needed to guarantee appropriate operations. The SLN maker can apply constraints of the following kinds:

(1) A constraint on nodes' attributes.
(2) A constraint on the relationships between attributes for coordinating nodes.
(3) A constraint between a node and its links, specifying a relationship between the links (such as *and*, *or*, *and-split*, *or-split* in the workflow model) and the conditions on links involved.
(4) A meta-constraint, that is, a constraint on constraints. The application of a given SLN is a particular form of explanation, reasoning or execution of the network under these constraints.

SLN can be displayed at the entity resource (e.g., document) level or the abstract semantic level, and can shift between these two levels.

An execution engine and monitor engine can be designed as in the workflow management systems (www.wfmc.org). The execution view should show an SLN satisfying at least the 3NF. Constraints can be set from application requirements when establishing the SLN and can be verified during execution. SLN reasoning can be carried out starting from a user's view of interest, and can be of the following two kinds:

(1) *View reasoning*. Reasoning between views of an SLN is carried out according to the transitivity of inclusion ($\subseteq$) and implication ($\rightarrow$) between views. For example:

$view_1 \subseteq view_2$, $view_2 \subseteq view_3 \Rightarrow view_1 \subseteq view_3$;
$view_1 \rightarrow view_2$, $view_2 \rightarrow view_3 \Rightarrow view_1 \rightarrow view_3$; and,
$view_1 \rightarrow view_2$, $view_3 \subseteq view_2 \Rightarrow view_1 \rightarrow view_3$.

(2) *Link reasoning* of the following five types:

- *Transitive*: reasoning about transitive semantic links: X_1—α→X_2, X_2—α→X_3, …, X_{n-1}—α→$X_n \Rightarrow X_1$—α→X_n.
- *Implication*: $\alpha \rightarrow \beta$, X—α→Y, Y—β→$Z \Rightarrow X$—β→Z.
- *Abstraction*: if X is an abstraction of Z, Z—α→$Y \Rightarrow X$—α→Y. For example, if *Professor* is an abstraction of the person named *Zhuge*, we have: *Zhuge*—*supervise*→*Zhang* ⇒ *Professor*—*supervise*→*Zhang*.
- *By analogy*: X—α→Y, $X \sim X'$, $Y \sim Y' \Rightarrow X'$—α→Y', where ~ represents a similar relationship. The result can be given a certainty degree *cd*, which can be determined from the similarity degrees of $X \sim X'$ and $Y \sim Y'$.
- *Hybrid*: any combination of the above types.

SLN reasoning is to derive implicit semantic links, which is likely to help people to act intelligently.

2.7 SLN Ranking

Nodes in an SLN have ranks differentiating their importance in the network.

2.7.1 *Hyperlink network ranking*

To rank and refine Web search results, search engines analyze hyperlink structure (M.R. Henzinger, Hyperlink Analysis for the Web, *IEEE Internet Computing*, 5(1)(2001)45-50). The HITS algorithm (J.M. Kleinberg, Authoritative sources in a hyperlinked environment, *Journal of the ACM*, 46(5)(1999)604-632) and the PageRank algorithm

(L. Page, et al., The pagerank citation ranking: bringing order to the web, *Technical report*, Stanford, Santa Barbara, CA 93106, January, 1998) are typical of the algorithms used to rank Web pages using hyperlink analysis.

Both algorithms calculate the scores for each page of the entire Web by taking link structure into account. A Web graph consists of Web pages as nodes and hyperlinks as edges.

The HITS algorithm measures the importance of a single Web page from two aspects: *authority* and *hub*. A page with high authority has many pages pointing to it, which implies its authority; a page with a high hub points to many other pages, which implies its richness in material. The authority of a page is computed by summing up the hub scores of all pages pointing to it, while the hub score is the summation of the authority of all pages pointing to it.

PageRank reduces the importance of a page to a single parameter — rank. It simply recognizes that a page with many other pages pointing to it is important because it is frequently cited by others. The rank of a page is evenly distributed among all pages it points to. A page gets a small donation from the rank of each page pointing to it. Then, its rank is calculated by summing up all the donations it gets.

HITS uses both in-links and out-links of a page, whereas PageRank only uses the in-links.

2.7.2 *SLN ranking*

A semantic link can be assigned a *certainty degree* (denoted by *cd*) to reflect the likelihood of a particular semantic relationship between its nodes or components (i.e., community). An inexact link can be shown as C_1—$(l, cd)\rightarrow C_2$, where C_1 and C_2 are components, l indicates a relation, and $cd \in (0, 1)$.

The certainty degree is valuable for ranking components according to their semantic importance when using an SLN. Semantic link structure reflects the relationships between components, just as hyperlink structure reflects the relationships between Web pages.

Naturally, we can adopt hyperlink analysis to rank nodes or components of an SLN, though in the PageRank algorithm every Web page has only one rank. But for an SLN we must take into account different semantic links, as *different semantic links may play different roles in structuring and evolving the network.*

For a semantic component *C*, we can devise an overall rank, called the *T*-rank, and a set of individual *l*-ranks for different *l*-links.

2.7.3 *A ranking algorithm*

Let *C* and *D* be components of an SLN, *l* be a semantic link in semantic link set *L*, F_C^l be the set of components of the SLN *l*-linked from *C*, and B_C^l be the set of components of the SLN *l*-linked to *C*. The sum of the certainty degrees of all *l*-links from *C* can be denoted by

$$N_C^l = \sum_{v \in B_C^L} cd_l^{v \to C},$$

where $cd_l^{v \to C}$ is the certainty degree of an *l*-link from *v* to *C*. The rank of a component *C* can be defined as follows:

$$R(C) = \beta \sum_{l \in L} R_l(C),$$

where $R_l(C) = \beta_l \times \sum_{D \in B_C^l} \frac{R_l(D) \times R(l)}{N_D^l}$,

and

$$R(l) = w_l \times cd_l^{D \to C}.$$

R(*C*) is the *T*-rank of *C*. $R_l(C)$ is the *l*-rank of *C*, derived from all the *l*-links pointing to *C*. β_l and β are normalization factors such that the total *l*-rank (or *T*-rank) of all components is constant. *R*(*l*) is the rank of an *l*-link, being the product of the weight of the link type (w_l) and the certainty degree of the link ($cd_l^{D \to C}$).

A vector of ranks of different semantic links on a semantic node $(R_l(C)|l\in L)$ is useful as applications may concern the roles of different semantic links. So both $R(C)$ and $R_l(C)$ should be known for a semantic link network application.

The l-rank of a component is shared among its outward l-links to contribute to the l-ranks of the components pointed to. As far as the inexactness of semantic links is concerned, the certainty degree is attached to the corresponding l-link.

The above formulas show that the rank of a component C depends recursively on the ranks of the semantic links pointing to it. Individual l-ranks $R_l(C)$ can be calculated by using an iterative algorithm similar to that of PageRank.

The ranking algorithm

Let A^l be an $n\times n$ matrix with rows and columns corresponding to semantic components, where N is the total number of components of the SLN under consideration.

Assume $A^l_{i,j}=cd_l^{C_i\to C_j}/N^l_{C_i}$ when there is an l-link from C_i to C_j with certainty degree $cd_l^{C_i\to C_j}$, otherwise $A^l_{i,j}=0$. A^l is called the l-link adjacency matrix.

If we treat R_l as a vector over semantic components, then we have $R_l=\beta_l(A^l)^T R_l$. So R_l is an eigenvector of $(A^l)^T$ with eigenvalue β_l. We want the dominant eigenvector of $(A^l)^T$.

The algorithm for computing T-rank is described as follows:

Function T-rank (S, A^l)

where A^l is the l-link adjacency matrix and S is an initial vector of l-ranks. S can be almost any vector over the semantic components, so we can simply set S to an n-dimensional vector with every element s_i equals to $1/n$.

```
{
    For each l in Ω
    {
        R_l^0 ← S
        Do
            R_l^{i+1} ← (A^l)^T R_l^i
            δ ← ||R_l^{i+1} − R_l^i||_1
                // note: ||R||_1 is the l_1 norm of vector R
        While δ > ε
    }
    R ← O        // note: O is a zero vector
    For each l in Ω
        R ← R + w_l × R_l
    Return R
}
```

Relevant experiments have been described in (H. Zhuge and L. Zheng, Ranking Semantic-linked Network, *Proceedings of WWW2003*, Budapest, May, 2003, available at www2003.org/cdrom /papers/poster/p148/P148-Zhuge/P148-Zhuge.htm).

It is an interesting topic to infer additional semantics from users' behavior (including interactions between users) when they browse an SLN.

Chapter7 will present a new approach to ranking.

2.8 Implementation of SLN Operations

2.8.1 *Matching between SLNs*

The SLNs discussed here contain no isolated nodes. Matching between two structures of SLN, expressed as graphs $G = (V, E)$ and $G' = (V', E')$, needs to distinguish the following five types of relationship:

(1) *Intersection*. There exists at least one edge that is contained in both E and E'. An edge expressed as $e = \alpha(x, y)$ comprises two vertices x and y and a semantic indicator α.
(2) *Null*. The intersection of E and E' is an empty set.
(3) *Equal*. Every edge in E is also in E', and every edge in E' is also in E.
(4) *Inclusion*. Every edge in E is also in E'.
(5) *Inverse inclusion*. Every edge in E' is also in E.

Let G be a 2NF SLN with at least two vertices and one edge, and $mat(G)$ be the SRM of G, with the nodes of rows and columns in the same sequence. Every element of $mat(G)$ is a possibly empty set of semantic indicators.

Let SLN $G' = (V', E')$ be a subgraph of $G = (V, E)$. For every edge $e = \alpha(x, y) \in E$, if $x \in V'$ and $y \in V'$, and $e \in E'$, we call $G' = (V', E')$ a *fully induced sub-semantic-graph* of G with vertex set V', denoted by $G_V(V')$.

For two SLNs $G = (V, E)$ and $G' = (V', E')$, if $V = V'$, let $R(G)$ be the result of a subtraction, $R(G) = mat(G) - mat(G')$.

$$R(G_{ij}) = mat(G_{ij}) - mat(G'_{ij}) = \bigcup_{i=1}^{3} W_i^k,\ k = 0, 1;$$

where $W_i^k = \{0\}$ if $<i = 1, k = 1>$; $W_i^k =$ null if $<i = 1, k = 0>$, $<i = 2, k = 0>$, and $<i = 3, k = 0>$; $W_i^k = \{+\}$ if $<i = 2, k = 1>$; and, $W_i^k = \{-\}$ if $<i = 3, k = 1>$.

$R_{ij}(G) = \{0\}$ means that there is at least one element that is contained in both $mat_{ij}(G)$ and $mat_{ij}(G')$. $R_{ij}(G) = \{+\}$ means that there is at least one element that is contained in $mat_{ij}(G)$ but not in $mat_{ij}(G')$. $R_{ij}(G) = \{-\}$ means that there is at least one element that is contained in $mat_{ij}(G')$ but not in $mat_{ij}(G)$.

The relationship between G and G' can be determined by the following steps.

(1) Let $V_{int} = V \cap V'$, and G_1 and G_2 be their fully induced sub-graphs with vertex set V_{int}. $G_1 = (V_{int}, E_1) = G_V(V_{int})$ and $G_2 = (V_{int}, E_2) = G'_{V'}(V_{int})$. The relationship between G_1 and G_2 can be determined by the algorithm *Rel_SLN_Vertex* described below, which is also suitable for determining the relationship between any two networks that have the same vertex set.

(2) From the relationship between G_1 and G_2, algorithm *Rel_SLN* determines the relationship between G and G'.

Algorithm 2.8.1 Let $G_1 = (V, E_1)$ and $G_2 = (V, E_2)$ be two structures of SLN, and let $mat(G_1)$ and $mat(G_2)$ be $n \times n$ SLN matrices. The following algorithm sets up a correspondence between the vertices of each SLN and the rows and columns of its matrix, and also a mapping between rows and columns of the two matrices. It is then used to determine the relationship between the structures of SLN.

Rel_SLN_Vertex (SLN G_1, SLN G_2)
{ *Pre_Process* ($mat(G_1)$, $mat(G_2)$); // establishes the node correspondence between G_1 and G_2.
Subtract ($mat(G_1)$, $mat(G_2)$, $R(G)$); // $R(G) = mat(G_1) - mat(G_2)$
RtnStr = *Result* ($R(G)$);
Return *RtnStr*;
}

The algorithm *Rel_SLN* determines the relationship between any two structures of SLN $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ and returns a value in {"intersection", "empty", "equal", "inclusion", "inverse inclusion"}.

Algorithm 2.8.2

Rel_SLN (SLN G_1, SLN G_2)
{ If (V_1 == V_2) Return *Rel_SLN_Vertex* (G_1, G_2);
  If ($V_1 \subset V_2$)
  { Let $G_3 = (V_1, E_3) = G_{2\ V2}\ (V_1)$;
    rtn = *Rel_SLN_Vertex*(G_1, G_3);
    If (*rtn* == "empty" || *rtn* = "inclusion") Return *rtn*;
    If(*rtn* == "equal") Return "inclusion";
    If ((*rtn* == "intersection") || (*rtn* == "inverse inclusion"))
      Return "intersection";
  }
  If ($V_1 \supset V_2$)
  { Let $G_3 = (V_2, E_3) = G_{1\ V1}\ (V_2)$;
    rtn = *Rel_SLN_Vertex* (G_2, G_3);
    If (*rtn* == "empty") Return *rtn*;
    If ((*rtn* == "inclusion") || (*rtn* == "equal"))
      Return "inverse inclusion";
    If ((*rtn* == "intersection") || (*rtn* == "inverse inclusion"))
      Return "intersection";
  }
  Let $V_{int} = V_1 \cap V_2$; // let V_{int} be the intersection of set V_1 and V_2
  If $V_{int} == \Phi$ Return "empty";
  If ($V_{int} != \Phi$)
  { Let $G_3 = (V_{int}, E_3) = G_{1\ V1}\ (V_{int})$ and
    $G_4 = (V_{int}, E_4) = G_{2\ V2}\ (V_{int})$;
    rtn = *Rel_SLN_Vertex*(G_3, G_4);
    If *rtn* == "empty" Return "empty";
    If (*rtn* != "empty") Return "intersection";
  }
}

From the relationship returned by algorithm *Rel_SLN*, we can find which SLN contains richer semantics. For example, if the returned value is "inclusion", then G_2 has richer semantics; if the returned value is "inverse inclusion", then G_1 contains richer semantics.

Algorithm *Rel_SLN* can only be applied to a simple SLN — one containing only atomic nodes.

A complex SLN is one containing a complex node — a node that is itself an SLN. Algorithms for determining the relationship between complex SLNs can be designed by viewing the complex nodes as atomic nodes, and then determining the relationship between corresponding complex nodes.

2.8.2 *The union operation*

The union operation is a kind of *semantic integration* of SLNs, which can combine semantic components. The union operation is useful for forming a complete semantic image (*single semantic image*) during browsing and reasoning.

The union of two structures of SLN $G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$, $G_3 = (V_3, E_3)= G_1 \cup G_2$, can be constructed as follows:

(1) View all the nodes in G_1 and G_2 as atomic nodes, $V_3 = V_1 \cup V_2$ and $E_3 = E_1 \cup E_2$.
(2) If a node $V_{1c} \in V_1$ is a complex node, and $V_{1c} \notin V_1 \cap V_2$, then the SLN expanded by V_{1c} yields the SLN expanded by the node V_{3c} (corresponding to V_{1c}) of G_3.
(3) If a node $V_{2c} \in V_2$ is a complex node, and $V_{2c} \notin V_1 \cap V_2$, then the SLN expanded by V_{2c} yields the SLN expanded by the node V_{3c} (corresponding to V_{2c}) of G_3.
(4) If the node $V_c \in V_1 \cap V_2$ is a complex node (let G_{1Vc} and G_{2Vc} be the SLNs expanded by V_c in V_1 and V_2 respectively, $V_{3c} \in V_3$ be the complex node corresponding to V_c, and G_{3V3c} be the SLN expanded by V_{3c}), then we have $G_{3\ V3c}= G_{1Vc} \cup G_{2Vc}$.

Algorithm 2.8.3 The algorithm *Union_SLN* for uniting two structures of SLN.

Union_SLN (SLN G_1, SLN G_2, SLN G_3)
{ $V_{int} = V_1 \cap V_2$;
$d = |V_1| + |V_2| - |V_{int}|$;
$V_3 = V_1 \cup V_2$;
Let L_1, L_2 and L_3 be arrays with one dimension;
Set all the nodes in V_1, V_2, and V_3 to arrays L_1, L_2, and L_3 respectively and ensure each location of L_i only contains one node;
Initialize (*mat*(G_3));
// *mat*(G_3) is a $d \times d$ SLN-matrix, every $mat_{i,j}(G_3)$ is set to *null*
For every $mat_{i,j}(G_3)$ ($1 \le i \le d$, $1 \le j \le d$)
{ Let $node_i = L_3[i]$ and $node_j = L_3[j]$;
If (both $node_i$ and $node_j$ belong to V_{int})
{ Let v_{1_i}, v_{1_j}, v_{2_i}, v_{2_j} satisfy the following equations:
$node_i = L_1[v_{1\_i}]$, $node_j = L_1[v_{1\_j}]$,
$node_i = L_2[v_{2\_i}]$, $node_j = L_2[v_{2\_j}]$;
$mat_{i,j}(G_3) = mat_{v1\_i,\, v1\_j}(G_1) \cup mat_{v2\_i,\, v2\_j}(G_2)$;
// the union of two sets.
}
If (only one node belongs to V_{int})
{ Suppose that another node belongs to V_k;
// k belongs to {1, 2}
Let v_{k_i} and v_{k_j} satisfy the following equations:
$node_i = L_k[v_{k\_i}]$, $node_j = L_k[v_{k\_j}]$;
$mat_{i,j}(G) = mat_{vk\_i,\, vk\_j}(G_k)$;
}
If (neither $node_i$ nor $node_j$ belongs to V_{int})
{ if (both $node_i$ and $node_j$ belong to V_k)
// k belongs to {1, 2}
{ Let v_{k_i} and v_{k_j} satisfy the following equations:
$node_i = L_k[v_{k\_i}]$, $node_j = L_k[v_{k\_j}]$;
$mat_{i,j}(G) = mat_{vk\_i,\, vk\_j}(G_k)$;

```
                }
                If (node_i and node_j are not in the same SLN)
                mat_{i,j}(G_3) = null;
            }
        }//End for
}
```

The algorithm *Union_SLN* only applies to simple SLNs. For complex SLNs, the algorithm *Rel_HyperSLN* needs two more steps:

(1) View the complex nodes in two networks as atomic nodes, and then use *Union_SLN* to unite them.
(2) Use *Rel_HyperSLN* to decide how to unite complex nodes.

2.8.3 *SLN-level reasoning*

For a given set of SLNs $S = \{G_1, G_2, \ldots, G_n\}$, the algorithm *Rel_SLN* can yield a set of SLNs that have *inclusion* or *inverse inclusion* relationships. Suppose that $\{G_{s1}, G_{s2}, \ldots, G_{sm}\}$ ($1 \le si \le n$, $1 \le i \le m$) is the result that satisfies $G_{s1} \subseteq G_{s2} \subseteq \ldots \subseteq G_{sm}$. So, when users want to know something from G_{si} ($i = 1, 2, \ldots, m-1$) in applications, G_{sm} can be used to replace G_{si}.

The detailed implementation of SLN reasoning and its role in implementing an intelligent browser are described in “Semantic Link Network Builder and Intelligent Semantic Browser” (H. Zhuge and R. Jia, *Concurrency and Computation: Practice and Experience*, 16(13)(2004)1453-1476).

The SLN inclusion relationship can be extended to an *implication-inclusion relationship* $\le$: $G_1 = \langle N_1, L_1, Rule_1 \rangle \le G_2 = \langle N_2, L_2, Rule_2 \rangle$ if and only if (1) $N_1 \subseteq N_2$; (2) for any $e \in L_1$, there exists $e' \in L_2$ or e' can be derived from $Rule_2$ such that e' is equal to or implies e; and, (3) $Rule_1 \subseteq Rule_2$. Such an implication-inclusion relationship extends SLN reasoning.

To extend SLN applications, the following problems are worth thinking about:

(1) How might the semantic relationship between SLNs be determined? An SLN provides the semantic context for its nodes. There are also semantic relationships between SLNs. Algorithms for finding such semantic relationships would help advance use of the SLNs.
(2) How might a closely related subset of a large SLN be found that satisfies a user's query? Especially for a dynamic SLN.
(3) How might semantic links in a given set of resources be automatically found or established? An approach to solving this problem by extending data mining algorithms with analogical and deductive reasoning was early suggested (H. Zhuge, et al., An Automatic Semantic Relationships Discovery Approach, *WWW2004*, www2004.org/proceedings/docs/2p278.pdf).

2.9 SLN Analogical Reasoning

Analogical reasoning is an important way of thinking. It is non-deductive, that is, the conclusion does not deductively follow from the premises. Its use can uncover semantic relationships that are not available to deduction.

Analogical reasoning was investigated in AI and cognitive science (R.E. Kling, A Paradigm for Reasoning by Analogy, *Artificial Intelligence*, 10(1978)147-178; D. Genter, Structure Mapping: A Theoretical Framework for Analogy, *Cognitive Science*, 7(1983)155-170). Structural analogical reasoning works by structure mapping between related objects. However, research progress has been limited in recent years due to a bottleneck in capturing the structure of objects. *By its very nature, an SLN yields structural and semantic data that can be used to support analogical reasoning*.

2.9.1 *Analogical reasoning modes*

In many cases, only one component of a large SLN is necessary for reasoning. A semantic component consists of semantic nodes and semantic links on the same topic (under the same direct super-class in classification hierarchy). Sharing concept hierarchy with other components, a semantic component can be represented as $SC = <C, L, Rules>$. C is a set of fine-grained semantic components or nodes. L is the set of semantic links between members of C. $L = \{<c_i, l, cd, c_j> \mid c_i, c_j \in C\}$. $<c_i, l, cd, c_j>$ denotes c_i—$<l, cd>$→ c_j. *Rules* are about L and C. Let U be the set of all semantic components in an SLN.

The following three relations between semantic components in U can be defined:

(1) *Inclusion*. The relationship between a semantic component SC' and its fine-grained components SC. Such a relationship can be denoted by $SC \subseteq SC'$. SC is called a sub-component of SC'. The inclusion degree $IncD(SC, SC')$ denotes the proportion of SC' taken up by SC.

(2) *Similarity*. Two semantic components SC and SC' are similar if their graph structures are isomorphic, that is, if there exists a one-to-one correspondence between their nodes and edges. We denote such a relationship by $SC \cong SC'$.

(3) *Partial similarity*. Two semantic components SC and SC' are partially similar if they have similar sub-components. We denote such a relationship as $SC \approx SC'$. $SimD(SC, SC')$ denotes the degree of similarity between SC and SC'. The inclusion degree reflects the degree of similarity between a semantic component and its sub-component. If $SC \subseteq SC'$, then $SimD(SC, SC') = IncD(SC, SC')$.

A structural mapping can be established from the *illustrative problem-solution pair* to the *target problem-solution pair* as shown in Fig. 2.9.1.1, where the dashed arrow from the solution S' to the problem P' means that S' is the solution to P' discovered by structural analogy.

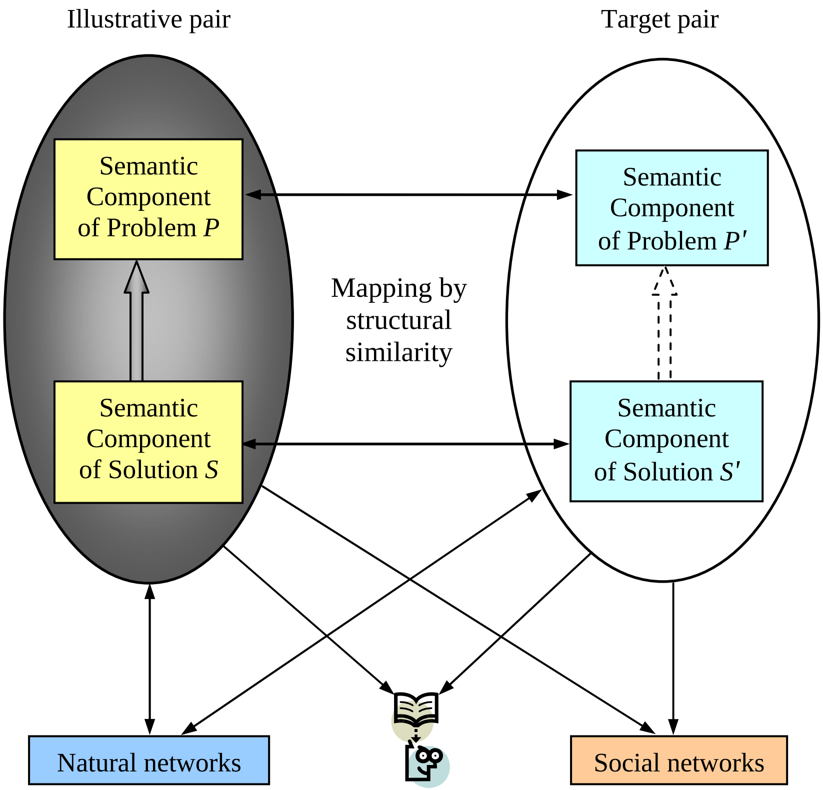


Fig. 2.9.1.1 Structural analogy in different networks.

The set of SLN analogical rules, *RULES*, is a subset of $U \times U$. A rule $r \in RULES$, denoted by SC_1—$<l, cd>$$\rightarrow$$SC_2$ ($SC_1, SC_2 \in U$), means that there is an l-type semantic link with certainty degree cd from SC_1 to SC_2 established manually, or computed by deduction or analogy.

The following are some propositions:

(1) In solutions to l-type problems, for SC_1, $SC_1\text{—}\langle l, cd\rangle\rightarrow SC_2 \Rightarrow SC_2$, means that SC_2 includes all resources (components) that have an l-relationship with a resource of SC_1.
(2) If $SC_1\text{—}\langle l, cd\rangle\rightarrow SC_2$ and $SC_1 \subseteq SC_1'$, then $SC_1'\text{—}\langle l, cd'\rangle\rightarrow SC_2$, where $cd' < cd$.
(3) If $SC_1\text{—}\langle l, cd\rangle\rightarrow SC_2$ and $SC_2' \subseteq SC_2$, then $SC_1\text{—}\langle l, cd'\rangle\rightarrow SC_2'$, where $cd' < cd$.

SLN analogical rules for two semantic components are stronger than rules for their sub-components. Analogical reasoning should use the strongest rules possible.

For *rule*: $SC_1\text{—}\langle l, cd\rangle\rightarrow SC_2$ and *rule'*: $SC_1'\text{—}\langle l, cd'\rangle\rightarrow SC_2'$, if $SC_1 \subseteq SC_1'$ and $SC_2 \subseteq SC_2'$ then *rule'* is semantically stronger than *rule*, denoted by $rule < rule'$ or $rule' > rule$. The degree to which *rule'* is stronger than *rule* can be measured by $PD = 1/(IncD\,(SC_1, SC_1') \times IncD\,(SC_2, SC_2'))$.

Analogical reasoning goes from premises to conjecture. The premise portion includes existing SLN analogical rules and some relations (e.g., *inclusion*, *similarity* and *partial similarity*) between semantic components. The conjecture is the rule reasoned from the premises. The certainty degree of the conjecture depends on various degrees in the premises, the certainty degrees of the rules, the inclusion degrees, and the similarity degrees. Because of the uncertainty inherent in analogical reasoning, the certainty degree of the concluding rule takes the uncertain type $\sim cd$.

Based on the above propositions, we offer the following analogical reasoning modes of SLN.

(1) *Fidelity enforcement mode*

$$\frac{\text{Premises:}\quad SC_1\text{—}\langle l, cd\rangle\rightarrow SC_2,\ \ SC_1' \cong SC_1,\ \ SC_2' \cong SC_2}{\text{Conjecture:}\quad SC_1'\text{—}\langle l, \sim cd\rangle\rightarrow SC_2'}$$

Since a semantic component could be a hierarchy, in some cases it is difficult to determine if $SC_1' \cong SC_1$. Sometimes, a transformation function φ (for example, semantic reconstruction by adding, deleting, splitting and merging sub-components) can help. So, we have:

$$\frac{\text{Premises:}\quad SC_1 \text{—}\langle l, cd\rangle \rightarrow SC_2 \qquad \varphi(SC_1') \cong SC_1,\ SC_2' = \varphi'(SC_2''),\ SC_2'' \cong SC_2}{\text{Conjecture:}\quad SC_1' \text{—}\langle l, \sim cd\rangle \rightarrow SC_2'}$$

where φ and φ' reflect a kind of invariance between components.

(2) *General mode*

$$\frac{\text{Premises:}\quad SC_1 \text{—}\langle l, cd\rangle \rightarrow SC_2,\ SC_1' \approx SC_1,\ SC_2' \approx SC_2}{\text{Conjecture:}\quad SC_1' \text{—}\langle l, \sim cd'\rangle \rightarrow SC_2'}$$

where $\sim cd' = cd \times min\ (SimD\ (SC_1, SC_1'), SimD\ (SC_2, SC_2'))$.

As in fidelity enforcement mode, we have:

$$\frac{\text{Premises:}\quad SC_1 \text{—}\langle l, cd\rangle \rightarrow SC_2, \qquad \varphi(SC_1') \approx SC_1, SC_2' = \varphi'(SC_2''), SC_2'' \approx SC_2}{\text{Conjecture:}\quad SC_1' \text{—}\langle l, \sim cd'\rangle \rightarrow SC_2'}$$

where $\sim cd' = cd \times min\ ((SimD\ (SC_1, \varphi(SC_1')),\ SimD\ (SC_2, SC_2''))$.

(3) *Multiple analogy modes*

For two semantic components $SC = \langle C, L\rangle$ and $SC' = \langle C', L'\rangle$, the union of SC and SC' is $SC \cup SC' = \langle C \cup C', L \,\Delta\, L'\rangle$, where $L \,\Delta\, L'$ is the result of eliminating redundant links from $L \cup L'$.

Given $SC_{1i} \subseteq SC_1$, $SC_{2i} \subseteq SC_2$, $SC_{1i}' \subseteq SC_1'$, and $SC_{2i}' \subseteq SC_2'$, if the following fidelity enforcement reasoning is true for $i = 1, 2, \ldots$, and k.

$$\frac{\text{Premise:}\quad SC_{1i} \text{—}\langle l, cd\rangle \rightarrow SC_{2i},\ SC_{1i}' \cong SC_{1i},\ SC_{2i}' \cong SC_{2i}}{\text{Conjecture:}\quad SC_{1i}' \text{—}\langle l, \sim cd\rangle \rightarrow SC_{2i}'}$$

Then, we have:

$$\frac{\text{Premise:} \quad SC_1 \text{—} \langle l, cd \rangle \rightarrow SC_2}{\text{Conjecture:} \quad SC_1' \text{—} \langle l, \sim cd' \rangle \rightarrow SC_2'}$$

where $\sim cd' = cd \times Min\ (\ (\ |\bigcup_{i=1}^{k} C_{1i}\ |\ /\ |C_1|\ ) \times (\ |\mathop{\Delta}_{i=1}^{k} L_{1i}\ |\ /\ |L_1|\ ),\ (\ |\bigcup_{i=1}^{k} C_{2i}\ |\ /\ |C_2|\ ) \times (\ |\mathop{\Delta}_{i=1}^{k} L_{2i}\ |\ /\ |L_2|\ ),\ (\ |\bigcup_{i=1}^{k} C_{1i}\ '\ |\ /\ |C_1'|\ ) \times (\ |\mathop{\Delta}_{i=1}^{k} L_{1i}\ '\ |\ /\ |L_1'|\ ),\ (\ |\bigcup_{i=1}^{k} C_{2i}\ '\ |\ /\ |C_2|\ ) \times (\ |\mathop{\Delta}_{i=1}^{k} L_{2i}\ '\ |\ /\ |L_2'|\ )\ )$.

Multiple analogy employs inductive reasoning.

(4) *Inexact analogy mode*

$$\frac{\begin{array}{l}\text{Premise:} \quad SC_1 \text{—} \langle l, cd \rangle \rightarrow SC_2, \\ \qquad\qquad SC_1' \subseteq SC_1,\ IncD\ (SC_1', SC_1) > \sigma,\ \ SC_2' \subseteq SC_2\end{array}}{\text{Conjecture:} \quad SC_1' \text{—} \langle l, \sim cd' \rangle \rightarrow SC_2'}$$

where $\sim cd' = cd \times min\ (IncD\ (SC_1', SC_1), IncD\ (SC_2', SC_2))$ and σ is the lower bound of the inclusion degree.

Correspondingly, we have:

$$\frac{\begin{array}{l}\text{Premise: } SC_1 \text{—} \langle l, cd \rangle \rightarrow SC_2,\ \ \varphi(SC_1') \subseteq SC_1,\ \ IncD\ (\varphi\ (SC_1'), SC_1) > \sigma, \\ \qquad\qquad SC_2' = \varphi'\ (SC_2''),\ \ SC_2'' \subseteq SC_2\end{array}}{\text{Conjecture:} \quad SC_1' \text{—} \langle l, \sim cd' \rangle \rightarrow SC_2'}$$

where $\sim cd' = cd \times min\ (IncD\ (\varphi(SC_1'), SC_1), IncD\ (SC_2'', SC_2))$.

The proposed analogical modes still hold if we replace “semantic component” with “SLN”.

2.9.2 *Process of analogical reasoning*

Suppose that developers construct some semantic components and store them in a *semantic components base* (*SCB*). Then, they store the (problem-solution) component pairs with an *l*-type relationship in an *l*-

type illustrative pairs base (*l_IPB*) in a three-tuple form (ID_S, ID_P, *cd*), where *l* is a semantic relation, ID_S is the solution component, ID_P is the problem component, and *cd* is the certainty degree of the *l*-link from ID_S to ID_P.

The general procedure of the fidelity enforcement analogical mode is as follows.

(1) *Discover similar components in the SLN*. The analogical agent analyses the *SCB* and computes the similarity degree of pairs of components. Pairs (c_i, c_j) with similarity degree greater than a certain lower bound are regarded as similar and are stored in the *similar components base* (*simB*) in a form such as (IDc_i, IDc_j, $SimD\,(c_i, c_j)$).
(2) *Prepare a new problem*. When a user presents a new problem *r* with the *l*-type relationship, a human or virtual agent checks if it matches an existing component *p* in the *SCB*. If it does not exist, an attempt is made to derive a semantic component *p* from *r*. If *p* exists, it is added to the *SCB* and step (3) is skipped. Otherwise, if there are items like (ID_S, ID_P, *cd*) in the *l_IPB*, the best of these is used to select s_{best} from the *SCB*, s_{best} is returned to the user as the best solution to *r*, and the process ends.
(3) *Find components similar to the new problem*. The analogical agent computes the degree of similarity between the new *p* and other components in the *SCB*. Any similar component pairs are stored in the *simB* as (ID_P, $ID_{P\mathrm{sim}}$, $SimD\,(p_{sim}, p)$). If no similar pairs are discovered, the process ends without a result.
(4) *Analogy by similarity*. The matching agent selects the component p_{max} with the highest degree of similarity to *p* from all components in the *simB*. If there are solutions to p_{max} in the *l_IPB* then the agent finds the best solution s_{best}' as in step (3), selects the component s_{max} with the highest degree of similarity to s_{best}' from the *simB*, adds it to the *l_IPB* as a newly discovered problem-solution pair (IDs_{max}, ID_p, $\sim(cd_{s_{best}'-<l,cd>\to p_{\max}} \times min\,(SimD\,(IDs_{best}',$ $IDs_{max}),\ SimD\,(p,\ p_{max}))))$, and returns s_{max} from the *SCB* to the user as the solution to *r*. If there is no s_{max}, then the pair (IDp, IDp_{max}, $SimD(p_{max}, p)$) is taken from the *simB* and step (4) is

repeated until all p_{sim} have been considered. If there is no s_{max} for any p_{sim}, the process ends without a result.

The processes for other analogical modes can be obtained by appropriately modifying the process for the fidelity enforcement mode. To ensure the reliability of conjectures, the validity of any computed solution should be verified manually, so that only valid solutions are added to the *l_IPB*.

Useful conjectures can be drawn when we look into the relationship between reasoning and rank. Reasoning can affect a semantic component's rank since semantic links added by reasoning will change components' ranks. In reasoning on an SLN, the rank of its components can be different before and after the computation.

Problem-solving applications usually work on multiple candidate semantic components. So, the candidate solution (semantic component) link with the strongest semantic link ($l_{strongest}$, if there is any semantic priority order) and the greatest certain degree for the problem (semantic component) should be taken as the best solution. Now we take the component's rank into account. The component's rank provides an overall view of the SLN to help in choosing the best candidate solution, especially in analogical reasoning. Since analogical reasoning is uncertain in nature, semantic link deduction is in general more reliable than analogical reasoning.

In analogical reasoning, using rank to select the best solutions is more rational than using semantic priority or certainty degree. The candidate solution component with the highest *T*-rank or with the highest $l_{strongest}$-rank should be selected as the best solution. Thus, we need to introduce component rank to modify the definitions of inclusion and similarity degrees.

For $SC \subseteq SC'$, the inclusion degree can also be measured as follows:

$$IncD_r\,(SC, SC') = \left(\sum_{sc\in\{C\}} R(sc) \Big/ \sum_{sc\in\{C'\}} R(sc) \right) \times \left(\sum_{l\in\{L\}} w_l \Big/ \sum_{l\in\{L'\}} w_l \right).$$

Characteristic 2.9.1 Let $SC_1 = <C_1, L_1>$ and $SC_2 = <C_2, L_2>$ be two semantic components. The symbol $\propto$ denotes: >, = or <. If $SC_1 \subseteq SC$, $SC_2 \subseteq SC$, $SC_1' \subseteq SC'$, $SC_2' \subseteq SC'$, $SC_1 \cong SC_1'$, $SC_2 \cong SC_2'$, and $IncD_r\,(SC_1, SC) \propto IncD_r\,(SC_2, SC)$, we do not always have $IncD_r(SC_1, SC) \propto IncD_r\,(SC_2, SC)$.

The similarity degree is based on the inclusion degree, as follows:

$$SimD_r\,(SC, SC') = IncD_r\,(SC_k, SC) \times IncD_r\,(SC_k', SC'),$$

where $1 \leq k \leq n$ and $IncD_r\,(SC_k, SC) \times IncD_r\,(SC_k', SC') = max\,(IncD_r\,(SC_1, SC) \times IncD_r\,(SC_1', SC'),\ IncD_r\,(SC_2, SC) \times IncD_r\,(SC_2', SC'),\ \ldots,\ IncD_r\,(SC_n, SC) \times IncD_r\,(SC_n', SC'))$.

A strategy of raising the effectiveness of analogical reasoning is to *make use of the minimum semantic cover for analogy*.

Analogical reasoning is an important reasoning mode of SLN, which can help establish an autonomous semantic link network model (H. Zhuge, Autonomous semantic link networking model for the Knowledge Grid, *Concurrency and Computation: Practice and Experience*, 19(7)(2007) 1065-1085), recommend useful information, discover knowledge, and promote imagination in many applications such as e-learning and e-science. With the evolution of an SLN, the semantic components may change, so the results of analogical reasoning may be different.

Study of analogical reasoning in psychology provides some evidence for developing the right computing model (R.J. Sternberg and B.Rifkin, The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, 27(2)(1979) 195-232;

D.E.Rumelhart and A.A.Abrahamson, A model for analogical reasoning, *Cognitive Psychology*. 5(1)(1973)1-28).

Analogy can be extended in the Cyber-Physical Society as it can carry out in and through the cyber space, physical space, socio space and mental space rather than just in the cyber space.

2.10 Dynamicity of SLN

A dynamic SLN reflects dynamic semantic relationships between its resources. Here presents the notion and principles of dynamic SLNs, and describes the functions of a browser.

In a dynamic SLN, the semantic relationships between resources can be changed at any time. A semantic link can carry a temporal relationship between its nodes, which can be either a document or another dynamic SLN.

Such a dynamic semantic link can change its nature under different conditions and at different times. The condition can be given by an event name or a Boolean expression. The time duration can be in one of the following three forms: [-, t] (effective up to time t), [t, $t+\Delta t$] (effective between time t and $t+\Delta t$), or [t, -] (effective from time t on). If there is no condition, then the link is static during the given time duration. If there is no time duration, then the link depends only on its condition.

A single dynamic semantic link consists of the link definition and a dynamic semantic description.

$$r_A \rightarrow r_B < [l_1\,(condition_1, presence\text{-}time, duration_1), \ldots,$$
$$l_n\,(condition_n, presence\text{-}time, duration_n)]$$

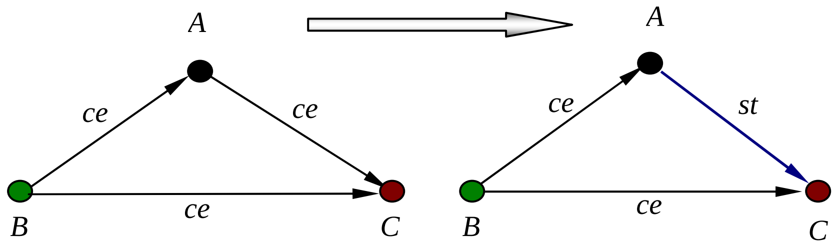
By changing the condition, presence time, and duration, we can dynamically adapt a link to reflect change.

A dynamic SLN can be given a life cycle, during which its content can be changed by dynamic semantic links or by adding new content. *The presence time also records the order of adding semantic links to SLN*.

Using SLN to organize the resources of the future Web, *users may obtain different contents when inputting the same request at different times*. This would enable people to get up-to-date contents and to put the contents to new uses.

Relationships between semantic links impose constraints on dynamic change if change of one semantic link would create *incompatibility* with another semantic link, and so affect the whole or a portion of an SLN. If change does not conflict with the transitive and implicative relationships between existing semantic links, then it is a *compatible change*.

Fig. 2.10.1 shows an example of semantic link change. The change of the semantic link between *A* and *C* does not conflict with the semantic link between *B* and *C*, and the link between *A* and *B*. We can still derive $B\text{—}ce\rightarrow C$ from $B\text{—}ce\rightarrow A$ and $A\text{—}st\rightarrow C$. So, the change is compatible.



(a) before change (b) after change

ce: cause-effect; *st*: subtype

Fig. 2.10.1 Compatible change of semantic link.

The following measures ensure the compatibility of semantic link change:

(1) Limiting change within the semantic scope determined by the transitive and implicative relationships between all types of semantic links. Change can also be carried out smoothly by adjusting the certainty factors of links.
(2) Limiting link change towards lower abstraction level.
(3) Limiting the condition and duration change of a link according to the conditions and durations of its neighbors and of the SLN at the next higher level.
(4) Changing all affected semantic links to maintain overall semantic compatibility when incompatible change is otherwise inevitable.

A dynamic SLN can use the following reasoning paradigms.

Forward dynamic semantic link chaining under a single condition:

$r_A \rightarrow r_B$ [l ($condition_1$, $presence\text{-}time_1$, $duration_1$)]
$r_B \rightarrow r_C$ [l ($condition_2$, $presence\text{-}time_2$, $duration_2$)]

$r_A \rightarrow r_C$ [l ($condition_3$, $presence\text{-}time_3$, $duration_3$)], and $duration_3 = duration_1 \cap duration_2$.

Analogical reasoning:

$r_A \rightarrow r_B$ [l_1 ($condition_1$, $presence\text{-}time_1$, $duration_1$), …, l_n ($condition_n$, $presence\text{-}time_n$, $duration_n$)]
$r_A \sim r_{A'}$, $r_B \sim r_{B'}$

$r_{A'} \rightarrow r_{B'}$ [l_1 ($condition_1$, $duration_1$), …, l_n ($condition_n$, $duration_n$)], CF

Where ~ denotes the similarity relationship between nodes, and the similarity degrees determine the certainty factor CF of the result.

In a connected SLN, any node is accessible from any other node via semantic links. A connected SLN is said to be semantically interconnected if there is no conflict between the semantics of its links. A semantic link will disappear when its conditions and timings do not satisfy its dynamic description. Disappearance of semantic links may destroy the connectivity.

Dynamic change of semantic links may cause an SLN to be continually changing between connected and disconnected states. A connected network may even break into several isolated semantic fragments, some of which may reconnect after changes to their semantic links.

Two criteria — *disruption* and *focus* — can be used to measure the connectivity of an SLN.

Computation of these involves *fragments* (the number of fragments), *nodes* (the number of nodes), and *topics* (the number of topics, the corresponding concepts at higher abstraction level). The smaller the degree of disruption and the larger the degree of focus, the better the connectivity.

The *degree of disruption* of an SLN depends on the number of fragments and the number of nodes. The degree of disruption for a particular topic depends only on the number of fragments and nodes relevant to the topic.

The *degree of focus* for a topic depends only on the number of fragments relevant to the topic. The degree of focus of the overall network depends on the sum of the degrees of focus of all topics and the total number of topics. The minimum number of fragments is one, so the degree of focus on a topic is equal to or less than one.

$$disruption\ (time) = fragments\ (time) \,/\, nodes\ (time)$$
$$disruption\ (time, topic) = fragments\ (time, topic) \,/\, nodes\ (time, topic)$$
$$focus\ (time, topic) = 1 \,/\, fragments\ (time, topic)$$

$$focus\ (time) = \sum_{topic} Focus(time, topic) \,/\, topics$$

A well-behaved dynamic SLN should evolve to increase the focus, and decrease the disruption, of its topics.

What are the advantages if we build an SLN browser?

The Web browser retrieves Web pages from their Web locations as provided by the user or through hyperlinks. The browser displays Web pages only in human readable not in "machine" (software mechanism) understandable forms. On the other hand, the contents of Web pages and their hyperlinks are not easily adapted to new content or new uses.

Unlike the Web browser, the semantic browser is a retrieval mechanism that can generate and display a view of a large SLN. The view can be displayed either as a network or as a page of text. The SLN browser will have the following key features.

(1) *Real-time update*. While they are being browsed, semantic links may change as events happen or as conditions are satisfied. The browser updates the view on display in real-time.
(2) *Change tracing*. Each node (or link) will record semantic link (or node) changes so that the nature of the evolution of the SLN can be studied.
(3) *Dynamic semantic link reasoning*. Semantic link reasoning is carried out to provide hints to help users anticipate results. The hints may vary as the semantic links vary.
(4) *Explanation*. The browser can explain a result, and any changes in it, from the reasoning rules and the record of changes.
(5) *Accessibility checking*. Since semantic links between resources change dynamically, accessible resources may become inaccessible. The browser will check the accessibility of the views on display whenever its links change. A resource becoming inaccessible may fragment the view of a topic. The browser will retrieve and display the fragments.
(6) *Evaluation*. The browser will evaluate the degrees of focus and disruption of the SLN for the topic of interest and display these with the content being browsed.
(7) *Emerging semantics*. The browser will display different structures with the evolution of SLN. In each display, the nodes and links in the tightly connected components are about the same topic.

The dynamic SLN reflects dynamic semantic relationships between various resources. This enables users to browse up-to-date content and to easily put that content to new uses. The dynamic semantic browser can intelligently steer the browsing of any view of the network using various semantic link reasoning paradigms, trace content change and evaluate the degrees of focus and disruption.

An SLN can be regarded as a dynamic process of sequentially adding semantic links to the current SLN.

The following characteristic reflects the dynamicity of SLN evolution.

Defintion 2.10.1. Removing semantic links *l* and *a* from the minimum semantic cover of SLN constructs SLN'. If the closure of SLN' does not include *a* while the closure of SLN'$\cup\{l\}$ includes *a*, we say that *l determines a* (or *a* is determined by *l*). All semantic links determined by *l* forms a *determining set* of *l*.

Characteristics

(1) *The determining set of two semantic links may have intersection*.

(2) *Semantic links are only determined by the semantic links in the minimum semantic cover*.

(3) *If the semantic link to be deleted is not in the minimum semantic cover, the SLN equals to the original SLN after deletion.*

(4) *If the semantic link to be added is in the closure of the SLN, the SLN equals to the original SLN after addition.*

(5) *Deleting a semantic link and no other accompany operations and then adding the same semantic link between the same pair of nodes does not change the SLN*.

(6) *Adding a semantic link and then deleting it may change the minimum semantic cover of the SLN, therefore the SLN is changed.*

(7) *If a semantic link l in the minimum semantic cover can be derived out by relational reasoning* $l_1 \times l_2 \rightarrow l$, *and l's determining set does not include* l_1 *or* l_2, *l should not be in the minimum semantic cover*.

(8) *Any removed semantic link in the minimum spanning graph cannot be derived out from the graph any more.*

The following are some concepts and characteristics on the spanning graph of SLN.

Definition 2.10.2. The *spanning graph of a semantic link l* is a subgraph of the minimum semantic cover and its closure includes *l*.

Definition 2.10.3. The *minimum spanning graph of a semantic link M* is a subgraph of the minimum semantic cover such that its closure includes *M*, and there is no smaller subgraph whose closure includes *M*.

A semantic link may have several minimum spanning graphs, and these minimum spanning graphs may have intersection.

No more semantic links can be derived from the minimum spanning graph of SLN.

2.11 SLN Abstraction

Abstraction plays an important role in the transition from the perceptual image of objects to rational thinking about them. Different individuals may come to different conclusions when abstracting from the same perceptions. Such differences can produce diversity of knowledge, and promote healthy evolution from the viewpoint of ecology.

Abstraction on concepts carries out in the classification hierarchy of the semantic space of SLN. Abstraction on semantic nodes and abstraction on semantic links are based on abstraction on concepts. Therefore, abstraction between SLNs is based on the abstraction on concepts.

Definition 2.11.1 The structure of a semantic link network $SLN'=<N', L'>$ is called an abstraction of $SLN=<N, L>$ if and only if there exists an onto mapping A: $<N, L> \rightarrow <N', L'>$ such that for any semantic link n_i—l→n_j of SLN, there exists a corresponding link $A(n_i)$—$A(l)$→$A(n_j)$ of SLN', $A(n_i)$ and $A(n_j)$ are the abstraction of n_i and n_j respectively in the classification hierarchy, and, $A(l)$ is the same as l or the super-class of l.

The abstraction of $A(n)$ from n concerns the following three cases:

(1) Both n and $A(n)$ are atomic concepts. In this case, node $A(n)$ is the abstraction of n if and only if $A(n)$ is the parent of n in a conceptual abstraction hierarchy.
(2) Node n is an SLN and $A(n)$ is an atomic concept. In this case, node $A(n)$ is the abstraction of the SLN.
(3) Both n and $A(n)$ are SLNs. In this case, the abstraction between the nodes becomes the abstraction between SLNs (definition 2.11.1).

From the above definition, we can define an abstraction operation $\cap_A$ that generates an abstract SLN from n SLNs: $\cap_A (SLN_1, SLN_2, \ldots, SLN_n) = SLN$, where A stands for the topic area of the abstraction. The algorithm for realizing this operation is based on the intersection operation of SLNs but using abstraction as outlined above to replace equivalence.

For a given set of SLNs as the original source of abstraction, repeating the following steps can generate an abstraction tree as shown in Fig.2.11.1:

(1) Generate a set of abstract SLNs by abstracting any combination of source SLNs.
(2) Generate the source set at a new level by clustering the abstract SLNs that are similar to each other.
(3) If the result has more than one abstract SLN, repeat from step 2.

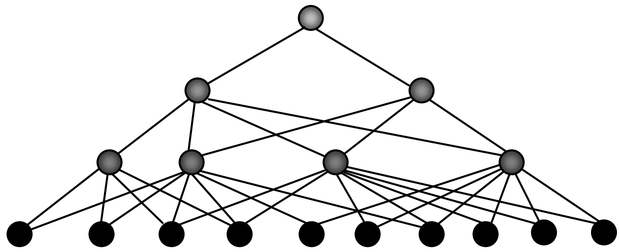


Fig. 2.11.1 An abstraction tree.

Abstraction usually works with analogy (H. Zhuge et al, Analogy and Abstract in Cognitive Space: A Software Process Model, *Information and Software Technology*, 39(1997)463-468). Fig. 2.11.2 shows the relationship between analogy and abstraction. While abstractions are carried out on sets of objective existence, analogy can provides links and references between abstractions.

The abstraction process is a process of semantic selection. The abstraction tree reflects the epistemology of the topic area. If the abstraction tree is derived from many diverse SLNs, and adding new SLN abstractions doesn't change the tree, then the epistemology of the topic area is completed. Once completed, the abstraction tree can be independent of the SLNs and evolve during use.

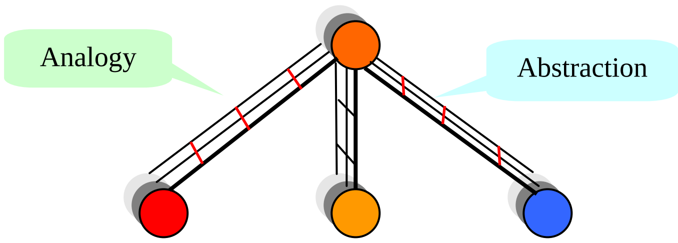


Fig.2.11.2 Abstraction with analogy.

A *semantic selection* is of the form $\alpha \leq_C \beta$, which means that the topic area selects β from candidate set C as more abstract than α. A set of semantic selection results is generated in one's mind as his/her epistemology is formed.

Given an epistemology, the abstraction of an SLN can be automatically generated from the epistemology. The abstraction tree and the selection set constitute the epistemology of the topic area. An epistemology can guide the computation of the abstract SLN of a given SLN. The formation of a community's epistemologies from their common SLNs reflects a kind of social selection process based on shared values.

A society consisting of resources, resource producers and consumers can build a mutual understanding between individuals from knowledge of each others' epistemologies. If the epistemologies can be processed by machine, mutual understanding will thereby be easier to attain.

Now, we can view a scenario of resource management in the Knowledge Grid environment: The resources' providers (humans or machines) need to attach their epistemologies to the resources. The resource consumers (humans or machines) can use and understand the resources by using the attached epistemologies. Each consumer is also a provider. Therefore, the environment needs to manage the epistemological attachments in addition to resources. With the attachment, service providers like search engine will be able to provide not only the links and documents but also multi-facet navigation and explanations.

Fig. 2.11.3 depicts interaction between providers and consumers through multi-layer abstractions. Images are the direct reflection of the physical space. Texts are a kind of abstraction, which indicate certain semantics. Attributes can be extracted from entities or sensed directly from the physical space. Hyperlinks and semantic links establish the relevancy between entities. High-level knowledge includes models and methods.

Current research into ontology seeks to establish consensus between different understandings in domains. The effectiveness of this effort is questionable if we look at the nature of the World Wide Web and the background of the evolution of culture and emerging new domains.

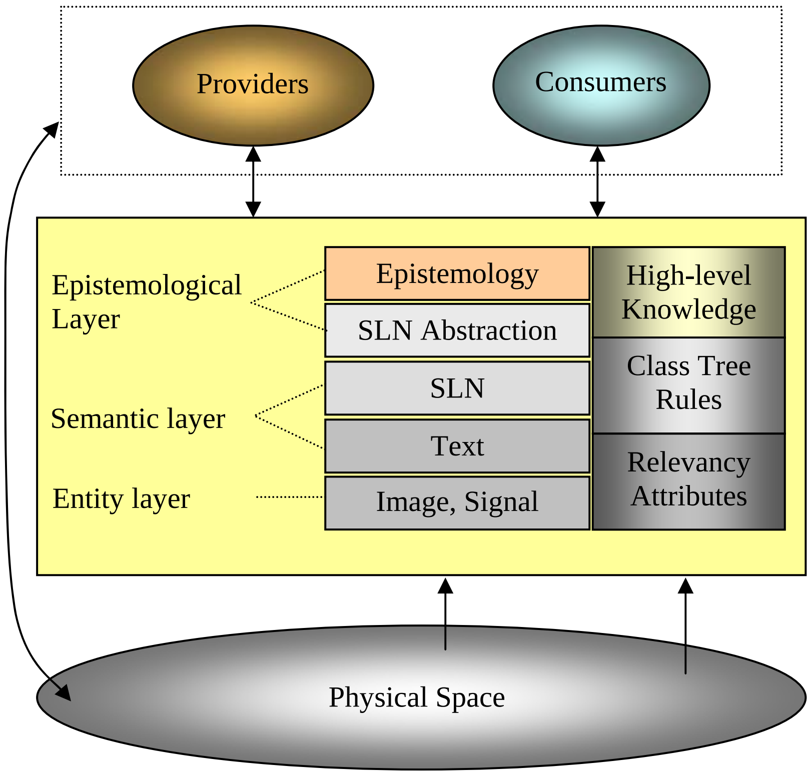


Fig.2.11.3. Interaction through multi-layer abstractions.

Epistemology as described here respects differences between individuals in a social context. All participants construct their own epistemologies. This is in line with the diversity requested by the Knowledge Grid environment.

Epistemology will have an important part to play in establishing the future Cyber-Physical Society.

SLN pursues diversity and emphases on dynamicity rather than correctness and accuracy. It enables smart applications to work in self-organized semantic context.

2.12 Application: SLN-based Image Retrieval

Semantics-based image retrieval has the following advantages:

(1) The retrieval result is a semantic clustering of relevant images rather than a simple list of images put out by current search engines.
(2) Users can browse images by wandering along semantic hints with supporting semantic reasoning and explanation.

Effective retrieval of images from the Web has attracted many researchers. The three main approaches to Web image retrieval are the text-based, the content-based, and the hyperlink-based (V. Harmandas, et. al., Image retrieval by hypertext links, *ACM SIGIR*, 1997; R. Lempel and A. Soffer, PicASHOW: Pictorial Authority Search by Hyperlinks on the Web, *WWW Conference*, 2001).

The text-based approach applies text-based Information Retrieval (IR) algorithms to keywords in annotations of images, captions of images, text near images, the entire text of pages containing an image and filenames. These approaches support a specific natural language for queries.

The content-based approach applies image analysis techniques to extract visual features from images. The features are extracted in a preprocessing stage and stored in the retrieval system's database. The extracted features are usually of high dimensionality, and need fewer dimensions to allow scalability.

The hyperlink-based approach makes use of the link structure to retrieve relevant images. The common basic premise is that a page displays or links to an image when its author considers the image to be of value to the viewers (H.Zhuge, Retrieve Images by Understanding Semantic Links and Clustering Image Fragments, *Journal of Systems and Software*, 73(3)(2004)455-466).

These approaches are almost independent of the semantics of the image itself. The main obstacle to semantics-based image retrieval is that it is hard to describe an image semantically. But the semantics of an image can be implied by related images and their semantic relationships. Thus, semantic links can help realize semantic image retrieval.

The following three sets of semantic links list position relationships:

(1) *X is-above Y*; *X is-below Y*; *X is-left-of Y*; and, *X is-right-of Y*
(2) *X is-north-of Y*; *X is-south-of Y*; *X is-east-of Y*; and, *X is-west-of Y*.
(3) *X is-ahead-of Y*; and, *X is-behind Y*.

The positional links can be described in a *simple grammar* as *X is-β-of Y*, where $\beta \in$ {*above*, *below*, *left*, *right*, *north*, *south*, *east*, *west*, *ahead*, *behind*} is called a semantic indicator. The pairs *above*/*below*, *left*/*right*, *south*/*north*, *east*/*west*, *ahead*/*behind* are each symmetric.

Other positional links can be composed from two different semantic indicators, depending on their meaning. For example, *X is-north-west-of Y*; *X is-south-east-of Y*; *X is-south-west-of Y*; and, *X is-north-east-of Y* are meaningful semantic links, but *X is-north-south-of Y* isn't. In general, the composite links can be described as: *X is-α-β-of Y*.

Rules can be derived from the positional links as shown in Table 2.2. From the rules in the table, we can derive two generalization rules and one transitivity rule.

Generalization Rule 1. X *is*-α_1-α_2-*of* $Y \Rightarrow Y$ *is*-β_1-β_2-*of* X if and only if
(1) *is*-α_1-α_2-*of* and *is*-β_1-β_2-*of* are both meaningful;
(2) X *is*-α_1-*of* $Y \Rightarrow Y$ *is*-β_1-*of* X; and,
(3) X *is*-α_2-*of* $Y \Rightarrow Y$ *is*-β_2-*of* X.

A relation is meaningful if it has a corresponding concept in the mental space.

Generalization Rule 2. If α and β are mutually symmetric, then X *is*-α-*of* $Y \Rightarrow Y$ *is*-β-*of* X.

Transitive Rule. X *is*-α-*of* Y and Y *is*-α-*of* $Z \Rightarrow X$ *is*-α-*of* Z.

Table 2.2 Rules for positional links.

No	Rule	Category
1	X *is-above* $Y \Rightarrow Y$ *is-below* X	X *is*-α-*of* $Y \Rightarrow Y$ *is*-β-*of* X
2	X *is-below* $Y \Rightarrow Y$ *is-above* X	ditto
3	X *is-left-of* $Y \Rightarrow Y$ *is-right-of* X	ditto
4	X *is-right-of* $Y \Rightarrow Y$ *is-left-of* X	ditto
5	X *is-north-of* $Y \Rightarrow Y$ *is-south-of* X	ditto
6	X *is-south-of* $Y \Rightarrow Y$ *is-north-of* X	ditto
7	X *is-east-of* $Y \Rightarrow Y$ *is-west-of* X	ditto
8	X *is-west-of* $Y \Rightarrow Y$ *is-east-of* X	ditto
9	X *is-ahead-of* $Y \Rightarrow Y$ *is-behind* X	ditto
10	X *is-behind* $Y \Rightarrow Y$ *is-ahead-of* X	ditto
11	X *is-north-west-of* $Y \Rightarrow Y$ *is-south-east-of* X	X *is*-α_1-α_2-*of* $Y \Rightarrow$ Y *is*-β_1-β_2-*of* X
12	X *is-south-east-of* $Y \Rightarrow Y$ *is-north-west-of* X	ditto
13	X *is-south-west-of* $Y \Rightarrow Y$ *is-north-east-of* X	ditto
14	X *is-north-east-of* $Y \Rightarrow Y$ *is-south-west-of* X	ditto

Orthogonal semantic relations exist between positional indicators. We use $\alpha_1 \perp \alpha_2$ to denote that relation α_1 is orthogonal to relation α_2. Such orthogonal relationships indicate the characteristics of concepts that help image retrieval and assist high-level applications on images.

The following are six sets of orthogonal relationships determined by the spatial concepts in mind:

(1) *below* $\perp$ *left*, *below* $\perp$ *right*, *above* $\perp$ *left*, and *above* $\perp$ *right*;
(2) *south* $\perp$ *west*, *south* $\perp$ *east*, *north* $\perp$ *west*, and *north* $\perp$ *east*;
(3) *behind* $\perp$ *left*, *behind* $\perp$ *right*, *behind* $\perp$ *above*, *behind* $\perp$ *below*;
(4) *behind* $\perp$ *south*, *behind* $\perp$ *east*, *behind* $\perp$ *west*, *behind* $\perp$ *north*;
(5) *ahead* $\perp$ *south*, *ahead* $\perp$ *east*, *ahead* $\perp$ *west*, *ahead* $\perp$ *north*; and,
(6) *ahead* $\perp$ *left*, *ahead* $\perp$ *right*, *ahead* $\perp$ *above*, *ahead* $\perp$ *below*.

Two more rules follow from these orthogonalities.

Symmetry Rule. If $\alpha_1 \perp \alpha_2$, then we have $\alpha_2 \perp \alpha_1$.

Orthogonal Rule. In a two-dimensional space, if $\alpha_1 \perp \alpha_2$, $\alpha_3 \perp \alpha_1$ and $\alpha_4 \perp \alpha_2$, then $\alpha_3 \perp \alpha_4$.

The position of an image can be determined by using the following rule.

Position Determination Rule. The position of image *X* can be determined by two positional links *X is*-α_1-*of A* and *X is*-α_2-*of B* if and only if α_1 and α_2 are positionally orthogonal.

Besides the semantic and positional links discussed so far, the following *existence semantic* links also help cluster images.

(1) *X is-coincident-with Y*;
(2) *X is-not-coincident-with Y*;
(3) *X is-compatible-with Y*;
(4) *X is-not-compatible-with Y*; and
(5) *X is-complementary-to Y*.

Further, experience with layout, hierarchy, importance and relevance also help image retrieval. Capturing such experience for computation is a challenge.

The semantic link network approach can be used for coordinating and fusing various contents by incorporating specific rules, because the contents of resources are not isolated. The key is to find the semantic links between the contents. Semantics-based content fusion is an important intelligent mechanism.

Actually, semantics was often lost, distorted, or hidden in texts when they were made. So, discovering hidden semantics, correcting distortion, and inferring semantics is also a challenge.

2.13 Application: Active Document Framework ADF

Humans create symbol languages and rely on them to indicate semantics — *the mapping between the symbol space, the physical space, and the social space (including the mental space)*. An active semantic link network of concepts can simulate some functions of minds. Readers rebuild the semantics of authors by reading the symbols in passive documents.

People need to learn languages to build the ability of using symbols. But symbols do not uniquely indicate physical objects and concepts. Although grammars and sentences can be used to raise the accuracy of indication, readers often misunderstand authors.

Current e-documents are as passive as hardcopy documents. People need to use a search engine to retrieve documents of possible interest and to browse document(s) page by page in the same way as reading the hardcopy documents. Passive documents have two shortcomings: first, the user will feel it is difficult to efficiently get appropriate content when browsing a large document (especially large-scale documents); second, the search engine does not have any context about the document, which prevents it from providing an ideal content service.

An active document (*AD*) *is a mechanism that can help humans to efficiently construct, maintain, read and interact with the active documents, which are based on semantic link networks of language components (words, sentences, and paragraphs). It includes interactive interfaces for constructing, maintaining, and browsing the semantic link networks of language components and a set of engines for implementing intelligent behaviors. It can interact with each other and with humans through diverse roles. The interactions wave semantic link networks of active documents, which provides the ground for active documents.*

As compared in Fig. 2.13.1, active documents can better reflect semantics as they represent various relations in document and can uncover more semantics and interact with author and reader.

Interactions between active documents enable an active document to accurately provide contents for readers, from itself SLN, from other active documents through interacting with them, and from the authors or other readers through interacting with humans. Interactions will also form and evolve communities of active documents, which provide the semantic environment for semantic interaction.

In an active interactive environment, various advanced applications can be built on the basis of active documents.

An *AD* can be regarded as a function of the *input requirement* (denoted by *I*), the *SLN* of language components, and a repository of *engines* (denoted by *E*) as follows:

$$O = AD\ (I, SLN, E).$$

AD can search for new engines and put them into the engine repository. An *AD* includes the following basic engines.

(1) The *search engine*, responsible for searching and analyzing documents, extracting the components, and adding them to SLN.

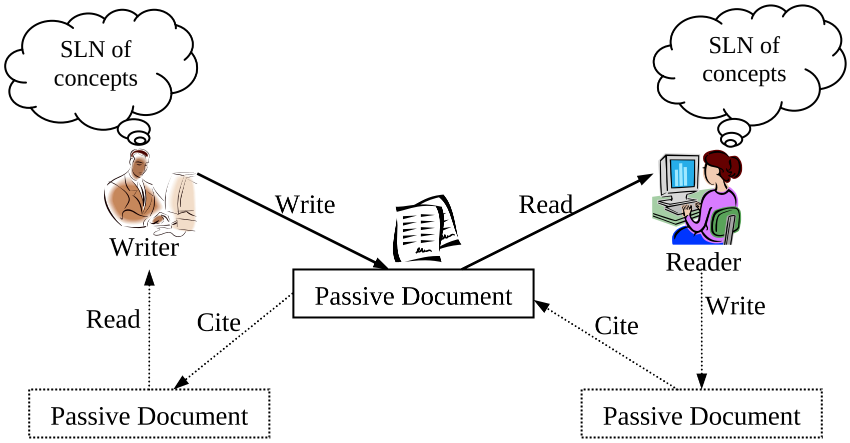
(2) The *execution engine*, responsible for zooming and displaying SLN components with an integrated content on a certain topic according to requirement.
(3) The *interaction engine*, responsible for interacting with the user and the other active documents.
(4) The *reasoning engine*, responsible for reasoning and explanation.
(5) The *management engine*, responsible for managing SLN and the repository of engines.

Active Document Framework ADF helps humans to easily create active documents. A model of ADF is discussed in "Active e-Document Framework ADF: Model and Tool" (H. Zhuge, *Information and Management*, 41(1)(2003)87-97).

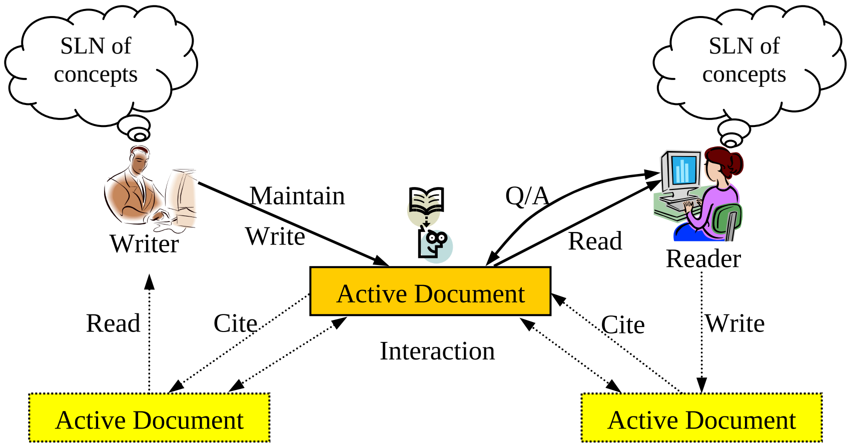
To implement the idea of ADF, designing the engines for easily establishing and maintaining the semantic link network and the Interactive Semantic Base are the key (H.Zhuge, Interactive Semantics, *Artificial Intelligence*, 174(2010)190-204).

We can foresee the scenario of ADFs in the Knowledge Grid environment through the following behaviors (H.Zhuge, Clustering Soft-devices in the Semantic Grid, *Computing in Science and Engineering*, 4(6)(2002)60-62):

(1) Designers create and maintain active documents.
(2) Users publish requirements through an human-machine interface.
(3) Active documents find requirements.
(4) Users interact with active documents to learn content by browsing through semantic links, add semantic links on documents under constraints, or express further requirements.
(5) Active documents interact with each other.



A library of passive documents



An interactive environment of active documents

Fig. 2.13.1 Comparison between passive document and active document.

2.14 Application: e-Learning

An e-learning system using the semantic link network to organize learning resources has the following advantages:

(1) Learner's profile can be described in a semantic link network of concepts. A concept is a basic semantic image of behaviors, events, or resources, or an abstraction of a class of semantic images. The ranks of concepts evolve with the operations on the interested learning resources and the network. The ranks of the concepts corresponding to the often operated resources will become higher.
(2) Learners can be provided with a semantic map of learning resources while operating the system, not only at the instance level but also at multiple abstraction levels (regard nodes as classes). Reasoning mechanisms can help users to know implicit links and foresee the result of operations.
(3) Learners can be provided with not only the required learning resources but also the reference resources.
(4) The system can explain a learning resource through diverse reasoning.
(5) The relational, inductive, and analogical reasoning can help understand implicit semantic links, and can inspire creative thinking and broaden knowledge of learners.
(6) Learners can know the formation process of a semantic link network of learning resources by recording the evolution history of the network.
(7) The learning processes can be analyzed by retrieving the formation processes of the learners' individual semantic link networks. The learning processes can be improved by discovering the influence of operating the semantic link networks (e.g., adding or deleting a semantic link). It is useful for learners to know which semantic links or nodes are influenced by operations.

Semantic links can help learners to know learning contents from multiple aspects and at different levels, for example: the *causeEffect* link can help learners to know the effect when they know the cause, and know the cause when they know the effect.

Semantic links can be chained and compared for relational reasoning and analogical reasoning. The *similar* link can help learners to know similar contents when learning. The *sequential* link can help arrange learning contents sequentially step by step. The *reference* link can help learners get appropriate references about the learning contents. The implication link can help learners to know the underlying content about the current learning resources (H.Zhuge, Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering,* 21(6)(2009)785-799).

Discovering semantic communities and implicit links in the semantic link network of learning resources can provide an e-learning system with the following functions:

(1) Enable learners to focus on learning relevant contents within a semantic community that matches interests, which are rendered by the often indicated concepts.
(2) Explain or indicate a concept through diverse semantic links.
(3) Explain a semantic link network of learning resources by its community hierarchy.
(4) Explain a semantic community by its minimum semantic cover and maximum spanning tree.
(5) Filter out the special SLN with the links of definite types to match learning interest.
(6) Expand the given concept to the linked concepts through given semantic links.
(7) Automatically generate learning paths and find the shortest path.
(8) Automatically cluster and recommend learning resources according to the emerging semantic communities and links.
(9) Discover local emerging semantic communities, and reflect the law of forming and evolving semantic communities.

As depicted in Fig.2.14.1, extending the learning system from the cyber space to the cyber-physical society can obtain more advantages.

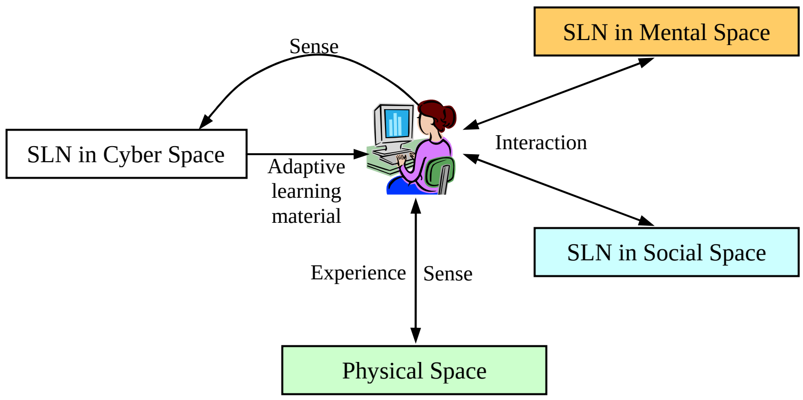


Fig. 2.14.1. Panoramic learning in Cyber-Physical Society.

For example, the e-learning system can search and organize learning resources not only in the cyber space but also in the social space and the physical space, and provide the learning resources in rich forms according to the characteristics of the learner. The system can sense the expression, location and performance of the learner through sensors so that it can adapt to change and provide appropriate learning resources for the learners at appropriate time, at appropriate location, and in appropriate space.

A learner can interact with people to obtain indirect experience through social network. As the consequence, his/her profile will be updated, and the cyber space will also be updated. He/She can also experience in the wider scope in the physical space through sensors. His/her mental space reflects the learning resources and fuses them with the active SLN in the mental space. So, learners can obtain maps of contents rather than isolated learning resources, and can learn in a panoramic learning environment.

2.15 Potential Applications, Relevant Work and Q&A

There is great scope for the SLN to develop.

When nodes and links only represent concepts and cause-effect relations respectively, SLN is reduced to fuzzy cognitive maps (Z.Q. Liu and R. Satur, Contextual Fuzzy Cognitive Maps for Decision Support in Geographic Information Systems, *IEEE Transactions on Fuzzy Systems*, 7(10)(1999)495-502; Y. Miao, et al., Dynamic Cognitive Network, *IEEE Transactions on Fuzzy System*, 9(5)(2001)760-770; H. Zhuge and X. Luo, Automatic Generation of Document Semantics for the e-science Knowledge Grid. *Journal of Systems and Software*, 79(7) (2006)969-983). As there is only one type of relation in the fuzzy cognitive map, reasoning on fuzzy cognitive maps can be easily carried out by matrix multiplication. However, the concepts in the fuzzy cognitive maps are indicated by simple words, and there is no semantic space to regulate their semantics. In the semantic link network, words are regarded as the indicators of concepts and semantic images in mind. Section 2.20 will discuss the mental concept in detail.

SLN can also be reduced to other models. As an abstraction model and method, the study of SLNs can help the study of other models.

SLN is also an approach to realizing a semantics-rich Web. It is significant in method and theory, but its real application relies on the transition to SLNs from current standards (www.w3c.org) as well as on industry efforts to establish standards for semantic grounds.

More applications of SLN are worth mentioning:

(1) Use of SLN to organize useful resources and support cooperative research and learning. Applications have shown that it is feasible to use SLN to organize teaching materials to support interactive and adaptive learning.

(2) Use of SLN to organize resources in a semantics-rich network to realize semantic peer-to-peer resource management (H. Zhuge, et

al., Query Routing in a Peer-to-Peer Semantic Link Network, *Computational Intelligence*, 21(2)(2005)197–216).

(3) Use of SLN to organize resources to support relational query (H. Zhuge, Autonomous Semantic Link Networking Model for the Knowledge Grid, *Concurrency and Computation: Practice and Experience*, 19(7)(2007)1065-1085).

(4) Prediction of the missing semantic links in some semantic link networks similar to the work of predicting missing links (A. Clauset et al., Hierarchical Structure and the Prediction of Missing Links in Networks, *Nature*, 453(2008)98-101).

In general, SLN is a self-organized and evolving space of semantic links.

As a knowledge representation approach, a semantic net is a network that represents semantic relations between concepts. It is first invented by R. H. Richens in 1956, developed by R. F. Simmons, A.M. Collins and colleagues in 1960s (A. M. Collins and M.R. Quillian, Retrieval Time from Semantic Memory. *Journal of Verbal Learning and Verbal Behavior*, 8 (2) (1969) 240–248; A. M. Collins and E. F. Loftus, A Spreading-Activation Theory of Semantic Processing. *Psychological Review*, 82 (6) (1975) 407–428). It was proposed as a human associative memory model and applied to natural language processing in the early 1970s. Some expert systems adopted semantic net as their knowledge representation mechanism. Reasoning on semantic network/net is based on the matching between semantic nets.

From the basic expression, the subject-predicate-object triple expressions of the Resource Description Framework (RDF, www.w3.org/RDF/) is similar to the semantic net and the classic conceptual modeling approaches such as Entity-Relationship (ER model) or class diagrams. The subject denotes the resource, and the predicate expresses a relationship between the subject and the object. The representation of RDF is based on the Extensible Markup Language (XML, www.w3.org/XML/), which facilitates the exchange of data cross platform. RDF can be regarded as the development of the semantic network in Web platform. RDF syntax was recommended by

W3C in 2004.

The Web Ontology Language OWL is for use by applications that need to process the content instead of just presenting text to humans. It provides additional vocabulary along with a formal semantics. It has three increasingly-expressive sublanguages: OWL Lite, OWL DL, and OWL Full (www.w3.org/TR/owl-features/).

Standardized in 1999, the topic map that organizes information by using the concepts of *topics*, *occurrence* and *association* (www.topicmaps.net/pmtm4.htm; P. Auillans, et al., A Formal Model for Topic Maps, *Proc. 1st International Semantic Web Conference*, LNCS 2342(2002)69-83).

In database area, the foreign key establishes reference relation between tables. A foreign key is a field in one table that matches a candidate key of another table. It actually links tables to provide wider access scope of data.

Linked Data is a technique proposed by Tim Berners-Lee for exposing, sharing, and connecting pieces of data on the Semantic Web using URIs and RDF. There are many projects on this topic. Linked Data can be regarded as the extension of the foreign key in database on the Web.

The following are some questions and answers about SLN.

Question 1: It is difficult to determine the exact semantic link between resources sometimes.

Answer: The first cause is that people do not really recognize the resources, e.g., the meaning of paintings. The second cause is that using a simple word to indicate a relation is difficult sometimes. To overcome this difficulty, we can use a set of words to indicate one relation just like people use a set of words to search web pages. Words could be restricted by simple program. Different people could use different sets of words to indicate the same link to support the diversity of viewpoint and culture. The third cause is that the explanation of a semantic link sometimes is difficult. Chapter 8 will introduce a decentralized networking system making use of implicit semantic links.

Question 2: People consciously or subconsciously use versatile semantic links, but people usually are not specialized in determining the semantic links between resources.

Answer: The first cause is that there may exist many semantic links on multiple facets between two objects, people may not know or cannot decide immediately the exact relation in certain situation. The second cause is that different people may use different words to indicate the same relation.

Question 3: It is not easy to determine the general reasoning rules on semantic links.

Answer: Most semantic links and rules are domain specific, so only those who are familiar with the domain can establish the SLN schema. If different users define different rules, the consistency between the rules should be checked.

Question 4: There are several approaches such as semantic net and RDF, do we still need the Semantic Link Network?

Answer: First, scientific exploration is endless, new models will emerge with the development of IT technology. Second, *developing a single dominant model is not the right development trend of technology*. The Knowledge Grid in the Cyber-Physical Society will allow diverse technologies co-exist and co-evolve. SLN is a promising semantic model that could co-exist with other semantic models. Third, SLN is not only a technique but also a methodology of modeling the self-organized social space.

Question 5: What technologies can be used to study and develop the Semantic Link Network?

Answer: As a theoretical framework and method, SLN will develop with the evolution of the Web and the major advanced networking systems. Research on real social networks will significantly influence the development of the SLN. Research on culture in the Cyber-Physical Society will influence the evolution of the SLN.

Question 6: *How to establish a semantic link network on the Web*?

Answer: The Web is designed for humans to read. Before clicking the link, people can guess the meaning from the words (anchor's content) on the link in the displayed text, but the links in the linked page is unknown. It will be very helpful if a semantic map (an abstract semantic link network) of the linked page can be seen when people point to the link. Furthermore, an automatic navigation system can be designed to guide people to browse the interested contents.

A more advanced system can self-organize the interested contents through semantic links and display the contents in a meaningful order. Developing a new Web browser based on this idea can fundamentally improve the current approaches (browse and search) to accessing the Web and completely change user experience.

The following steps constitute one solution to implement a semantic link network on the Web:

(1) Discovering the semantic links between Web pages within the selected Web site.
(2) Adding rules to the semantic link network, and using standard rule representation like RuleML (http://ruleml.org/) to facilitate rule exchange on the Web.
(3) Building classification, manually or by online extraction, for example, from Wikipedia (www.wikipedia.org) and ODP (www.dmoz.org).
(4) Linking semantic indicators to classes.
(5) Carrying out reasoning to derive implicit semantic links.
(6) Discovering communities.
(7) Displaying Web pages according to the communities and the semantic links.

2.16 SLN 2.0: Autonomous Semantic Data Model

A semantic data model is an abstraction of the real world by defining the relations between data. The types of data and the types of relations between data are predefined and regarded as commonsense. It is an important research topic in database and software engineering (M.Hammer and D.McLeod, The Semantic Data Model: A Modeling Mechanism for Database Applications, *SIGMOD*78).

Usually, a semantic data model like the ER model (P.Chen, The Entity-Relationship Model: Toward a Unified View of Data, *ACM Transactions on Database Systems*, 1(1)(1976)9-36) is established at the analysis stage. It should be fixed before the design stage as it will be transformed into the data structure of the information system, and it should not be changed at the system development stage and execution stage. A certain change of domain business will lead to the failure of the system. This is the main cause of low success rate of information system development. Traditional static semantic data models are for closed systems, which cannot reflect the dynamic nature of society.

A graph database uses graph structures with nodes, edges, and properties to represent and store information. It is often faster for associative data sets, and it is more directly relevant to the object-oriented data model (M.Gyssens, et al., A graph-oriented object database model, *IEEE Transactions on Knowledge and Data Engineering*, 6(4)(1994)572-586). It can scale more naturally to large data sets as they do not require expensive join operations. Avoiding rigid schema, they are more suitable to manage ad-hoc and changing data with evolving schemas. Relational databases are good at performing repeat operations on large numbers of data elements. Graph database supports graph queries such as sub-graph, reachability, and the shortest path between two nodes.

Graph data model research can trace to the early network data model, i.e., Data Structure Diagram (C.Bachman, The Evolution of Storage Structures, *Communications of the ACM*, 15(7)(1972)628-634;

C.Bachman, The Programmer as Navigator. ACM Turing Award Lecture. *Communications of the ACM*, 16(11)(1973)653-658).

Social network research provides more graph features for graph data models to support advanced queries, for example, about the degree distribution, diameter, and community (S.Wasserman and K. Faust, Social Network Analysis: Methods and Applications, Cambridge University Press, 1994).

However, database based on graph models has the following major limitations:

(1) *It neglects the semantics in graph. Although it can query the features of the graph, it knows little semantics of the graph. Therefore, it is limited in ability in supporting queries on semantics of the graph.*
(2) *It does not support query on implicit relations.*
(3) *It does not know the effect of operating graph.*
(4) *It does not support abstraction on graph.*
(5) *It does not support relational and analogical reasoning on graph.*

A social SLN is dynamic and open. Any semantic node and semantic link can be added to and removed from the network. Users are diverse, and there is no obvious difference between users and designers — *they are all explorers in essence*. Users can maintain the semantic link network. The temporal links reflect the control flows through semantic links. Open domain applications usually do not need global central control.

The semantics of the SLN evolves with the evolution of the network structure. In the meanwhile, the mental spaces of users, designers and programmers change. The users' interests and the content of queries will be also changed. Therefore, the application system including the interface needs to be adaptive.

Fig. 2.16.1 compares the static semantic data model and the dynamic semantic data model. The static data model usually needs to be transformed into a data structure like relational table. The dynamic semantic data model allows systems to directly operate on the data model. The self-organized model allows users to define and maintain individual SLNs. Queries will be routed in a peer-to-peer network when individual SLN is not able to answer.

The following discussions support the implementation of an autonomous semantic data model.

Semantic Distance and Extensible Concept Hierarchy. The semantic distance between two classes (concepts) in a classification tree can be defined as the sum of their distances to the nearest common ancestor. From the root, the first-level of the classification tree regulates commonsense, and the second-level regulates domain commonsense. Users are encouraged to use commonsense to indicate semantics. They can use their own keywords to indicate the semantics of the semantic nodes and semantic links by extending the classification. Individual indicators should be given individual classes or instances, which can be up-graded to commonsense in terms of its popularity during using or confirmation from authorities.

Metric Space. It values the semantic nodes and semantic links as well as the possibly attached attributes. The value of a semantic link is in positive proportion to the following three factors: (a) the values of its two ending nodes; (b) the times of its occurrence in SLN; and, (c) the times it participates in reasoning. The value of a semantic node is in positive proportion to the values of its neighbor nodes. The metric space is also responsible for measuring the probability over the SLN.

Abstract SLN. An abstract SLN reserves the abstract semantic nodes, abstract semantic links, and the rules defined in the semantic space.

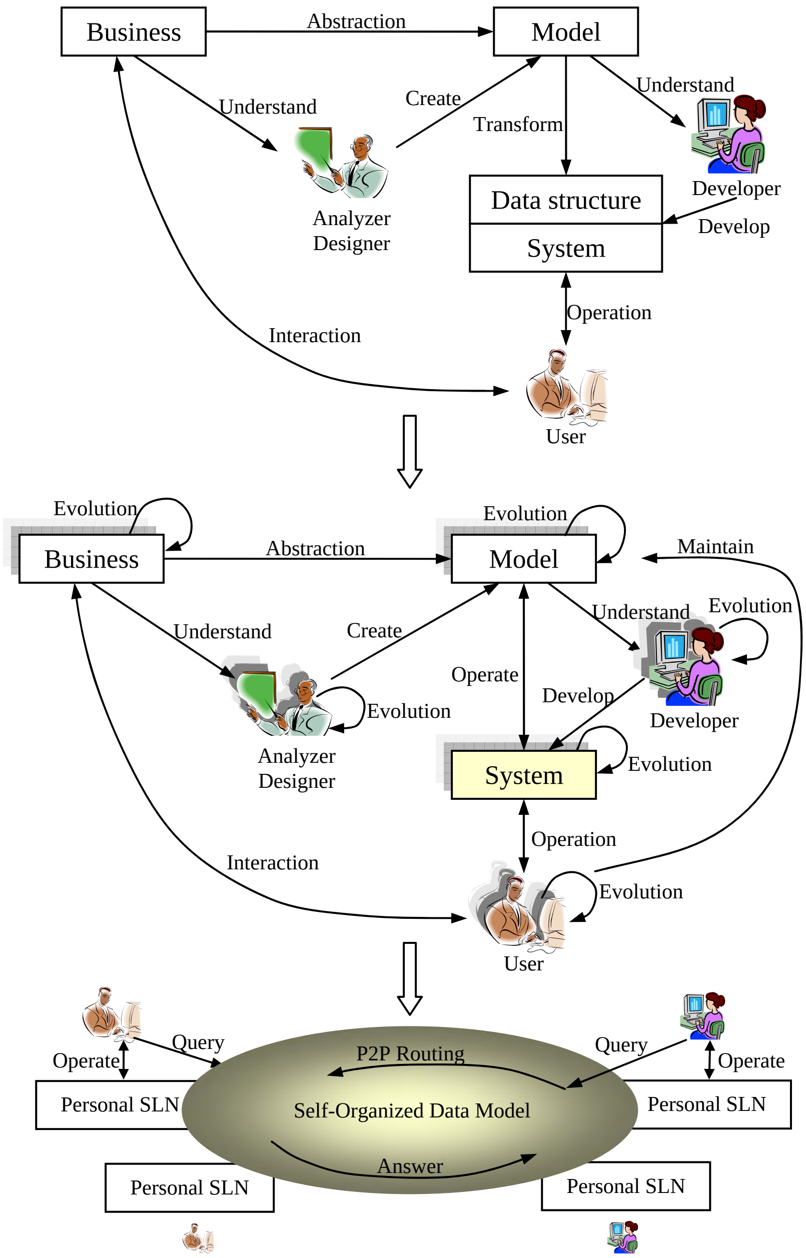


Fig. 2.16.1. Semantic data models: from static to dynamic.

Instances. An instance SLN consists of instances of semantic node types and instances of semantic link types. An abstract SLN can generate several SLN instances by instantiating its semantic nodes and semantic links.

The schema of relational database defines the structure of database, which consists of a set of relations with attributes and the dependencies between attributes. The relational schema can be normalized to ensure consistency, non-redundancy and efficiency. A form of SLN schema is defined in (H. Zhuge, Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering*, 21(6)(2009)785-799).

SLN Schema is a triple denoted as *SLN-Schema* = <*NodeTypes*, *LinkTypes*, *Rules*> under the given *concept hierarchy* in the domain. *NodeTypes* is a set of resource types, each of which is represented as *NodeType* = [*name*: *field*] | [*name*: *field*, … , *name*: *field*]. *LinkTypes* is a set of semantic link types belonging to $NodeTypes \times NodeTypes$, each of which is represented as *name*(*linkType*)(*NodeType*, *NodeType*). *Rules* is a set of reasoning rules on *LinkTypes*, denoted as $Rules=\{\alpha\cdot\beta\Rightarrow\gamma \mid \alpha, \beta, \gamma \in LinkTypes\}$. The field can be defined by the basic data type, classification trees, or rules in the semantic space.

A schema view is a subset of the node type, the link type, and the rules of the schema.

Maintenance operation of schema includes appending or deleting node type or link type. Deleting a link type and node type in the schema should check whether there are corresponding instances exist, or whether there is corresponding node type and link type in the schema view.

A domain SLN schema reflects consensus on the basic semantics of a domain. Users can append SLN instances to the system by instantiating the schema or a schema view. Reasoning on instances is based on the reasoning rules defined in the schema or schema view.

Another form of schema was defined in (H. Zhuge and Y. Sun, The Schema Theory for Semantic Link Network, *Future Generation Computer Systems*, 26(3)(2009)408-420). The algorithms for SLN schema extension and reduction and reasoning algorithms for deriving more semantic links were introduced.

A decentralized, easily extensible data management architecture in which any user can contribute new data, schema information, and mappings between other peers' schemas was proposed (A.Y.Halevy, et al., Schema Mediation in Peer Data Management Systems, *ICDE 2003*, USA).

The following are two strategies to construct an SLN as a data model.

Schema-based Strategy. Create SLN schema first and then instantiate it by giving the names and values of semantic nodes and semantic links. This strategy requires users to contribute to and share the same schema. This also implies that users have consensus on the primitive semantic space**.**

Appending a new node (e.g., "XiaopingSun") to or deleting a node in the SLN is implemented by determining the type of the node (e.g., "person"), selecting a semantic link type from the schema and linking the new node to the existing node (e.g., "Hai Zhuge"). If there is no corresponding link type or node type, the user should create a *personal schema view*, and append the new link type or the new node type to the schema first. Then, the user can publish the new link type, node type, or rules. If most users can use them, they can become the elements of public schema.

SLN schema is useful in defining SLN for closed domain applications. But, it is not appropriate to define a rigid schema in open domain, especially, for self-organized and frequently changing applications.

Self-Organized Strategy. Users create their individual schemas, and then define SLN instances. Users can only maintain their own schemas. There are two ways to realize cooperation between individual schemas: (1) Individual schemas work in peer-to-peer paradigm. (2) Creating a global schema by unifying individual schemas. The global schema needs to be maintained when any individual schema is changed.

No-Schema Strategy. Users freely link nodes one another. Semantic communities can be discovered to limited operations. Change is limited bottom-up from the semantic community hierarchy. Change is firstly limited within semantic communities. If an operation leads to the damage of a community, the change will be limited within the corresponding up-level community. New communities will be linked to an up-level community. Abstract SLNs can be created by making abstraction on semantic nodes and semantic links according to the semantic space. This strategy can adapt to the change of domains. Semantic links such as *equal*, *specialization*, and *generalization* can be established between nodes in different individual SLNs.

Semantic distance can be used to discover and measure communities. Semantic distance between nodes within community should be shorter than those between communities. The semantic link within the same semantic community takes higher priority to route queries than the semantic link through communities.

Abstraction and community are two dimensions to recognize a network as depicted in Fig. 2.16.2. Each concept can consist of a hierarchy of concepts representing abstraction at different levels. Each community can be a hierarchy of communities representing communities of different scales.

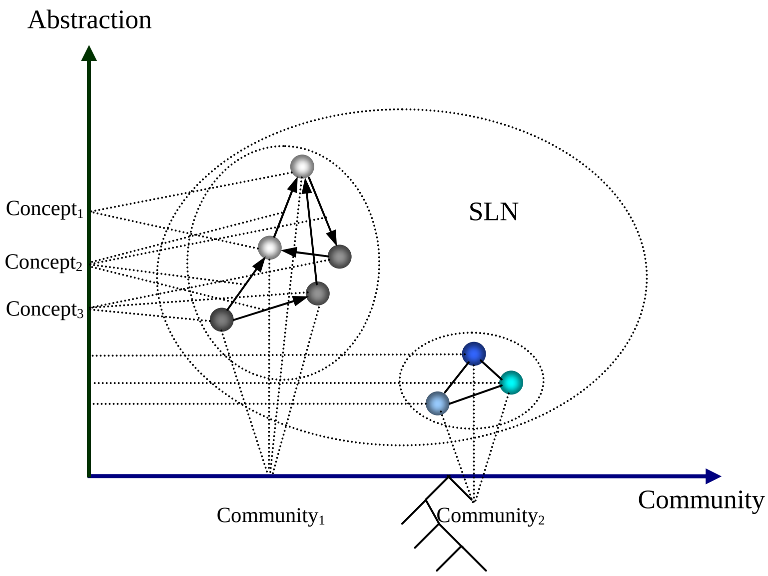


Fig. 2.16.2. Abstraction and community as dimensions of SLN.

2.17 Probabilistic Semantic Link Network

Semantic links sometimes are uncertain due to the following causes:

(1) Multiple semantic links may exist between two resources (things or nodes) and between clusters of resources. These links may play different roles in the network.
(2) Semantic links may be assigned by the users who are not sure.
(3) Semantic links may be predicted by inference rules, derived by relational, statistical or analogical reasoning.

So, SLN needs to reflect the uncertainty.

A probabilistic Semantic Link Network takes the following form:

$<\{R, C, T\}, L, Rules>$, where

(1) Resource set $R = \{r_1, \ldots, r_x\}$, a resource cluster set $C = \{c_1, \ldots, c_n\}$, and an indicator set $T=\{t_1, t_2, \ldots, t_q\}$. An indicator t_i is a word or a set of words under a light-weight grammar for indicating the semantics of a resource or cluster.
(2) Semantic link set $L = \{s_1[l_1, u_1], s_2[l_2, u_2], \ldots, s_y[l_m, u_m]\}$, where s_i ($1\leq i \leq m$) is a semantic link indicator with lower bounded probability $l_i \in(0, 1]$ and upper bounded probability $u_i \in(0, 1]$. This range reflects the probability of the semantic link derived from multiple semantic paths.
(3) Rule set *Rules* = {*LR*, *IR*, *CR*, *AR*}, where *LR* is a set of *relational reasoning rules*, and each takes the following form $\alpha[l_\alpha, u_\alpha]\times\beta[l_\beta, u_\beta]\xrightarrow{cd}\gamma[l_\gamma, u_\gamma]$, where *cd* is the *certainty degree* of the rule, $l_\gamma=f(l_\alpha, l_\beta)$ and $u_\gamma=f(u_\alpha, u_\beta)$ are determined by function *f*. *IR* is an *inference rule set* consisting of statistical inference rules and assertion rules on semantic links. *CR* is a set of *classification rules* for classifying resources according to the probable relations between the indicators, resources and clusters. *AR* is a set of *attribute-to-resource rules* for deriving relationships between resources from relationships between attributes. Each rule takes the following form: $Att(X)\text{—}\alpha\rightarrow Att(Y)\Rightarrow X\text{—}\beta\rightarrow Y$. Attribute-to-resource rules are mainly given by humans or interactions between humans and machines through statistics.

Classification rules can be obtained from the cluster networks and the indicator–resource networks.

Inference rules predict semantic links between resources according to the semantic links between existing clusters. This type of rules is usually acquired by statistical method from existing SLN.

Reasoning rules reflect the relations between semantic links. A general reasoning rule can be specialized in different domains by specializing semantic links, for example, the *equal* link can be specialized as the *sameTopic* link, and the *partOf* link can be specialized as the *subSite* link. SLN supports the following relational reasoning: $R_1\text{—}\alpha[l_\alpha,u_\alpha]\rightarrow R_2$, $R_2\text{—}\beta[l_\beta,u_\beta]\rightarrow R_3 \Rightarrow_{cd} R_1\text{—}\gamma[l_\gamma,u_\gamma]\rightarrow R_3$, where R_1, R_2 and R_3 are resources, α, β and γ indicate the semantics of the semantic links, and $[l_\gamma, u_\gamma]=f([l_\alpha,u_\alpha], [l_\beta, u_\beta], cd)$, e.g., $f([l_\alpha,u_\alpha], [l_\beta, u_\beta], p)= [cd\cdot l_\alpha\cdot l_\beta, cd\cdot u_\alpha\cdot u_\beta]$.

The *attribute-to-resource rules* can be given by users or learned from a set of samples for some purposes. A resource has attributes of multiple facets.

P-SLN concerns the following types of semantic link networks:

(1) *Cluster–resource network* consists of resource clusters, resource entities, the *instanceOf* link between resource and cluster, the *equal*, *similar*, and *subCluster/partOf* links between clusters and between resources.
(2) *Citation network* mainly consists of *reference* links. The generalization of *cocite*, *cocited*, *sequential* and *similar* links.
(3) *Attribute network* consists of such links as *sequential* and *equal* between attributes of resources. Different types of attribute may be measured by different functions.
(4) *Indicator–resource network* consists of the *indicate* links between indicators and resources. Its probability is the weight of the indicator. The *co-occur* link exists between a pair of indicators if they are used to indicate the same resource.

Table 2.16 shows the difference between the SLN and the P-SLN.

The probabilistic values of semantic links take part in reasoning. When new resources are added to the P-SLN, semantic links between the new resource and the existing nodes are inferred by the inference rules and the reasoning rules.

As the consequence, the statistical inference rules of semantic links are updated, the indicators of resource clusters are updated, and the resource classification rules are also updated.

Table 2.16. Comparison between SLN and P-SLN.

Components	SLN	P-SLN	Explanation
semantic links	$X—\alpha\rightarrow Y$	$X—\alpha[l_\alpha, u_\alpha]\rightarrow Y$	l_α and u_α are the lower bounded and upper bounded probabilities of semantic indicator α.
reasoning rules	$\alpha\times\beta\rightarrow\gamma$	$\alpha[l_\alpha, u_\alpha] \times\beta[l_\beta, u_\beta] \xrightarrow{cd} \gamma[l_\alpha l_\beta, u_\alpha u_\beta]$	l_α and l_β are the lower bounded probabilities of semantic indicators α and β; u_α and u_β are the upper bounded probabilities of α and β; *cd* is the *certainty degree* of the rule.
classification rules		$p(c \mid t_i)$	The probability of an indicator t_i in a resource belongs to cluster c.
inference rules		$src—r[l_r, u_r]\rightarrow tgt$	l_r *and* u_r are the lower bounded and upper bounded probabilities of semantic link r between clusters *srt* and *tgt*.
Attribute-to-resource rules	$Att(X)—\alpha\rightarrow Att(Y) \Rightarrow X—\beta\rightarrow Y$	$Att(X)—\alpha\rightarrow Att(Y) \Rightarrow X—\beta\rightarrow Y$	$Att(X)$ is an attribute of resource X

2.18 Discovering Semantic Link Network

2.18.1 *The General Process*

Human civilization helps individuals establish worldview through lifetime learning. Humans can learn rules to discover various relationships. Machines do not have any worldview, so a semantic worldview is needed to help automatically discover semantic links (H.Zhuge, Interactive Semantics, *Artificial Intelligence*, 174(2010)190-204).

Fig. 2.18.1.1 shows the process of automatically discovering semantic links between resources. A general discovery mechanism should separate the rules from the mechanism.

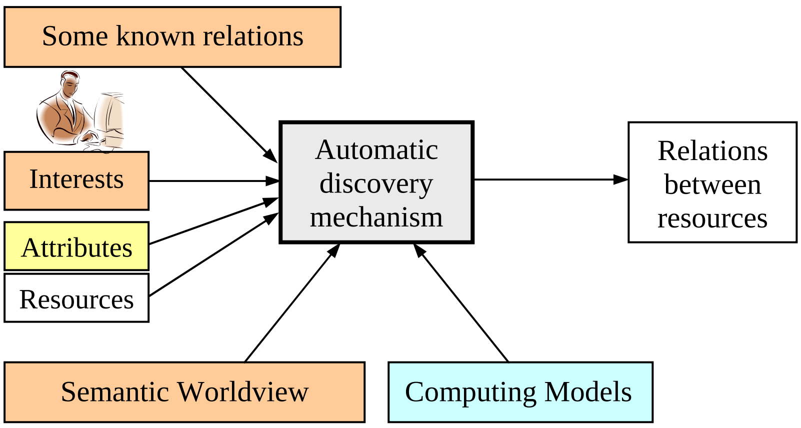


Fig. 2.18.1.1. The process of automatically discovering relations.

If users only want to discover the interested links, the following information is helpful:

(1) The superclass of the interested links, which enables the discovery mechanism to find all of the subclasses by searching the multi-disciplinary classification top-down.
(2) A class *C* and its ancestor *C*', which enables the discovery mechanism to search classes top-down from *C*' and bottom-up from *C*.
(3) A pair of interested relations (e.g., *coauthor* and *cite* are commonly interested links in e-science applications), which enables the discovery mechanism to find the closely relevant relations by searching the multi-disciplinary classification bottom-

up until reaching a common concept, then top-down from the relation pairs to the leaves.

(4) A semantic link instance (including the semantic link, the semantic nodes and their attributes), which enables the discovery mechanism to find the relations that are influenced by the given semantic link.

Discovering the interested semantic links among a set of resources can be regarded as the extension or enhancement of the communities in the individual semantic link networks waved during lifetime. As the consequence, the individual semantic images are extended and enhanced.

Some semantic links can also be discovered by content analysis. The approach to discovering semantic links among documents was suggested (H. Zhuge and J. Zhang, Automatically constructing semantic link network on documents. *Concurrency and Computation: Practice and Experience*, 23(9)(2011)956-971). A general approach needs to consider *the influence of human behaviors on the contents through various semantic links*.

2.18.2 Discover Semantic Links by Content Analysis

Semantic links can be discovered by finding the indicators of the contents of resources, resource clusters, indicator co-occurrence network, and the semantic links among resources, clusters and indicators.

Step 2. *Similarity calculation*. Resources can be clustered according to similarity, e.g., measured by $|T_1 \cap T_2|/|T_1 \cup T_2|$. In this step, co-occur links can be built between indicators and resources, and between indicators.

Step 3. *Hierarchical clustering*. A hierarchical SLN can be constructed according to the similarity between resources and the similarity between clusters. *InstanceOf* and *subCluster* are two major semantic links in the hierarchical structure of resource networks. The *subCluster* links may exist between the clusters of different levels. Clusters at the same level are similar to each other. Resources are clustered iteratively until all of the resources are in the same cluster or the times of iteration exceeds the predefined maximum value. Each resource cluster has a representative indicator set. Similar to the global indicator co-occurrence network, each resource cluster has a local indicator co-occurrence network recording the co-occurrence of indicators of the resources in the same cluster.

Step 4. *Estimating semantic links according to representative indicators*. Semantic links between resources or clusters x_1 and x_2 such as *similar*, *partOf* and *equal* can be estimated according to the following rules, where T_1 and T_2 are indicator sets.

(1) If $|T_1 \cap T_2|=0$, then x_1—*irrelevant*—x_2.
(2) If $0 <|T_1 \cap T_2|< min(|T_1|,|T_2|)$, then x_1—*similar*—x_2.
(3) If $|T_1 \cap T_2|= min(|T_1|,|T_2|) < max(|T_1|,|T_2|)$, then x_1—*partOf*→x_2.
(4) If $T_1 = T_2$, then x_1— *equal*—x_2.
(5) If $T_1 \subset T_2$, then x_1 is the *subCluster* of x_2, denoted as x_1—*subCluster*→ x_2.
(6) Besides, a resource r can be regarded as an instance of a cluster c (denoted as r—*instanceOf*→c) if r belongs to c.

Step 4. *Reasoning*. More semantic links can be derived from applying the reasoning rules to the semantic link network. If some semantic links exist between two nodes in the SLN, there would be one or more semantic link paths between them. The probability of a derived semantic link is the function of all probabilities of the semantic links in the path, which ensures that the probability of the derived semantic links reduces with the increase of the length of the semantic link path. The relations and the probability of the semantic links evolve with the semantic link reasoning.

Table 2.18 shows some relational reasoning rules on the indicator sets, which enable the semantic link network to carry out relational reasoning.

If there are semantic links between two resources, the two resources should share some indicators that imply the semantic links. The larger are the number and weights of the connecting indicators, the higher the probability of the semantic links between them.

Table 2.18. Some relational reasoning rules.

Relational reasoning rules	Characteristics
subCluster×*subCluster*⇒ *subCluster*	$T(c_1)\subset T(c_2)$, $T(c_2)\subset T(c_3)\Rightarrow T(c_1)\subset T(c_3)$
partOf× *irrelevant*⇒ *irrelevant*	$T(c_1)\subset T(c_2)$, $T(c_2)\subset T(c_3)\Rightarrow T(c_1)\subset T(c_3)$
partOf×*partOf*⇒ *partOf*	$T(r_1)\subset T(r_2)$, $T(r_2)\subset T(r_3)\Rightarrow T(r_1)\subset T(d_3)$
instanceOf×*subCluster*⇒ *instanceOf*	$T(r_1)\subset T(r_2)$, $T(r_2)\subset T(r_3)\Rightarrow T(r_1)\subset T(d_3)$
instanceOf×*subCluster*⇒ *instanceOf*	$r_1\in c_1$, $c_1\subset c_2\Rightarrow r_1\in c_2$
partOf×*instanceOf*⇒ *instanceOf*	$T(r_1)\subset T(r_2)$, $r_2\in c\Rightarrow r_1\in c$

The initial inference rules depend on the initial resource set. After initialization, the semantic links between resources can be discovered.

2.18.3 Enrich Semantic Links through Evolution

The P-SLN evolves with the changes of nodes and semantic links. New resources may change existing clusters. The indicators of new resources activate the changes of the cluster’s indicator sets.

When a new resource comes, its category can be determined according to the relations between its indicators and those of existing clusters.

A resource may belong to several clusters, so the indicator-cluster association rules can be used to infer the resource classifications. Then, new resources can be inserted into the resource cluster networks. New resources will change the cluster networks and the indicator networks, and the resource-cluster-indicator networks will influence the resource classification rules.

The evolution of cluster networks carries out with the following operations:

(1) *Addition of new resources*. Semantic links between a new resource and its clusters and those between the new resource and other resources can be established according to the indicators of the new resource, the existing resources, and the clusters. New resources may cause the change of the indicator set. The cluster's indicators evolve with the changes of the resource's indicators.
(2) *New resource clusters occurrence*. Semantic association degrees between the new cluster and the old clusters need to be calculated. If new clusters are clustered into the clusters at the higher level, the depth of the cluster networks may increase due to re-clustering.

An inference rule is influenced by the following factors:

(1) *Change of the source or target clusters of semantic links*. Addition of new resources will lead to the change of the clusters or the occurrence of new clusters.
(2) *Occurrence of new semantic link types*. When new resources are added, semantic link types may be increased. The change of the number of semantic links leads to the change of inference rules.
(3) *Change of classification rules*. The association rules between indicators, resources and clusters will evolve with the changes of the cluster's representative indicators. The classification rules change with the certainty degree of the semantic links between resources and clusters.

Since inference rules evolve with the changes of the probabilistic semantic link network, the semantic link types and the probability values are related to the insertion order. Semantic links and their certainty degrees may be different if resources are inserted into the semantic link network in different orders. When duplicate resources are added to the semantic link network, inconsistency may occur. But, this reflects the uncertainty of the network. Different inference results are caused by different initial resource sets. Newly added resources will influence the classification and the semantic link inference on the new resources. With the evolution of the network, the diversity of semantic links increases, and the intervals of the certainty degrees evolve.

The following are *major steps of discovering semantic links*:

Step 1. *Creating classification rules*

Resources can be clustered according to the representative indicators and their weights. The inference rules between indicators and classes can be obtained by the following statistic method.

The association between indicator t and cluster c_1 can be calculated by the following formula:

$$P(c_1 \mid t) = \sum_{i=1}^{n} P(c_1 \mid r_i) P(r_i \mid t),$$

where $P(c_1|r_i)$ is the probability that resource r_i belongs to class c_1, and $P(r_i|t)$ is the probability that t is used as one of the indicators to represent resource r_i. $P(c_1|r_i)$ is calculated by the cluster algorithms while $P(r_i|t)$ are calculated by the following Bayes formula:

$$P(r_i \mid t) = \frac{P(t \mid r_i)\, p(r_i)}{\sum_{t \in r} P(t \mid r)\, p(r)},$$

where $P(t|r_i)$ means the probability that t is used as one of the indicators to represent resource r_i.

Step 2. *Building semantic link inference rules*

One resource may belong to several clusters. The probability of α can be calculated as follows, where l and u are the lower bounded and the upper bounded certainty degrees of α, and *src* and *tgt* are the source and target clusters, $cd(l(\alpha), src, tgt)$ and $cd(u(\alpha), src, tgt)$ are respectively the lower bounded and the upper bounded certainty degrees of α between *src* and *tgt*, and A is a semantic link between *src* and *tgt*.

$$\text{Min-cd}(\alpha, src, tgt) = \frac{\sum cd(l(\alpha), src, tgt)}{\sum cd(u(A), src, tgt)}$$

$$\text{Max-cd}(\alpha, src, tgt) = \min\left(\frac{\sum cd(u(\alpha), src, tgt)}{\sum cd(l(A), src, tgt)}, 1\right)$$

If resources r_1 and r_2 are given, their classifications can be found by using the classification rules, and the certainty degree of semantic link β between r_1 and r_2 can be inferred if r_1 belongs to class *src* and r_2 belongs to class *tgt*. The following is the general semantic link inference rule.

$$src \text{—} \beta[Min\text{-}cd\ (\alpha, src, tgt), Max\text{-}cd\ (\alpha, src, tgt)] \rightarrow tgt.$$

Step 3. *Discovering semantic links*

Semantic links between the newly added resources and the existing resources can be inferred. Given resources r_1 and r_2, semantic links between them can be inferred as follows.

(1) Obtain the indicator sets T_1 and T_2 of r_1 and r_2.
(2) Find resource clusters for resources r_1 and r_2 according to the similarity between the indicator sets or the cluster's indicator sets, and the indicator–cluster association rules among resources, indicators and clusters.
(3) If r_1 and r_2 are in the same cluster, find the semantic links according to the indicator sets. Semantic links such as *irrelevant*, *similar*, *partOf* or *equal* can be discovered.

New semantic links can be derived from the evolving probabilistic semantic link network.

The resource classification rules, inference rules and cluster association networks automatically evolve with the evolution of the network.

After specialization, the proposed approach can be used to automatically construct a semantic overlay on any resource set to support advanced applications such as recommendation and relational query.

2.19 SLN 3.0: Cyber-Physical-Socio-Mental Network

Things are related not only in the cyber space but also in and through the physical space, social space and mental space. Some relations are objective, some are subjective, some are explicit, and some are implicit. Some networks are natural like the food web, while some are artificial like the World Wide Web. Consciously and subconsciously networking, maintaining and making use of versatile relations, humans act intelligently to evolve the social space and influence the other spaces.

The cyber-physical society operates with various self-organized semantic link networks and the method that can record, discover, understand, predict, indicate, maintain, and evolve various explicit and implicit semantic links.

2.19.1 *Origin of semantics*

Lived in the space where objects are interrelated explicitly or implicitly, humans recognize and learn the relations through cognitive processes, and use the relations to link external objects to internal concepts and link concepts to concepts to construct and maintain a semantic link network through spaces to support intelligent activities.

Human senses such as sight, hearing, touch, smell, taste, time, pain and temperature are the origin of human semantics. Categorizing the *values of senses into classes, naming these classes, mapping them into the mental space as concepts, and abstracting classes into class hierarchies.*

Human recognition on classification and set leads to the establishment of some basic relations such as *subclassOf*, *instanceOf* and *memberOf*. *These basic semantic links can be discovered by the equivalence relations between senses, between attributes, and between classes.*

The recognition of symbiosis, blood, and kinship relations also start from senses and classifying senses.

Interaction plays an important role in establishing consensus on classifications, naming, and mapping objects and relations in the physical space and social space into the mental space.

Natural languages are generated from indicating physical objects and concepts and further used to indicate complex meaning for effective interactions. They have evolved into a rich symbol space with diverse grammars. The symbol space evolves with the development of the cyber space, social space and mental space. Language has become an important function of these spaces for interaction. The study of language evolution concerns formal method, learning method, and evolutionary dynamics (M.A. Nowak, et al., Computational and Evolutionary Aspects of Language, Nature, 417(2002)611-617).

SLN can be regarded as a language on relations with various rules.

Taxonomies are established and evolved with the development of various domains. Human understanding is based on the internal structure and through multiple channels. The symbols used to indicate the basic relations should be explained by the scenes in the real world and the abstract characteristics on classification, set and number.

Localization of social behaviors in human history determines the diversity of languages. People can use their own words to indicate objects by extending the classification hierarchy. The frequently used words will be gradually accepted as commonsense and linked to the existing classes (concepts).

Individuals have freedom to use symbols to indicate physical objects, concepts and relations. It takes time to establish consensus on using symbols within community, especially in a multi-cultural community. Different indicators may be used to indicate the same relation, and one indicator may be used to indicate multiple objects or relations, so mappings between indicators are needed to establish uniformity on the diversity. Therefore, a relation may be indicated by a set of symbols. Relations on relations enable humans to derive new relations based on existing relations.

A semantic link network in the social space is waved through various social interactions. These interactions are the driven force to evolve the society. The relations keep evolving with the development of human recognition of the physical space and social space. Mapping the social semantic link networks in minds while interacting with each other, humans evolve the mental semantic link networks.

Fig.2.19.1.1 shows the structure of the interactive semantic base that facilitates understanding during interaction between individuals through machines (H. Zhuge, Interactive Semantics, *Artificial Intelligence*, 174(2010)190-204).

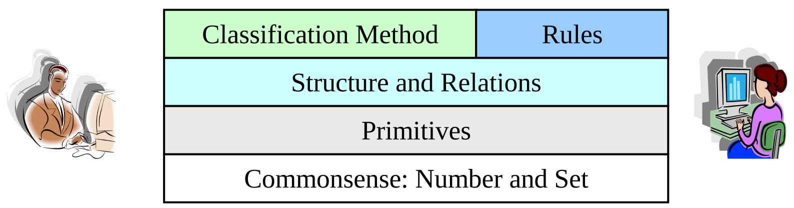


Fig. 2.19.1.1 The Interactive Semantic Base.

A semantic link network also bridges human semantics and machine semantics through the cyber space as shown in Fig. 2.19.1.2.

Various interactions enrich the diversity of semantic nodes and semantic links.

In the above discussion, both relation and semantic link are used. What is the difference between relation and semantic link?

Relation is general and static, while semantic links can be specific or general, explicit or implicit, and can be changed from time to time.

Various semantic link networks keep evolving with the development of society and culture.

Semantic link network's social and dynamic features make it different from traditional semantic net.

The semantic link network in the cyber-physical-social-mental complex space provides the semantic context for humans or machines to act meaningfully through sense.

A relevant phenomenon is that children's writing ability does not significantly rely on the amount of reading materials.

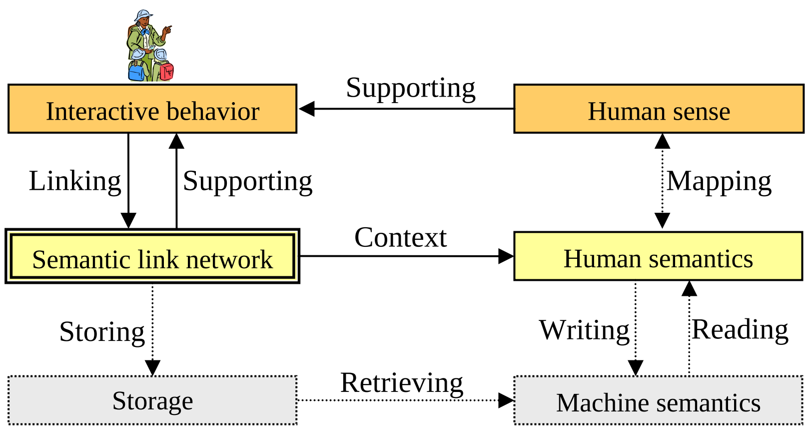


Fig. 2.19.1.2 Semantics of the semantic link network.

Establishing rich links between reading materials and between reading materials and the sense of writing plays an important role in the development of writing ability.

In macrocosm, semantic links go through the physical space, cyber space and mental space to enable these spaces to interact effectively and co-evolve harmoniously. Various closed-loop flows through the links are the basis of forming the ability.

Different people have different abilities of semantic networking and appropriately connecting to the existing networks of different types. Different people may involve in different levels of a semantic link network due to the difference of cognitive ability and social status.

To discover some unknown relations and rules in society and nature needs scientific research. Some intrinsic relations need long-term in-depth research.

2.19.2 *Characteristics*

Since SLN 3.0 will pass through not only the cyber space but also the physical space, socio space and mental space, semantic nodes represent physical objects, social individuals, and mental concepts, which obey socio rules. The following are its characteristics.

(1) *Autonomous*. Semantic nodes can actively connect to the appropriate nodes according to socio rules. Generally, there is no central control during semantic networking.
(2) *Self-evolution*. The structure and semantics of a semantic link network evolve with operating, networking and interacting behaviors.
(3) *Semantic community discovery*. Semantic communities emerge and evolve with the motion of the network. Community discovery is at a certain coordinate of the time dimension.
(4) *Reflecting uncertainty*. Uncertain semantic links reflect uncertain physical phenomena and relations.

(5) *Complex and temporal link.* Some semantic links reflect the complex relation that cannot be explained in a single concept. Some semantic links can pass through various flows such as material flow, information flow, and service flow. Some semantic links exist only within a period of time.
(6) *Relation prediction.* Potential and future relations can be derived or predicted according to relational and statistic reference rules and various reasoning mechanisms.

In addition to the basic maintenance operations, advanced operations on the semantic link network include the following:

(1) *Recommendation*. SLN 3.0 can recommend appropriate resources, links and communities to relevant individuals.
(2) *Complex reasoning*. SLN 3.0 supports relational reasoning, analogical reasoning and inductive reasoning as well as complex reasoning.
(3) *Influence* aware. Influence of operations will be aware when operating SLN 3.0.
(4) *Explanation*. SLN 3.0 can explain the semantics of a semantic link, path or a community according to the semantic space, complex reasoning and influence analysis.
(5) *Query answering*. SLN 3.0 can answer queries on various relations between semantic nodes, and can quickly locate specific community that includes some semantic nodes or links.

Research on SLN 3.0 mainly concerns the following issues:

(1) The method for discovering, predicting, establishing, and maintaining complex semantic link networks.
(2) The intrinsic characteristics and rules of semantic link network motion.
(3) The emerging semantics with the evolution of the network.
(4) Diverse models for reasoning (e.g., relational reasoning, analogical reasoning, inductive reasoning, and complex reasoning), prediction and influence.

(5) Various influences in the semantic link network, and the method of making use of influence.
(6) Various flows through semantic link networks and relevant rules.

2.19.3 *Intension and extension*

Some semantic models represent physical systems by specifying the *intensions* of objects: attributes and relations between attributes (e.g., function dependence in the relational data model) or relations between functions.

Some objects like ancient artifacts are hard to be directly indicated by symbols (because only some experts can know the real meaning, e.g., in some Chinese ancient cave paintings as shown in Fig.2.19.3.1), but they can be largely indicated by the known relations between artifacts. How to indicate the relations between paintings? What are the rules of linking one painting to the other?

Fig. 2.19.3.1 Semantically linking ancient paintings.

A certain range of semantic link network represents the *extension* of a semantic node as shown in Fig. 2.19.3.2.

The *minimum extension* of a semantic node (e.g., X in the figure) is determined by the directly connected semantic links (e.g., α, β, η and γ), and the directly linked semantic nodes (e.g., A, B, C and D).

The semantic community of a semantic node can be seen as the *maximum extension* of a semantic node in a large complex semantic link network. A definition of the semantic community is given in (H.Zhuge, Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering*, 21(6)(2009)785-799).

Analogical reasoning, inductive reasoning and relational reasoning are approaches to infer the intension and extension of a semantic node.

For example, the intensions of X and Y could be regarded as similar to each other if the extensions of X and Y are similar to each other (there exists a kind of structural mapping between the two extensions).

Semantic link network can be viewed from different granularities. Some semantic links are between objects like the citation relation between papers. Some semantic links are within objects, for example, the sequential links between sections, between paragraphs, and between words.

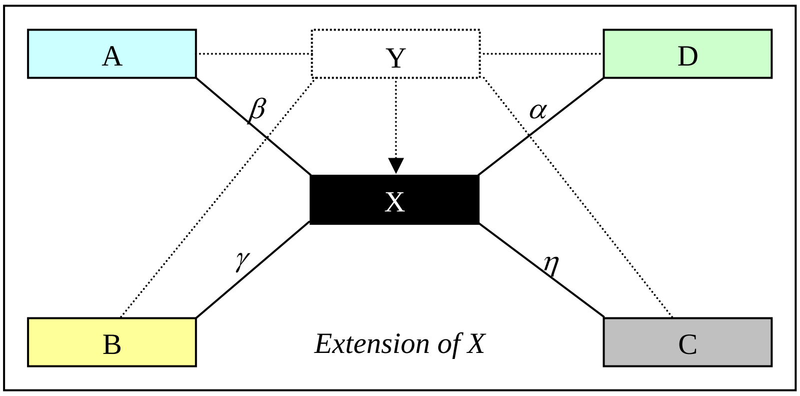


Fig. 2.19.3.2. The intensions of the neighbor semantic nodes and the directly connected semantic links determine the minimum extension of a concept.

2.19.4 *Semantic Link Network of events SLN-E*

Besides resources, the social space consists of versatile intangible and temporal events. An event is an interaction between resources *R* and individuals *I* at particular *time* and *venue*.

$$Event\ (time,\ venue,\ R,\ I).$$

The *time* is an interval consisting of the start time and the end time of the event. The *concurrent*, *overlap*, *inclusion*, and *sequential* links between events can be determined according to the time intervals of the events.

The *venue* can be a region or a space, according to which, the geographical links between events can be determined. The links may be about distance and direction. So, the *inclusion*, *overlap* and *disjoint* links can be referred according to the distance between regions.

The sequential events that a resource or an individual (human or agent) participated in within a time interval constitute the *extension of experience*. The *intension of experience* is the mental reflection of the physical space. The participants sharing the same or similar experience are likely to have some common or similar features. Semantic links exist between resources if they have the same experience. For example, people studied the same course or often discussed with the same professor are likely classmates.

Various resources, events, and individuals constitute a vivid social network. Classifying, storing and retrieving events enable individuals in the Cyber-Physical Society to share and retrieve not only resources but also events.

The *cause-effect* link exists between events. Reasoning can be carried out through the semantic link chain of the cause-effect links. Through the link or chain, the causes of an event can be found.

Events can be clustered into categories according to the types of participants and interactions. Categories can also be clustered into higher-level categories to form a class hierarchy. An event is an instance of a category. Two events *Event*=(*time*, *venue*, *R*, *I*) and

Event'=(*time*', *venue*', *R*', *I*') can be united into one as follows if there are semantic links between *time* and *time*' and between *venue* and *venue*': *Event*$\cup$*Event*'= (*time*—α→*time*', *venue*—β→*venue*', *R*$\cup$*R*', *I*$\cup$*I*').

As depicted in Fig.2.19.4.1, connecting the semantic link network of resources to the semantic link network of events enriches the semantic link network. The connection can be made since a resource can participate in one or multiple events. Events would enable some isolated resources to be linked. The semantic link network of events also supports relational reasoning.

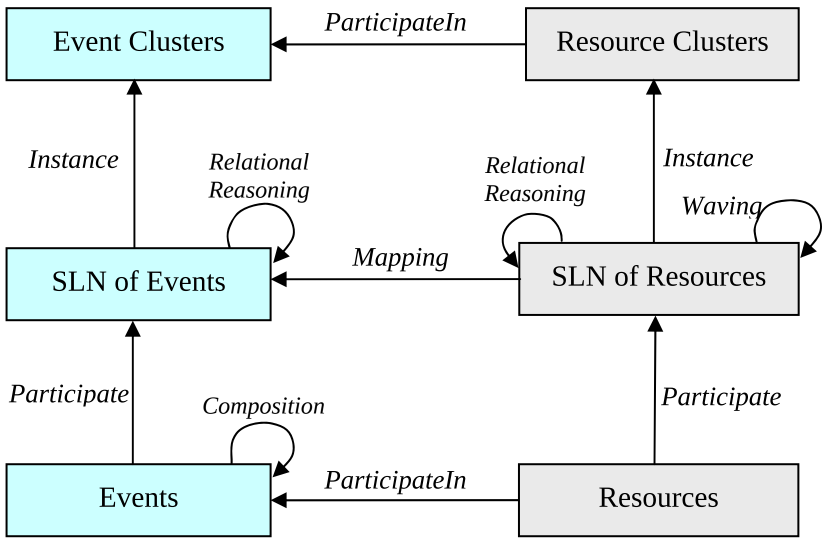


Fig. 2.19.4.1. Connecting the semantic link network of resources to the semantic link network of events enables applications to freely access the clusters of events or resources, the SLNs of events or resources, and the events and resources.

In human history, only very important events are recorded. Previous information systems only store and manage static resources like data and texts. Most events are lost during social development, so *it is important to explore the future with knowing a little story about the past*.

The semantic link network of events and resources provide applications with rich semantic links between events, between resources, and between events and resources, as well as various reasoning paradigms. Discovering rules in the semantic link networks of events and resources is important to support intelligent applications.

In the near future, a movie could be automatically generated by searching relevant resources (music, text, image, and video) and events and coordinating them for displaying sequentially. Reasonable coordination between resources and events is the key issue.

2.19.5 *Through minds via words*

A good author is able to use rich cues to connect simple or complex concepts to form diverse understanding routes.

The following is an example of selecting the most important keywords in the title and the first five paragraphs of this section from the author's point of view:

Section Title: {*semantic link network*}.
Paragraph 1: {*relation, network*}.
Paragraph 2: {*relation, semantic link*}.
Paragraph 3: {*relation, link, semantic link network*}.
Paragraph 4: {*semantics, class, classification*}.
Paragraph 5: {*classification, relation, class*}.

The above example indicates the following characteristics:

Characteristics

(1) *Words in the title take the priority of emerging keywords* (indicators).
(2) *Previously emerged keywords take the priority of emerging keywords*.
(3) *Words relevant to previously emerged keywords take the priority of emerging keywords*.
(4) *Sequential relation constructs the backbone structure*.

Characteristic (1) holds because authors carefully select the appropriate title words to represent the core idea of an article. Characteristics (2) and (3) hold because of the *relevance emerging principles* discussed in (H.Zhuge, Interactive Semantics, *Artificial Intelligence*, 174(2010)190-204).

Characteristic 4 indicates the following method:

The sequential relation is the main evidence to collect and link the text pieces distributed on the Web into a meaningful text.

Besides the explicit sequential relations, the following cues can be found:

The title links paragraph 1, 2, and 3 through the keywords *network* and *semantic link network*.

Paragraph 1 links paragraph 2, 3, and 5 through the keywords *relation* and *network*.

Paragraph 2 links paragraph 3 and 5 through the keywords *relation* and *semantic link*.

Paragraph 3 links paragraph 5 through the keyword *relation*.

Paragraph 4 links paragraph 5 through the keywords *classification* and *class*.

If the sequential relation is unknown, the five paragraphs can be retrieved according to the keywords in the title of this section, and then can be re-organized according to the implicit links. These links help readers understand the meaning of the paragraphs.

Fig.2.19.5.1 shows some links through the title and the five paragraphs. The sequential link between paragraphs indicated by $P_1 \xrightarrow{seq} P_2$ reflects the sequential browsing order.

The Vector Space Model and its improvement represent the general content of text (G. Salton, et. al., A Vector Space Model for Automatic Indexing, *Communications of the ACM*, 18(11)(1975)613-620), based on which, only the similarity between texts can be measured according to the angle between vectors. It is hard to reflect rich semantic links within text.

Readers need to scan the text to pick out the keywords and find the cues between paragraphs. To better understand the content, readers need to traverse the implicit semantic link network several times by different routes. If the implicit semantic link networks can be known by information service systems, an intelligent navigation system can be built to help readers to quickly understand the author's meaning. One way is to build a convenient authoring tool to help authors to describe the semantic link networks (H. Zhuge, et al., Semantic Link Network Builder and Intelligent Semantic Browser, *Concurrency and Computation: Practice and Experience*, 16(14) (2004)1453-1476). The other way is to automatically discover semantic link networks.

Let's look into how do writing and reading pass through semantic link networks.

A writer organizes the words in the symbol space to indicate the corresponding concepts in the mental space and control hand to write the words sequentially according to the following localization principle:

The distance between two words in one sentence is closer than that in two sentences, and the distance of two words in one paragraph is closer than that in two paragraphs. The concepts corresponding to the closer words take higher priority to emerge in reader's mind.

While forming the text, the author selects words from his/her memory to compose a text, and all of the selected words have certain probability of being changed during revision. This also fits human reading scope. The probabilities of some words are higher than others if some *related* words have higher probabilities. For example, if the word “semantic” occurs, the probability of selecting the word “link” as the following word is high since they often co-occur in the sequential order. Therefore, the two words are probably two semantic nodes of the same probabilistic semantic link in the text.

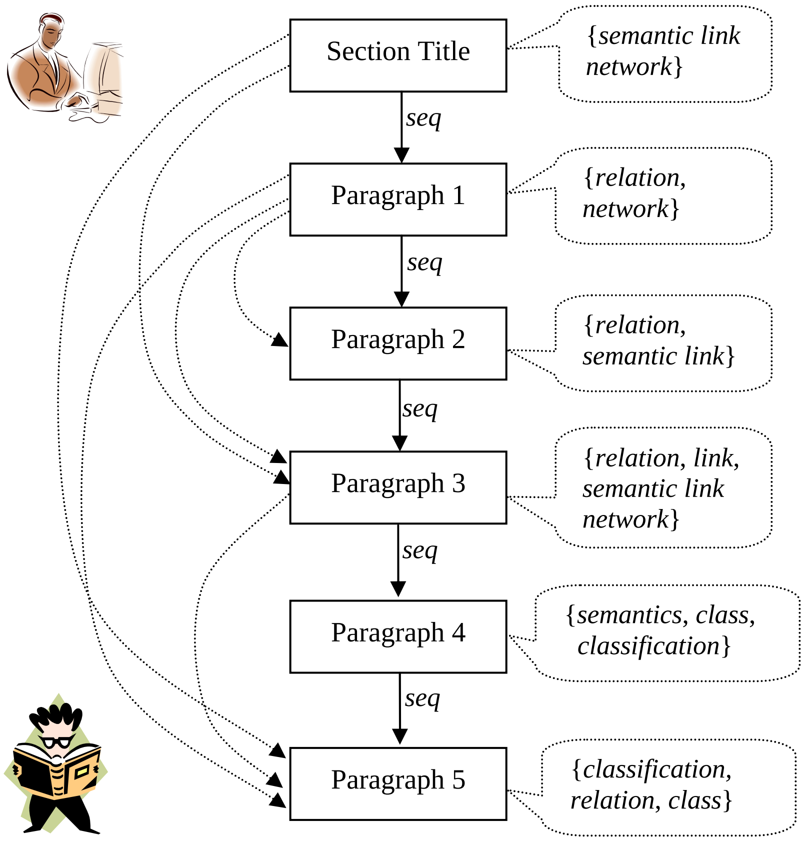


Fig.2.19.5.1 Semantic links through the first five paragraphs of section 2.19.

Reading is a process of constructing semantic link networks of concepts in reader's mind by discovering and browsing the semantic link networks of words waved by the author.

During browsing the text, multiple words can be probably selected within the range of reader's view. The new semantic link network can be the shrink, expansion or revision of the original network.

Semantic link network can be used to improve the quality of text summarization. Previous text summarization approaches need quality criteria (R. Barzilay, K. R. McKeown and M. Elhadad, Information fusion in the context of multi-document summarization, *ACL '99*).

Usually, a summarization can be made by extracting the important sentences from the original text. An advanced summarization may reorganize sentences or even reorganize words. Cues interpret the core idea of authors, so important cues should be maintained in the generated texts. People read a text sentence by sentence, so we have the following principle:

An understandable text should ensure that there are semantic links or semantic link paths between words in two close sentences.

We can build semantic link networks of different levels from the given text: word level (SLN-W), sentence level (SLN-S), paragraph level (SLN-P), and text level (SLN-T) as shown in Fig. 2.19.5.2. The word order is the basis of interpreting meaning.

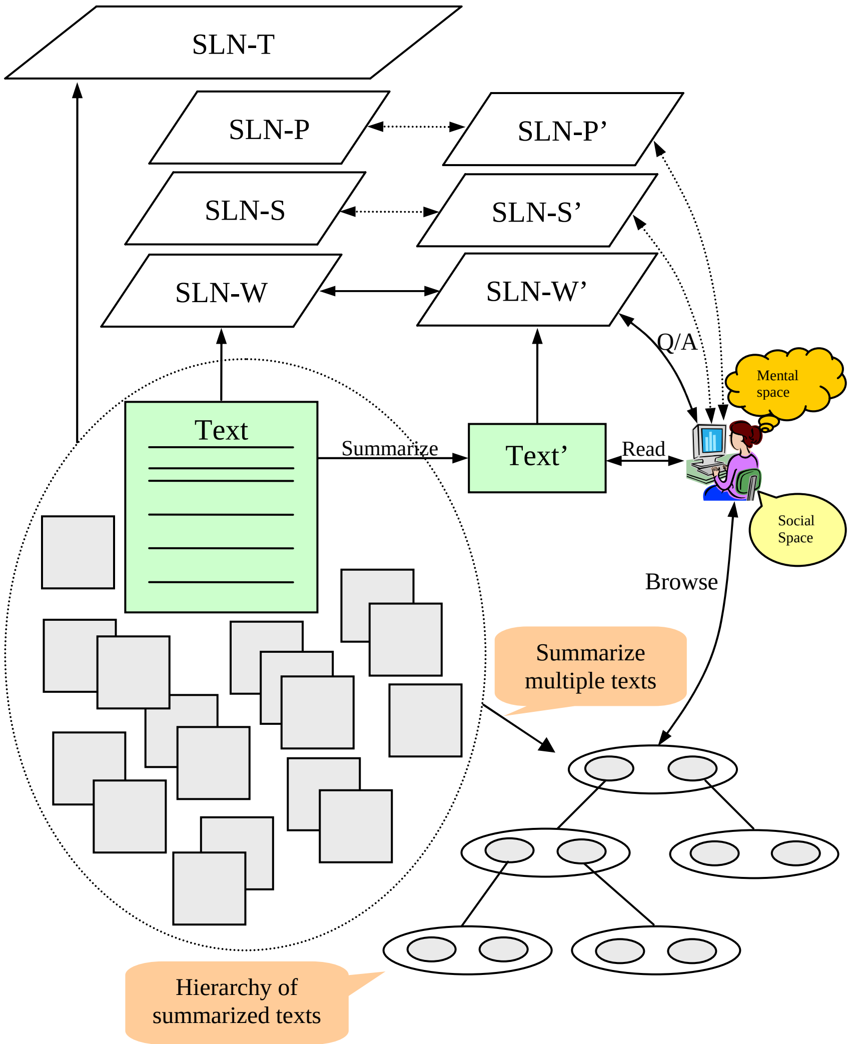


Fig. 2.19.5.2. SLN-based text summarization.

The following is the basic principles of SLN-based text summarization.

(1) *The important words and the semantic links between the words in SLN-W should be in the semantic link network of words SLN-W' in the summary*.
(2) *The important sentences and the semantic links between sentences in SLN-S should be in the semantic link network of sentences SLN-S'* in *the summary*.
(3) *The important paragraphs and the semantic links between paragraphs in SLN-P should be in the semantic link network of sentences SLN-P' in summary*.

The importance can be measured by various centralities.

One way to summarize multiple texts is based on the single text summarization approach: Summarizing each text first, and merging the summarized texts into one text, and then summarizing the merged text.

A semantic link network of texts (SLN-T) can provide a context for summarizing either multiple texts or single text. Since a semantic community in SLN-T includes closely related contents, the following is the SLN-based approach to summarizing multiple texts:

(1) Discover communities in SLN-T. As the result, a community hierarchy is formed.
(2) Summarize texts in the bottom-level (smallest) communities.
(3) Summarize each community according to the semantic links between the summarized texts.
(4) Replace the bottom communities with the summarized texts.
(5) Repeat from (2) until the top-level communities have been replaced.

The SLN-based summarization approach can output hierarchies of the semantic link networks of the summarized texts for readers to browse.

The generated SLNs and the community hierarchies in SLNs can also provide a kind of knowledge in texts for answering questions. The answers will be based on the semantic links, the structure of the community hierarchies, and the rules that can derive out implicit links.

So, the SLN-based text summarization approach provides the following three ways for users to efficiently link the author's mind and the reader's mind through texts:

(1) Read the summarized text to establish the general concepts and semantic links in mind.
(2) Browse the community hierarchies to know the contents top-down or bottom-up.
(3) Ask questions and get answers according to SLN reasoning and the hierarchies.

The above three ways can be either used separately or integrated into a single interface.

As shown in Fig. 2.19.7, a mental space and a social space accompany, evolve with, and influence the interaction process. The semantic images of words also evolve with the process and experience in the physical space and social space. Text summarization should be put into the interaction process so as to better know the real interest of the user.

2.19.6 *Through society, culture and thought*

Diverse types of semantic link networks need to be coordinated to provide semantic grounds for various socio behaviors as depicted in Fig.2.19.6.1.

The probabilistic semantic link network reflects the conditional probabilistic relation. The co-occurrence semantic link network reflects the basic co-occurrence relation between events. The two types of semantic links are objective and mainly between objects or events. Some networks are semantics-rich while some are semantics-poor.

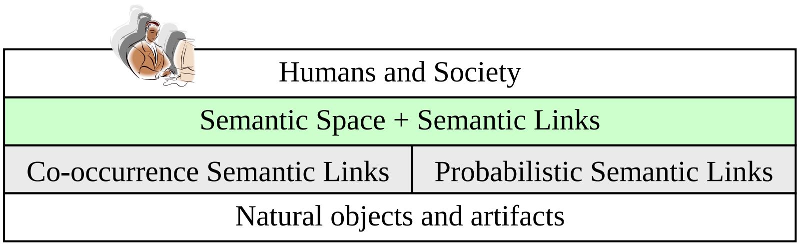


Fig. 2.19.6.1. Semantic links of different layers.

The social relation (denoted as *socialRel*) can be specialized for social networking as the following relations: *familyRel*, *workRel* and *communicationRel*. The *familyRel* can be specialized as the following links: *childOf*, *wifeOf*, *husbandOf*, *fatherOf*, *motherOf*, *brotherOf*, *sisterOf*, …, etc. The *childOf* link can be further specialized as the following links: *sonOf* and *daughterOf*. The *workRel* link is determined by the hierarchy of a work organization. In scientific research field, the *workRel* link can be specialized as the following links: *AuthorOf*, *CoAuthorOf*, *ColleagueOf*, *ReviewerOf*, *PIOf*, *memberOf*, …, etc.

Isolated node and link are meaningless. Semantic nodes and semantic links are regulated in a semantic link network, where the semantic space (classification hierarchy and rules) regulates their semantics. Reasoning on semantic link network is a kind of special operation for deriving semantic links. The derived semantic link network implies the original network in semantics. But, its semantics will change if new nodes or semantic links are added by humans or reasoning mechanisms.

Fig. 2.19.6.2 shows the scenario of waving semantic link network in society. Individuals establish their own SLNs by learning and understanding the relations between resources, and by establishing semantic links between individuals through interaction. Some semantic links are implicit, some are explicit, some are in mind, and some are record in text or other media.

The nodes and semantic links interact with each other to support intelligent behaviors. Interaction between individuals forms the global semantic link network, which may exhibit a certain pattern during evolution, e.g., semantic community (H.Zhuge, Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering*, 21(6)(2009)785-799). Relational reasoning usually carries out within a semantic community.

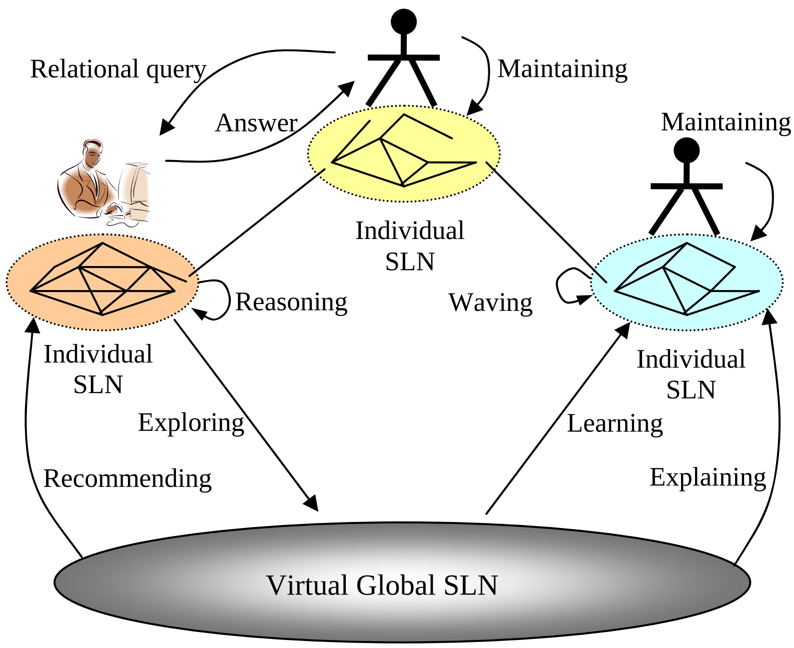


Fig.2.19.6.2. Waving semantic link networks in society.

Both semantic link and semantic node play an important role in evolving the semantic link network. A semantic node can be an abstract concept, a physical object, a digital object, an event, or an individual in society.

Nodes in different cultures play different roles in forming the general pattern of a semantic link network.

In some cultures, the family relation plays the central role. In some cultures, the work relation plays the central role. In some cultures, the friend relation plays the central role. In some cultures, two of or all of the above relations play the equal roles.

All individual SLNs constitute the virtual global SLN, which provides exploring and learning context for individuals. Relational reasoning activities happen more often within individual SLNs than between individuals. Analogical reasoning carries out in and through the cyber space and the mental spaces by comparing different semantic link networks.

Some common SLNs can be abstracted as an abstract SLN for effective sharing and operation. Schema is a kind of such an abstraction.

The following is a basic approach for an individual to establish semantic links with other individuals.

(1) Add the name and address of other individuals to his/her communication list, determine the category of the semantic link between them, and then make specializations. The general *association* relation can be specialized as such relations as *friend*, *co-occur*, *friendOf* and *studentOf*.
(2) Add the metadata of other individual resources to his/her peer list, and determine the semantic link according to the relation between metadata, e.g., the citation relation between papers. Knowing the metadata of neighbors, a peer will communicate with their peers more efficiently.

(3) Record query/answer pairs as well as the individuals or the types of individuals who asked and answered as the experience to support intelligent activities.

A cyber-physical-socio-mental semantic link network can be uncovered through analyzing various interactive behaviors in various spaces and the underlying networking mechanisms. Technical research and development can make use of current techniques such as social network, ontology, World Net, Web x.0, database, Semantic Web, domain modeling, uncertainty and semantics-based networking mechanisms like semantic peer-to-peer networks as depicted in Fig.2.19.6.3.

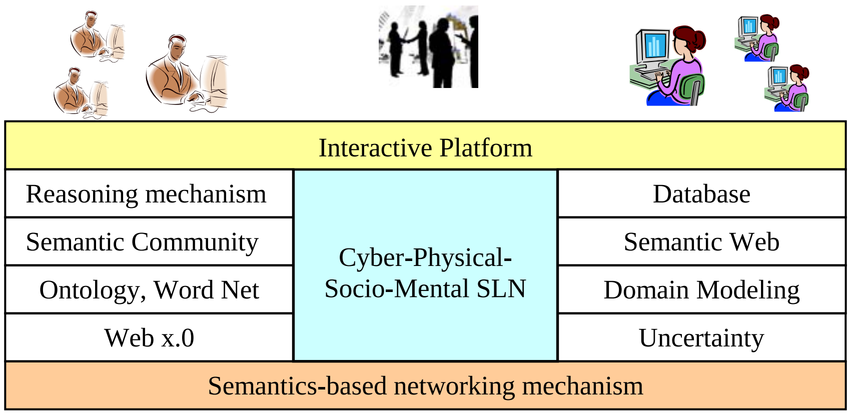


Fig.2.19.6.3. The cyber-physical-socio-mental SLN and relevant techniques.

The interaction between human individuals and their artifacts forms the following cultural networks:

(1) *Reference network of artifacts* — a network formed by observing the natural objects and artifacts, writing documents about the artifacts, or citing the documents about the artifacts. The number

of citations and readers determines the reputation of an artifact in the network.

(2) *Semantic link network of artifacts* — a network of relations between meta-information of artifacts. The attributes of an artifact can be inferred by relevant semantic links and the semantics of relevant artifacts.

(3) *Value network of artifacts* — The interactions between artifacts and the interactions between humans and artifacts constitute the value network of artifacts. The following factors influence the value of an artifact: the values of relevant artifacts as well as the authorities and the number of comments and citations.

Physically isolated artifacts interconnect with each other via relevant documents in the symbol space and the concepts in mind. So culture exists in and evolves with a semantic link network of artifacts and a network of interactions between humans. The two networks interact with each other by humans' making and exhibiting behaviors. Humans also write documents to explain artifacts.

Cultural semantic link networks can help people understand the form, content and creation of culture in multiple spaces through time, how people can benefit from them, and how they can inspire thinking in an advanced environment— carriers, forms, creation and evolution.

Market selection mechanism can be adopted to interpret the fitness in social and culture selection during their evolution. For example, the times of occurring culture features with the development of generations can be taken as selection criteria.

Although people wave and make use of semantic link networks consciously or subconsciously, we know little about their nature as a whole, how they operate and evolve, and how they help people to act intelligently.

The food web can be regarded as the degradation of the SLN as its links only reflect the trophic relationship. The rules on the food web reflect the influence between trophic levels.

The society can be abstracted as a SLN consisting of various social resources (including material resources and human resources), social relationships and social rules. Some social relations are explicit while some are implicit. Some social rules need to be explored. So, it is not easy to build a social SLN. It is a challenge issue to study the laws of the social SLN. Research on SLN can help understand the society.

SLN 3.0 concerns the fundamental issues about the world and human cognition: *How things in the social space (or other spaces) are related, maintained and evolving*? *How to make use of these relations to act intelligently*?

Previous research on brain network consists of three types of network: structural network, functional network, and effective network. *They are far from the semantics*.

Fuzzy Cognitive map represents causal relations between concepts (B. Kosko, Fuzzy Cognitive Maps, *International Journal of Man-Machine Studies*, 1986, pp. 65-75), but it is limited in ability to represent diverse relationships. SLN has potential to model the thought network by incorporating the semantic lens and solving some challenge issues discussed in (H.Zhuge, Discovery of Knowledge Flow in Science, *Communications of the ACM*, 49 (5) (2006) 101-107).

Since semantic relationships are not easy to be accurately indicated, statistical, inductive and analogical reasoning mechanisms are feasible candidate approaches. Complex network analysis could be useful implications (D. Liben-Nowell and J. Kleinberg, The Link-Prediction Problem for Social Networks, *J. Am. Soc. Inform. Sci. Technol*. 58 (2007) 1019; A. Clauset, C. Moore, and M. E. J. Newman, Hierarchical structure and the prediction of missing links in networks, *Nature*, 453(2008)98).

SLN 3.0 study also concerns modeling mental network or thought network. A dynamic SLN thought component has the following features:

(1) A semantic node can contain concepts, axioms, lemmas, rules, methods, and even theory. Semantic nodes are dynamically ranked to reflect the up-to-date importance of nodes.

(2) Semantic links mainly indicate the following relations: *subclassOf* (*subtypeOf*), *cause-effect*, *similar*, *reference*, *sequential*, and *partOf*.

(3) A component can have built-in rules for reasoning and evolving.

(4) The SLN evolves with adding and linking new concept nodes. Ranks of nodes change with the evolution. The evolution of SLN simulates the evolution of thought.

(5) Two components can be merged into one by the following two links: reference from one component to the other, and reference from the third component to the two components. A reference link would generate more semantic links and semantic links between the two components would be enriched through reasoning.

Biology research shows that navigation of brain information is in triangle, which resembles the triangle semantic link network reasoning in the cyber space, social space, and mental space.

2.19.7 *Principles of emerging semantics*

Linking a semantic node or adding a semantic link to an SLN could generate new semantic links. A new semantic node could immediately know relevant semantic nodes through reasoning or flows through links. The following gives a new measure for nodes and links.

The richness of a semantic node is in positive proportion to the number and diversity of the semantic links it has and the richness of its neighbor nodes.

A richer semantic node could provide richer contents and more semantic relations for others.

The richness of a semantic link is in positive proportion to the following factors:

(1) *The number and richness of the semantic links it can reason with, the more the richer*.
(2) *The times of the relation appeared in SLN, the more the richer*.
(3) *The richness of its two ending nodes, the richer the richer*.

The following is the massive emerge principle:

The more diverse the richer.

A richer link contributes more richness to the connecting nodes. A richer node supports the richness of its connected links. So, this principle implicates the following strategy for a new semantic node to be rich.

Linking to enrich semantic links.

That is, new semantic link should be able to influence or reason with the potential neighbor semantic links.

Adding a new semantic link to SLN reflects the purpose of the new node. If the semantic link is unknown, the strategy for a new node to become rich can be simplified as follows:

Link to the richer node, as the richer node owns more diverse semantic links, which offer higher probability to influence or reason with the new semantic link.

The above principle can be explained by the following example:

A person with only one relation, e.g., family relation, usually has low and unstable social status. A person with multiple relations (e.g., not only family relation but also friend relation) usually has higher and relatively stable social status. So, if a person wants to raise social status, he/she should link to the person who has diverse relations rather than link to the person who is isolated or has only single relation, because this leads to higher probability to make new relations that can raise social status of involved persons. Generally, a node should have at least two kinds of relations to maintain social status in a social network.

In depth analysis of social relation needs to consider the negative and positive influence through a semantic link.

Fig. 2.19.7.1 shows three kinds of nodes. The massive emerging principle suggests that nodes C and D should take the priority to emerge as the candidates for node A to link. Selection decision also relies on the type of the relation between node A and the candidates.

Adding semantic links to an SLN tends to make shorter semantic paths.

Since implicit semantic links may be derived out from time to time with the evolution of the network, the richness of semantic nodes will be changed from time to time. This helps the new nodes to share the richness of network. Selecting appropriate nodes to connect, the network provides the chance for a new semantic node to become rich. Meanwhile, the new node helps the old nodes become richer, which in turn helps itself become richer.

The massive emerging principle maps a flat network into the metric space to help discover the emerging semantic nodes and links.

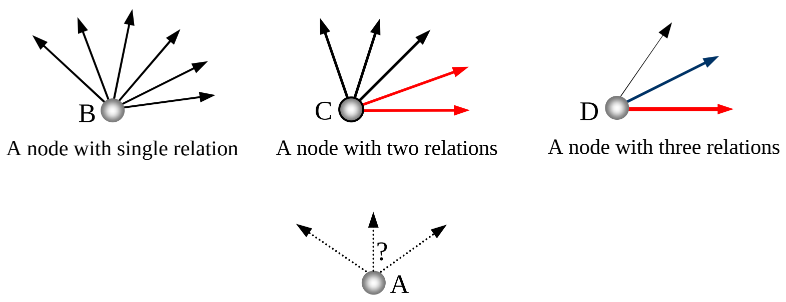


Fig. 2.19.7.1. Linking to the node with diverse relations.

Multiple semantic paths may exist between semantic nodes in SLN. It is harder to explain the semantics of the semantic link path containing more different types of semantic links. The simplest principle is indicated by the following commonsense:

The less information a semantic link path contains, the easier people understand and remember.

The following is the simplest principle.

Among multiple semantic paths, the shortest path with the least types of semantic links takes the priority of emerging.

The above principle can be explained by information entropy. The richness changes with the evolution of the network, while the entropy of a semantic path is relatively stable unless the path itself changes.

The simplest emerging principle focuses on a particular semantic path while the richness emphasizes on the status of a semantic node or a semantic link in the whole network.

The simplest semantic path is the semantic link path with the least types of semantic links between two semantic nodes in the corresponding semantic closure SLN^+.

This massive principle reflects the law of movement of SLN, and the simplest principle reflects the law of recognizing and understanding an existence.

The following are principles from different facets.

The Distinguish Principle. *A semantic node or community has the priority of emerging if it is distinguished from others*.

A semantic community needs to maintain its distinguished characteristics. A humanized cyber society should enable any individual to autonomously select appropriate friends, and enable any community to maintain appropriate structure. So, individual and community should be able to predict situation and actively select new semantic links.

The Relevance Principle. *A semantic node or community has the priority of emerge if it is linked to an emerging semantic node through an existing or a potential semantic link.*

The above principles of emerging semantics can apply to the emerging of thought in modeling thought network. If the emerging path of semantic nodes is regarded as a thought, the relevance principle can be extended to the following:

An emerging semantic link path has the priority of emerging if it is linked to an emerging path through an existing or a potential semantic link.

2.19.8 *Discovering semantic communities — semantic localization*

Previous research on community discovery focuses on various approaches to partition a static graph through operating edges or nodes. The formation process of a community is seldom considered.

A significant semantic community is formed according to some rules. *The key to semantic community discovery is actually to find the rules of forming community*.

A semantic community concerns two aspects: *structure* and *semantics* restricted by rules and reasoning.

A *reasoning-restrict semantic community* satisfies the following conditions:

(1) *It is a connected graph*.
(2) *It excludes such semantic links that do not participate in any reasoning*.
(3) *The intra-community semantic links participate in reasoning with each other much more times than with the inter-community semantic links*.

The second condition implies that a semantic link belongs to the other community if it cannot reason with its neighbor semantic links. This condition can be used to discover semantic communities, where every semantic link participates in reasoning at least once. The third condition ensures that semantic links should be tight within community and should be loose between communities. This notion allows semantic communities to share a semantic node. This also means that given different sets of semantic links over the same set of nodes may represent different semantic communities.

The total number of semantic links participated in reasoning with a semantic link reflects its importance in the network or the extent of other semantic links relying on it. The following definition reflects the importance or reliance between semantic links, which is useful in discovering semantic communities in SLN.

The semantic betweenness of a semantic link in a given SLN is the number of times it can participate in reasoning.

Some reasoning rules are closely related while others are loosely related. A set of closely related rules influences the formation of a semantic community.

A rule-restrict semantic community carries out reasoning mainly within community.

The basic assumption is that the concepts in the same classification tree should be closely related to each other. Therefore, they should be in the same semantic community. An SLN is classification-restrict if all of its nodes and links appear in the same classification tree.

A classification-restrict semantic community satisfies the following:

(1) *Semantic nodes and relations belong to a common class in the classification trees*.
(2) *The semantic distance between any pair of intra-community nodes ≤ the semantic distance between any pair of intercommunity nodes*.

A semantic node and a relation can belong to multiple classes. The following is the algorithm to discover semantic communities by removing the semantic links with the lowest semantic betweeness with reference to its closure.

(1) Construct the SLN^+ of the input SLN, record the semantic betweeness of all semantic links and list them in descending order. Record all of the semantic links that have reasoned with other semantic links.
(2) Remove the semantic links with zero semantic betweeness.

(3) Remove the semantic link(s) with the smallest semantic betweeness if this operation does not generate isolated nodes.
(4) Check the reasoning rule set and find all of the semantic links that have reasoned with the semantic link removed by step 3. Decrease the semantic betweenness of the semantic links that have reasoned with the removed semantic link(s) by 1.
(5) Repeat from step 2 until no semantic links is qualified to be removed or isolated node is found.

The above algorithm calculates SLN^+ once. It can also avoid recalculating the semantic betweenness for all semantic links after the removal of one semantic link by checking the influenced semantic links at step 4 according to the rules. A tree of communities could be formed during the discovery process.

Many networks have such features: some nodes play more important role than others in forming communities. Another idea of community discovery is to find some initial communities, adjusting and combining communities to discover more reasonable semantic communities according to the semantic links in the SLN^+.

The following algorithm inputs a community intensity η to help decide whether two communities should be combined to one. If more than η percent of nodes in one community are linked to the nodes in another community or vice versa, the two communities are likely to be combined to one in the evolution process.

(1) Calculate the degrees of all nodes (the total number of in-links and out-links) and rank them in descending order to form a degree queue (arbitrarily arrange the order of the nodes with the same degree).
(2) Construct semantic closure SLN^+.
(3) The node with the highest degree and its neighbors constitute an initial community C_0. Remove these nodes from the degree queue. Let t=0.
(4) t=t+1. Let the first node k in the queue be the central node of a new community C_t. Remove k from the queue.
(5) For every neighbor of node k, put the neighbor into one community

in $\{C_0, \ldots, C_t\}$, which has the largest number of nodes semantically linked to the neighbor in SLN$^+$.

(6) Check every community C_j (j=0, …, t). If more than η percent of the neighbors of node k belong to C_j or more than η percents of the nodes in C_j semantically link to node k in SLN$^+$, then merge C_t to C_j and t=t–1.

(7) Repeat from step 4 until the number of communities satisfies user requirement or all nodes have been assigned to the communities.

Semantic communities in SLN can also be discovered by making use of the clustering features of the reasoning rules and the classification trees.

An SLN is rule-restrict if all of its semantic relations appear in its *Rules*.

A rule cluster of the reasoning rules is a set of rules, which can only reason with the rules within the rule cluster. The minimum rule cluster on rules is such a rule cluster that cannot be further partitioned into smaller rule clusters.

The above definition implies the following characteristics.

Characteristics

(1) *The minimum rule cluster is a connected graph of rules that can reason with each other.*

(2) *If there exists a rule that can reason with the rules in two rule clusters, the two rule clusters are the same or can be merged into one.*

(3) *There is no overlap between the precondition set and post-condition set of rules.*

An SLN can be very large, but its rule set is usually much smaller. The rule cluster can help efficiently determine relevant semantic communities.

The maximum partition of *Rules* is such a reasoning rule set $\{Rules_1, \ldots, Rules_n\}$ that each of which is a minimum rule cluster.

For a given rule-restrict SLN and a rule set Rules, if Rules can be clustered into a set of rule clusters $\{Rules_1, \ldots, Rules_n\}$, *then SLN can also be partitioned into a set of semantic communities* $\{SLN_1, \ldots, SLN_n\}$ *and semantic reasoning of* SLN_k *depends only on* $Rule_k$ (k=1, 2, …, n).

Discovering semantic communities enables operations on SLN to be localized within a local SLN specific to a rule cluster. It also enables operations on the SLN to keep focusing on a specific semantic community while semantic communities keep changing.

The rule-restrictd SLN is partitioned into n semantic communities $\{SLN_1, \ldots, SLN_n\}$ *according to the maximum partition on the rule clusters* $\{Rules_1, \ldots, Rules_n\}$. *Finding SLN closure* SLN^+ *is equivalent to finding each* SLN_k^+ (k=1, …, n), *i.e.*, $SLN^+ = SLN_1^+ \cup \ldots \cup SLN_n^+$.

The following is a strategy for calculating semantic closure.

Discover semantic communities in SLN first, and then calculate the SLN^+ *of each semantic community*.

The following are some characteristics.

Characteristics

(1) *The operation result of deleting a semantic link in SLN is the same as deleting it in its semantic community*.
(2) *Adding a semantic link to SLN is the same as adding it to its semantic community. If a relation does not exist in Rules, itself is a community*.
(3) *If all relations in the new rule belong to a rule cluster, adding this new rule does not influence the other rule clusters*.

Strategies for adding a rule to Rules

(1) *If all relations in the new rule do not appear in the existing rule cluster, the new rule forms a new rule cluster*.
(2) *If all relations in the precondition of the new rule belong to a rule*

cluster, the new rule belongs to the rule cluster.

(3) *If all relations in the precondition of the new rule belong to a rule cluster, and all relations in the post-condition of the new rule belong to another rule cluster, the new rule and the two rule clusters can be merged into one.*
(4) *If relations in the precondition or postcondition of the new rule belong to different rule clusters, merge these clusters and the new rule into one cluster.*

Although a large SLN may be defined by different people and at different times, operations on SLN can be localized within appropriate semantic communities regulated by the rule clusters.

Using classification semantics to discover communities in the SLN is the most direct and efficient way if its semantic space is available.

Semantic distance between communities is the semantic distance between corresponding concepts in the classification tree, each of which is the nearest common concept of the nodes of a semantic community.

Given a semantic distance function, the following approach can construct a semantic community hierarchy bottom-up for SLN.

(1) Take every semantic node of the given SLN as a semantic community.
(2) Merge the semantically closest semantic communities into one community according to the semantic distance measure.
(3) Repeat from step 2 until there are only two communities left.
(4) Recover the semantic links between semantic nodes within each community according to the given SLN.

The top-down approach to constructing the semantic community hierarchy consists of the following steps:

(1) Take SLN as an initial community.
(2) For every semantic link, calculate the semantic distance between the connecting nodes.
(3) Delete the semantic link with the longest semantic distance until the SLN is separated into two communities.

(4) For each current community, do from step 3 until every community only includes an isolated node.

The basic premise of the above two approaches is that two semantic nodes far away from each other in the semantic space should belong to different semantic communities, and the removal of any semantic link does not influence the semantic distances defined in the semantic space. More algorithms can be designed for discovering the semantic communities in SLN by using different levels of semantics.

For a general SLN, we can find the SLN-restrict first by using the semantic space, set the non-restrict SLN apart as a non-restrict community, and then restrict the non-restrict semantic nodes that link to the restrict nodes by semantic links. The percentage of the restrict SLN to the SLN reflects the restrict degree of the semantic space.

The above discussion shows a *semantic facet of the localization principle*.

2.19.9 *SLN-based relation and knowledge evolution*

Putting relation, knowledge, requirement, and services into the integrated cognitive processes in the cyber-physical society can understand them in-depth and therefore can make full use of them.

Social semantic link networks can be established among humans, annotations (tags), metadata, symbolized contents, knowledge and behaviors as shown in Fig.2.19.9.1 (H. Zhuge, The Knowledge Grid Environment, *IEEE Intelligent Systems*, 23(6)(2008)63-71).

Relations and knowledge generate and evolve through the networks with the following behaviors.

(1) Query and answer.
(2) Recommendation.
(3) Cooperation.
(4) Social behaviors such as employment and marriage.

(5) Reasoning over the semantic link networks.
(6) Information retrieval, extraction and summarization.
(7) Clustering humans and indicators (words) according to the usage relations between humans and indicators, the resource-mediated relations between humans, and the human-mediated relations between indicators. These clusters represent massive selection preference of semantics.
(8) Accumulating relations by discovering relations between symbolized contents, between humans, and between humans and meta-data, by clustering texts and then constructing content classification hierarchies, and, by discovering communities in the semantic link network.
(9) Accumulating problem-solving knowledge by asking questions and obtaining answers, making generalizations, and linking questions and answers to the semantic images in minds.
(10) Using relations to explain contents, using problem-solving knowledge to explain symbolized contents and relations, and using relations to complete the problem-solving knowledge.
(11) Managing a scalable structure of resource organization, which can adapt to the evolution of community semantics due to continuous change of humans and their recommended contents. Humans do not have to know the underlying structure of resource organization.

The SLN–based evolution substantially changes the way of knowledge acquisition in traditional knowledge engineering, from individual experts or knowledge engineers to massive contribution of knowledge. This change realizes a harmony:

One for everyone and from everyone.

Web 2.0 provides a Web-based platform for massive codification of knowledge.

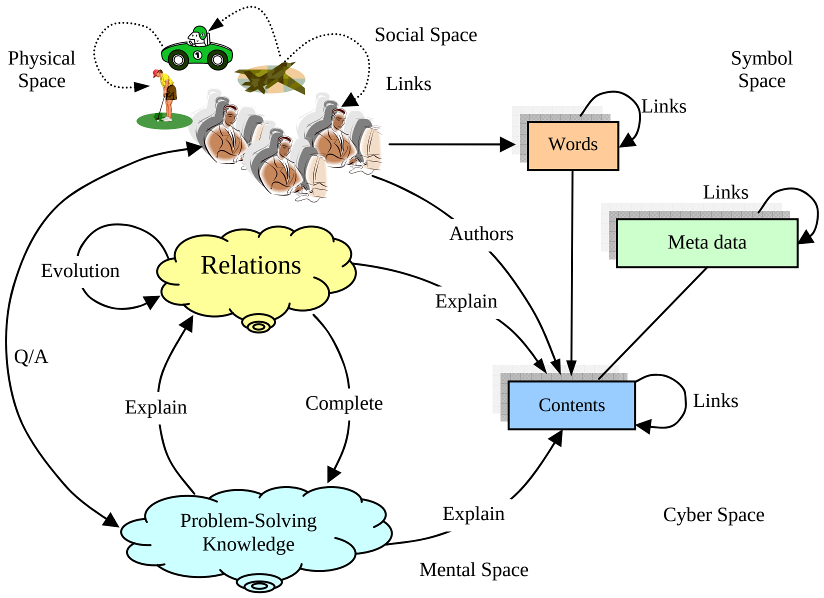


Fig. 2.19.9.1. Knowledge and relations evolve with operating a cyber-physical-social semantic link network.

2.19.10 *Building and performing semantic images*

Human individuals build and evolve semantic images in minds while interacting with each other in the social space and with the objects in the physical space and in the cyber space. Individuals may build different semantic images while involving in the same interaction process.

Let's observe how singers and audiences build and perform their semantic images.

Rhythm is the most basic language of human beings. It is widely used in performance arts. It is a kind of timed movement of such resources as event, sound and language in and through space.

Artists use rhythm and tune to organize music notes and words, and use speed to control the performing processes. Music instrument players and singers have established rich performing skill and semantic links between music notes and sounds in the mental spaces through learning process. A good music player or singer can quickly establish the semantic image of the music, and can perform with the emerging semantic image close to the author's semantic image, with the emotion that matches the author's emotion, and with the appropriate performing skill.

As depicted in Fig.2.19.10.1, a singer can build the semantic image, which consists of semantic link network of music notes, semantic link network of words in song, semantic link network of scenes, semantic links between music notes and sound, semantic links between pronunciation and word, semantic links between emotion and sound, and semantic links between scenes and words. Natural language is the grand for interaction between humans when establishing these semantic link networks.

The audiences build their semantic images when listening and watching. Audiences' semantic images may be different from the singer's semantic image as they may not have the same group of SLNs in mind. Different audiences may build different semantic images since they may have different links between emotion and music, and different links between scene and word.

Individuals evolve their semantic images through microscopic closed loops and macroscopic closed loops of interactions and flows (H.Zhuge, Semantic linking through spaces for cyber-physical-socio intelligence: A methodology, *Artificial Intelligence*, 175(2011)988-1019).

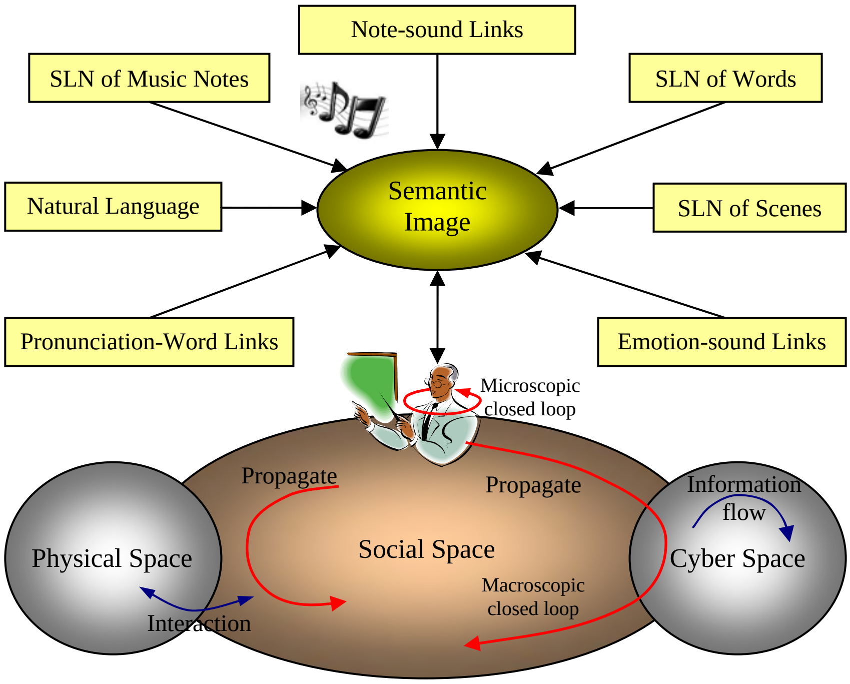


Fig. 2.19.10.1. Building and propagating semantic images.

Similar semantic images will be propagated through interaction between different individuals in the social space.

Although some audiences are not able to sing songs, they have built the similar semantic images because their emotion will respond to the songs to a certain extent when hearing them again. Some propagation will pass through the cyber space and social space. A music or song becomes popular when its semantic image has been built in large number of individuals in the society.

2.19.11 *Structure and networking rules of social space*

A society consists of *individuals*, *interactions*, *roles*, and *organizations*.

Individuals include humans and artifacts that provide material or mental services for humans.

Interaction includes *communication*, *flow*, *exchange*, and *influence*. Flow includes material flow, content flow, knowledge flow and service flow. Exchange consists of two-way flows, which can be of different types, e.g., using money to buy goods. Influence is the propagation of the changes of the statuses of individuals.

An individual can play multiple roles in different organizations. Interactions pass through organizations through individuals playing different roles.

Organization mainly includes the following types:

(1) *Family organization*, which is a natural organization. The most basic family organization consists of parents and children. It can be extended to a family tree of different scales.
(2) *Educational organization*, including schools of various levels and research institutions.
(3) *Work organization*, including public service institutions and companies.
(4) *Temporal organization*, which temporally contains individuals for special purpose, e.g., train, airplane, queue, and restaurant. In addition to the main purposes, temporal organizations also provide opportunities for individuals to communicate with each other.

Different organizations contain different social relations, e.g., family relation in family organization, colleague relation in work organization, and classmate relation in educational organization. The characteristics of different organizations are also suitable for propagating different types of knowledge.

Fig. 2.19.11.1 depicts the basic structure of social space, where the dotted arrows in red color representing flows. Flows will propagate through social relations (e.g., family relation, colleague relation, and classmate relation), the communication links, and the role links.

The following is the principle of flows through the structure.

Principle. *The effectiveness of propagating flows depends on the selection of appropriate individuals for interaction*.

The following are the rules for selecting appropriate individuals for communication:

Networking Rule 1. *The individuals who do not share concepts with the selector in the mental spaces should not be selected for communication*.

For example, a computer scientist should not select the individual without basic computer concepts for communication. The knowledge in an individual mental space is reflected by the roles of the individuals and the documents that he/she read or wrote.

Networking Rule 2. *The individuals who communicated before or communicated with an individual (or device) familiar with the selector should take higher priority to be selected for communication*.

This is because the individuals who communicated before or those communicated with a familiar individual should share some concepts. For example, the colleagues, classmates, or the individuals in the same community should take higher priority for communication than strangers.

Networking Rule 3. *The individuals who speak the same language as the selector should take higher priority to be selected for communication than those who speak different languages*.

This is because individuals who speak different languages need translators, who will slow down communication and may distort understanding.

The basic structure of the social space, the interactions through links, the principles of flow, and the networking rules reveal a pattern of knowledge society.

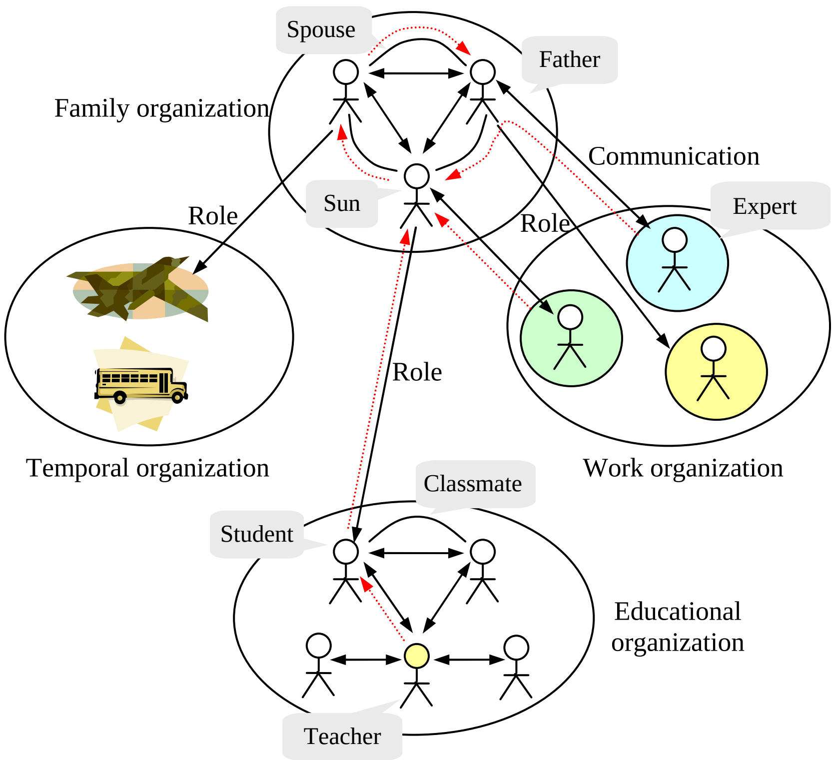


Fig. 2.19.11.1. Social space organization and knowledge propagation.

2.19.12 Communication rules and principles

Reasoning on semantic link network depends on the rules on semantic links. Semantic node, semantic link and semantic community involve in semantic networking. Communication through semantic links between human individuals is more relevant to the topic of communication content.

Fig. 2.19.12.1 depicts two communication rules. The left part depicts the following rule.

Communication Rule1: A←α→C, C←α→B ⇒ A←α→B, which means that if A and C are the appropriate individuals of communicating topic α, and, C and B are the appropriate individuals of communicating topic α, then A and B are the appropriate individuals for communicating topic α.

A←α→C implies that A and C share common mental semantic link networks on topic α, that is, α is a semantic node or indicated by a semantic community. C←α→B implies that C and B share common mental semantic link networks on α. Therefore, A and B share common mental semantic link networks on α, indicated by A←α→B. But, for semantic link reasoning, A—α→C, C—α→B ⇒ A—α→B may not be true because semantic link reasoning depends on reasoning rules. Using Rule1, no more topics can be generated during communication.

The left part of the figure depicts the following rule.

Communication Rule2: A←α→C, C←β→B ⇒ A←{α,β}→B, which means that if A and C are the appropriate individuals for communicating topic α, and, C and B are the appropriate individuals for communicating topic β, then A and B are the appropriate individuals for communicating topics α and β.

Using Rule2, individuals A and C can communicate with each other on topic β in addition to α, and individuals C and B can communicate on topic α in addition to β.

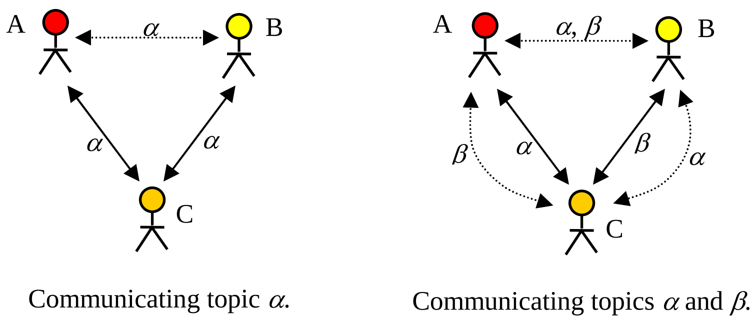


Fig. 2.19.12.1. Communication rules.

The Role of Language

Communication Rule 1 and Rule 2 assume that individuals use the same language. So far, language is the only way for humans to know each other. Language also plays the key role in constructing the structure of the social space, mental space and cyber space.

Fig.2.19.12.2 depicts communication through different languages, where "A: L_1", "B: L_2", and "C: L_1, L_2" denote that individual A speaks language L_1, B speaks language L_2, and C speaks language L_1 and L_2.

Individual A can communicate with C on topic α in language L_1, and individual C can communicate with B on topic β in language L_2. As the effect of knowledge fusion in C's mind, individual A can also communicate with C on topic β in language L_1, and the knowledge of B can flow to A through C. Individual C can also communicate with B on topic α in language L_2, and knowledge of A can flow to B through C. However, the knowledge flow through C depends on C.

Individual A and individual B cannot communicate with each other directly since they do not speak the same language. Individuals A and B can communicate with each other on topics α and β if C can play the translator role. In a knowledge creation organization, C can add own opinion when communicating with A or B rather than translate only.

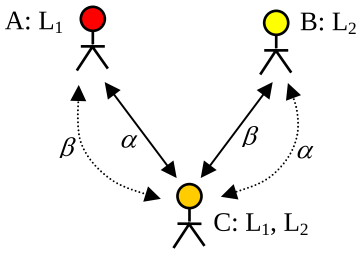


Fig. 2.19.12.2. Communication through different languages.

In a competitive self-organized society, C takes better position than A and B in obtaining knowledge because C can reserve some key symbolized contents and knowledge for keeping better position in competition. C can also attract more links because C can provide more content, knowledge, and services. In addition, C can get more profit by selling symbolized content (systemized semantic symbols), knowledge and services.

The above analysis indicates the following principles.

Communication Principle 1 (Language barrier of flows). The individual who knows more languages take better position in attracting flows of resources (contents, knowledge, services, etc).

Communication Principle 2 (Language barrier of centrality). The individual who knows more languages takes better position in gaining the centrality in the evolution of social networks.

Individuals will learn languages if the advantages of language become eminent. Therefore, we have the following principle.

Communication Principle 3 (Social effect of learning language). The advantages of the individuals who know more languages decrease with the increase of the number of individuals who know more languages. The advantage follows the law of marginal utility.

However, the above principles may not be eminent in a non-autonomous society because the activities of the individuals who know more languages may be restricted.

In addition to communication, languages profoundly affect the brain function and structure (J. Krizmana, et al. PNAS, 2012, DOI: 10.1073/pnas.1201575109; Crinion J, et al. Science 312(2006)1537–1540; Kim KHS, et al. Nature 388(1997)171–174).

However, language itself does not lead to intelligent behavior. Knowledge plays the key role in using language and performing intelligent behaviors.

Communication Principle 4 (Effectiveness). The effectiveness of communication depends on the consequence effects in the mental space, physiological space, psychological space and physical space.

Languages will evolve with various interactions in the Cyber-Physical Society.

2.19.13 Influence between mental space and social space

Humans communicate with each other through various languages, which are means of conveying concepts and thoughts in mind. Concepts play the key role in communication, and they are fundamental in philosophy although there are different explanations (http://plato.stanford.edu/entries/concepts/). What is the relation between concept and language? Some philosophers argue that natural language is necessary for having concepts, while others argue that concepts are prior to and independent of natural language.

The idea of interactive semantics argues that interaction takes place before the generation of language (H.Zhuge, Interactive Semantics, *Artificial Intelligence*, 174(2010)190-204), therefore *exploring interaction should take higher priority than studying language in semantics research*.

Interaction can be based on various indicators (language component, artifact, image, sound, etc). The following is the process of influence between the mental space and the social space through delivering indicators.

One person motivates concepts in mind according to internal motivation or in response to external stimulation, uses indicators (e.g., words) to indicate the concepts, and delivers the indicators to the other person(s) through the social network they are involved in. The person who receives the indicators activates the concepts in mind. The active concepts activate the other concepts through the mental semantic link network. The receiver may also use indicators to indicate the active concepts and deliver the indicator through social networks.

Continuously receiving indicators, the mind emerges semantic images. One semantic image can be linked to the other semantic images through the links between concepts to form a larger semantic image or a sequence of semantic images. The semantic images emerging one after another along the time dimension constitutes a *semantic movie*.

One or a set of indicators is meaningful if it can inspire semantic image or semantic movie in one's mind.

The following are some characteristics of the semantic link network of concepts in mind:

Characteristics

(1) *Diversity*. One indicator can activate multiple concepts in mind. Different orders in a set of indicators can render different semantic images. Sensing a sequence of indicators, different individuals may emerge different semantic movies in minds.
(2) *Priority*. The priority of emerging concepts is determined by the ranks of the concepts, the link structure, the times of being activated, and the current emerging concept. The often activated or high-rank concepts have higher priority. This priority is the basis of attention.

(3) *Temporality*. An active concept or semantic image can only last for a short period of time. Beyond the active period, the concepts will be inactive. The mind provides a stage for performing concepts or semantic images.
(4) *Relevancy*. A semantic image prepares to re-emerge when one of its concepts is activated. When one semantic image is emerging, the linked semantic images prepare to re-emerge. The priority of emerging a semantic image depends on the number of active concepts and their ranks.
(5) *Asynchrony*. There is a lag from sensing the indicator to emerging semantic image. If the speed of sensing is too fast, there will be no time to emerge semantic image. People can still sense the indicator even though no semantic image emerges. For example, if one scan text too fast, he/she has no time to emerge semantic images, therefore cannot understand it. This indicates that *the mind uses different components to process indicators and to emerge semantic images*.

Linking mental space to social space

Fig. 2.19.13.1 depicts the scenario that one mental semantic link network influences the other through social semantic link network. The larger nodes denote the currently more active concepts. A semantic image can be one concept or a set of related concepts. The nodes in yellow color denote semantic images.

Some concepts can be expressed in language while others are only in mind. *The concepts (i.e., internal-only concepts) that cannot be expressed in language may be linked to and influenced by those concepts that can be expressed by language*. The internal-only concepts can influence other concepts and render semantic images.

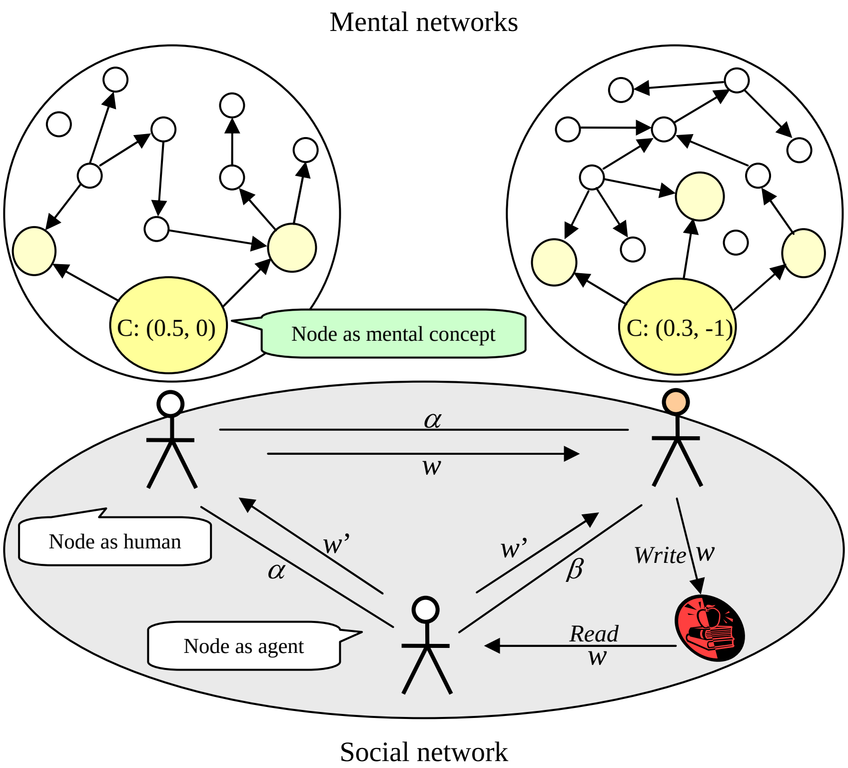


Fig. 2.19.13.1 Influence between mental semantic link networks through social semantic link network, where α denotes a semantic relation, *C* denotes a concept or semantic image in mind, and *w* denotes an indicator (e.g., word) that indicates *C*.

Activating a concept in different minds may have different effects because different minds have different semantic link networks and a concept has different emotional statuses in different minds. Therefore, a semantic link network of human individuals can pass through the influence of emotion.

The interesting issue is how to use the influence through links to control emotions.

The following is a semantic link in the mental space, where $Rank_A$ and $Rank_B$ are the ranks of concepts A and B, α is the relation between the two concepts (e.g., sequential, co-occurrence, cause-effect), and *ES* is the emotional status, which can be divided into multiple levels from negative to positive, e.g., negative (–1), neutral (0), or positive (+1):

$$\mathrm{A}(Attributes, Rank_A, ES_A) \text{ —}\alpha\text{— } \mathrm{B}(Attributes, Rank_B, ES_B).$$

A concept has a rank determined by the number of connected concepts and their ranks. The inactive concepts do not influence the other concepts. The emotional status of a concept is influenced by the emotional statuses of the connected active concepts. The following is an approach to computing the emotional status of concept *P* at time *t*+1, where *k* denotes the active concepts linked to *P*, and *f* is a function that determines the influence from these concepts:

$$ES_{\mathrm{P}}(t+1) = ES_{\mathrm{P}}(t) + f(Rank_{\mathrm{P}}(t), \cdot \sum_{k=1}^{n} Rank_{\mathrm{k}}(t) \cdot ES_{\mathrm{k}}(t)).$$

The rank of the emerging concept will get a certain rise. The rising rank will influence the rank of the connected concepts, and the influence will be propagated in the network.

In the semantic link network of concepts, the neutral concepts separate the positive concepts from the negative concepts. The high-rank negative concepts should reduce their ranks and then become neutral concepts before becoming positive concepts.

The emotional status of concepts could be changed by adapting links or ranks of the concepts.

Working in the social space, the social semantic link network consists of nodes as human individuals and links as social relations. It enables one person to communicate with the other person sharing common concepts by delivering indicators through certain semantic links. The following is the semantic link of the social semantic link network, where w is the indicator forwarding from person X to person Y at time t through relation α.

$$X(Rank_X, t) \text{—} \alpha{:}w(t) \rightarrow Y(Rank_Y, t).$$

The rank of the node in the social semantic link network is determined by the number of links it has and the ranks of the connected nodes. The potential energy of node and the motion energy of operating a semantic link network were proposed as a new measure of a dynamic network (H.Zhuge, Semantic linking through spaces for cyber-physical-socio intelligence: A methodology, *Artificial Intelligence*, 175(2011)988-1019).

A series of indicators W=$<w_1, \ldots, w_n>$ will be delivered from X to Y during a time period $\Delta t=[t_1, t_2]$, so the semantic link should consider a time period as follows:

$$X(Rank_X, \Delta t) \text{—} \alpha(\Delta t){:}W(\Delta t) \rightarrow Y(Rank_Y, \Delta t).$$

Y receiving W during Δt will emerge a concept (e.g, C in Fig. 2.19.13) or a semantic image consisting of the concepts indicated by these indictors.

2.19.14 Application in cognitive-behavioral therapy

Cognitive-Behavioral Therapy (CBT) is a kind of treatment for mental health problem (A.T. Beck, Cognitive Therapy and the Emotional Disorders. International Universities Press Inc., 1975). It combines cognitive therapy and behavioral therapy (M.B. Keller, et al., A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression, *New England Journal of Medicine*, 342(20)(2000)1462–1470). Some empirical evidences show that CBT is effective for treating the problems concerning mood, anxiety, personality, and psychotic disorders.

Depression is a state of low mood and aversion to activity that can affect a person's thought, behavior, feeling and well-being. CBT is one method of treatments, at least an assistant means.

The semantic link network can be an analysis means for CBT. *The basic viewpoint is that depression is not only a mental problem but also a social problem*. Wrong links, ranks and interactions in the social space and the mental space could lead to depression. Adapting the links, ranks and interactions is a way to solve the problem.

The emotional statuses of the often emerging semantic images influence human mental status. Depression happens if negative mental concepts and semantic images often emerge. As the consequence, the ranks of the negative concepts and semantic images and the connected links become higher. The high-rank concepts and semantic images take the priority in the competition of emerging. The “rich get richer” phenomenon forms an *internal negative loop*, which makes depression more serious. An inter-person negative loop is formed through continuous interaction between persons on negative concepts.

A solution is to change the statuses of the semantic images from negative to neutral or positive by operating the social network that the patient involved in.

The following operations can influence the emotional status of a person:

(1) *Adapt interaction*. Reduce the times of interaction between the patient and the person who has many common high-rank negative concepts (often delivers negative emotion), and increase the times of interaction between the patient and the person who has many common high-rank positive concepts (often delivers positive emotion).
(2) *Adapt rank*. Reduce the rank of the person who delivers negative emotion through the social network the patient involved in. Raise the rank of the person who often delivers positive emotion through the social network. For example, change the person’s position and responsibility in an organization.

(3) *Remove links*. Remove the link that often delivers negative semantic indicator in the social network. For example, change the patient's workplace to remove old work relations.
(4) *Increase links*. Increase the link that often delivers positive emotion in the social network. For example, encourage the patient to participate in more social activities to interact with the persons who can often deliver positive emotion. The linked persons should share common positive concepts.
(5) *Interact with appropriate resources*. Interact with the resources (e.g., text, video, audio) that indicate many positive semantic images.
(6) *Increase positive nodes*. Encourage the patient to learn new positive concepts and link to the persons that often deliver positive concepts.
(7) *Change the attributes of the negative concepts*. Abnormal attribute values of some concepts influence emotional statuses. Due to the bias of cognition, the attributes of some negative concepts do not appropriately reflect the external world. The values of attributes may be changed by knowing more instances of the concept.
(8) *Invite the third party*, which is important in forming the loop of delivering positive indicators, and getting rid of the loop of delivering negative indicators.

The key is to know the high-rank negative concepts and semantic images in the patient's mind and the involved social network so that the other persons (e.g., family members) can avoid talking to the patients about the negative concepts and help change the social network through the above operations.

A web-based interactive e-health system can help detect the negative, neutral and positive concepts and semantic images in the patient's mind through a set of querying and answering, and then give necessary suggestions or play games with the patient. It is particularly useful in the places where appropriate therapists are not available.

Analyzing the mental semantic link network and the social semantic link network of the patient as well as the influence between the mental space and social space provides a new means of therapy.

2.20 Principles of Mental Concepts

Mental concepts are the basic components of semantic images in mind. Some concepts are the reflection and generalization of physical objects or events, while some others are too abstract to correspond to the physical objects or visible events. A mental concept is formed and enriched through learning, interaction and experience in multiple spaces. The process concerns the internal microscopic close-loop and the external macroscopic close-loop (H.Zhuge, Semantic linking through spaces for cyber-physical-socio intelligence: A methodology, Artificial Intelligence, 175(2011)988-1019).

A concept is rendered by the following facets of elements:

(1) *Attributes*, which render the concept through the attributes reflected in the physical space (e.g., car's color, shape, and internal space) and in the social space (e.g., price, reputation, and sales).
(2) *Links*, which connect to the other concepts through semantic links such as generalization and specialization. For example, the concepts of car should be linked to the concepts of road and human.
(3) *Functions*, which transform input into output, or change the statuses of attributes.
(4) *Operations*, which reflect the behaviors of changing statuses, and instruct the physiological organisms to act (e.g., instruct hands and legs to drive a car with certain speed).
(5) *Experiences*, which link feeling to the attributes and operations, e.g., the feeling of driving or moving.

Concepts are linked to form semantic images or semantic movies through interactions between the elements of multiple facets. The interactions form closed loops of experience. Fig. 2.20.1 shows the links between the concept of human individual and the concept of the external object. For example, human individual uses eyes to see the operation mechanisms and read the specifications, and use hands and legs to operate them. The eyes and the senses of time and space feel the effect of operation (e.g., motion). The relations between the process of motion and the process of the operations are established through continuous feeling and operations.

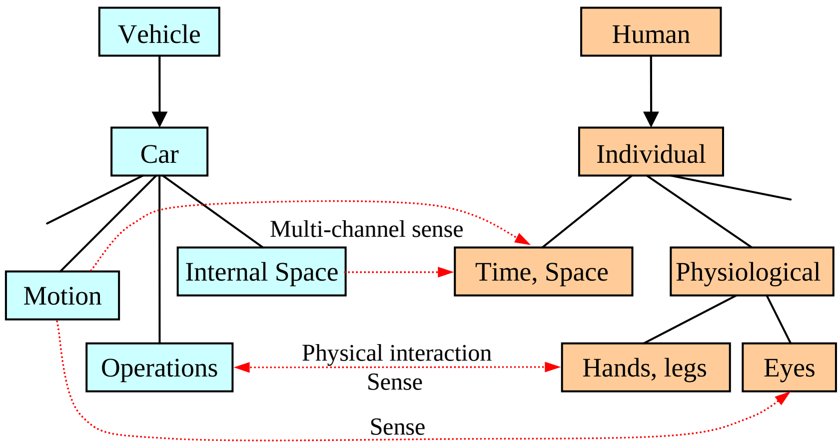


Fig.2.20.1. Linking concepts through interactions.

As shown in Fig.2.20.2, semantic link networks of concepts are built and keep evolving in lifetime with continuous learning and experiencing in the physical space and socio space and reading and writing through the symbol space.

Learning language helps the brain build the function of processing symbols according to the grammar and experience of reading and writing, and linking symbols to mental concepts.

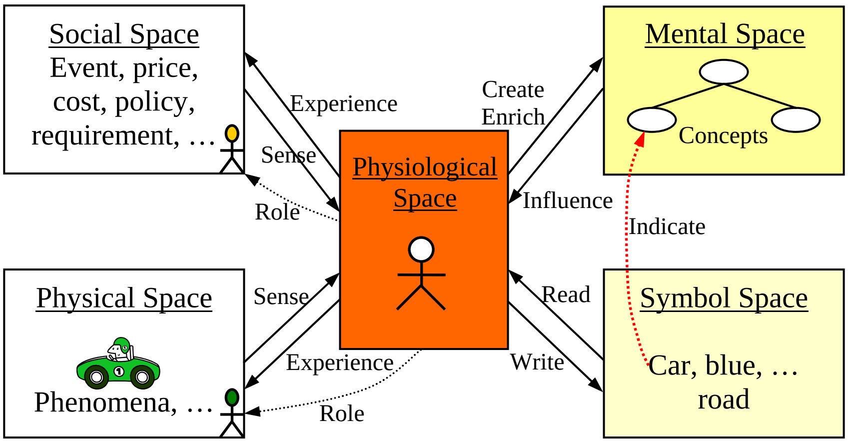


Fig. 2.20.2. The formation of mental concepts.

The most grand challenge problem is that there is no direct way for any person to know the mental concepts of the other.

The following principle explains why humans can indicate mental concepts.

Principle of concept 1. People with similar learning process and experience have similar concepts in minds.

The learning process includes learning classifications and language. Co-experience in different spaces generates feelings, which enable minds to make generalization and to enrich concepts. Generally, humans share similar genes and inherit similar culture, so have the inner condition and social condition to generate the similar concepts in minds. Therefore, we can generalize the principle by incorporating the physiological, social and physical aspects as follows:

Individuals have similar concepts in mind if they have similar physiological space, learning process and experience in physical space and social space.

Although there may have some innate concepts in mind, most concepts are generated through learning and experience. To interact with each other, humans use symbols (or sound) to indicate mental concepts. Learning establishes the mapping between words and mental concepts. However the mapping is not one-one mapping, one word can be used to indicate multiple mental concepts, and multiple words can be used to indicate one mental concept.

Emerging concepts with different emotions may have different influences in the psychological space and physiological space through behaviors. Studying the relation between mental concepts and behaviors can help people to understand each other.

People use a sequence of words to indicate a mental semantic image consisting of multiple concepts. Reading an article can emerge a network of semantic images in mind. People sharing more concepts have more common learning process and experience, so they know more common symbols.

Principle of concept 2. If two individuals have used more common symbols in multiple times of communication, their minds have more common mappings between concepts and symbols.

This principle indicates that people sharing more concepts can better understand each other when using symbols for communication, e.g., via emails. More times of communication provide stronger support for the mappings.

People in the same community should have used more common symbols for communication than those in different communities, so they share more concepts than those in different communities. According to the above principle, they can understand each other better than those in different communities.

People who have similar statuses in the social space should have similar activities, for example, students involve in learning activities, professors involve in teaching activities, and group leaders involve in similar activities although they are in different areas.

Principle of concept 3. Individuals have similar statuses in the social space often involve in similar events.

The statuses concern the ranks in the network, the link structure, and the statuses of the connected persons. This principle explains that great minds think alike.

Current technologies enable people to display the physiological behaviors of brain by using fMRI (functional magnetic resonance imaging) and EEG (electroencephalogram is a measure of brain waves), but there is a big gap between physiological behaviors and mental concepts. Fig. 2.20.3 shows the big gaps between mental concepts, physiological behaviors (e.g., brain images), and symbols.

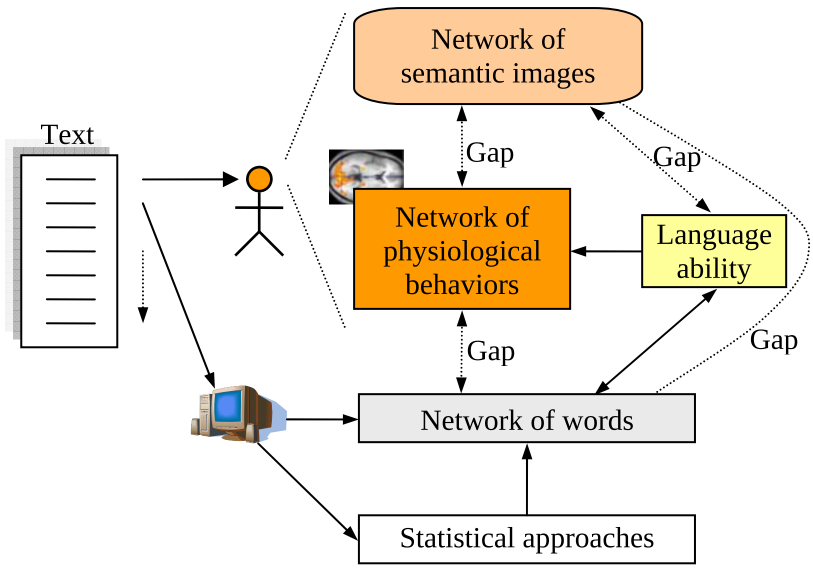


Fig. 2.20.3 Gaps between semantic image, brain image and symbols.

Scanning text, computers can easily remember all of the symbols and can form a network of symbols, but computers are not able to generate the semantic images as humans because the generation of the semantic images in mind concerns the spaces (e.g., social space, physiological space, and psychological space) that machines do not have and cannot experience, even if humans assign rules to them.

Only in some specific domains, machines can reduce the gaps by establishing one-one mapping between symbols and concepts. In the open domain, establishing one-one mapping is not feasible due to the diversity and richness of humans and society.

Knowing the semantic link networks of mental concepts is a way to effective interaction.

2.21 Discussion: Philosophy, Psychology, Language, and Semantics

Social relations are firstly studied by Laozi (576–BC), Confucius (551–479 BC), and Socrates (469 BC–399 BC). The formal generalization of relation can trace to the invention of set theory in 1874.

Hume (1711-1776) regarded *causation* as *a kind of mental association by associating constantly co-occurred events.*

"A caused B" is equivalent to "Whenever events of type A happen, events of type B follow", where the word whenever means by all possible perceptions.

He thought that humans have no perceptual access to the connection, but humans naturally believe in its objective existence. The causal relation can be also regarded as an expression of a functional change in mind, by which some events in experience could be referred.

He also believes that *two things are relational because they have constant conjunction* (*emerge into one, or one after the other at time, in space, or on logic*), *and one thing will remind people of another.*

SLN concerns all relations, the models of networking, and the laws and effects of networking. In the SLN method, relations can be in the social space, the cyber space, the physical space and the mental space. Relations in one space reflect the relations in the other space. Different minds may reflect different relations from the same set of individuals (*objects, humans, events, etc.*) *in the same space. The mental space can reflect relations not only in one space but also through spaces.* For example, the relations between behaviors in the social space can be detected by mining the data in the cyber space.

Wittgenstein (1889–1951) argued that the meaning of words is constituted by the function they perform within any given language-game (i.e., use cases). SLN uses a semantic space to determine the meaning of indicators (e.g., words) that indicate relations at the first stage. In the cyber-physical society, *the meaning of an indicator is rendered from not only the symbol space but also the physical space, social space and mental space*.

SLN is also relevant to the notion of *self*.

Hume regarded the self as nothing but *a bundle of interconnected perceptions linked by the property of constancy and coherence*.

John Locke (1632–1704) defines the self as conscious thinking thing that is sensible, concerned for itself, and conscious of emotion. He argued that the mind grows from empty through experience (J. Locke, An Essay Concerning Human Understanding, 1690).

Self is a basic cognition mechanism that reflects relations and determines the target of linking and the content of interactions. SLN can model and analyze self by creating the internal semantic link network and external semantic link network. *The internal network and the external network are linked through sense, reflection, and language interaction*.

How is self generated?

Physiological space and social space constitute the ground of generating self. Physiological system is the inner factor of generating self. For example, animals generate the motivation of eating food in the external physical space because of the sense of hungry generated from the physiological system. Social space is the external factor of generating self. For example, social competition and natural evolution generate the motivation of survival.

A self needs a reflection, from the cyber space, physical space, or social space. Self also needs confirmation in a community.

The cyber space has no physiological and social grounds to grow self.

Psychology is the study of mind and behavior. Research is often limited by the following factors: single dimension, isolated objects, small sample, and pre-defined tasks. The SLN provides tool and method for modeling and analyzing mind and behavior in a large relational network. In the cyber-physical society, behaviors and sense will be reflected in more spaces, and can pass through spaces. Therefore, psychologists can observe the real phenomena in addition to do experiments on behaviors with pre-defined tasks.

William Molyneux raised the following issue to Locke in 1688: "If a blind person were suddenly able to see, would he be able to recognize by sight the shape of an object he previously knew only by touch? Presented with a cube and a globe, could he tell which was which just by looking?" Locke answered: "He would not be able with certainty to say which was the globe, which the cube, whilst he only saw them," he said: "though he could unerringly name them by his touch."

This question may lead to fundamental thinking: *Whether there is an innate conception space common to both sight and touch or not*? Or, *whether we learn that relationship only through experience or not*? This concerns the core of philosophy of mind. Locke's opinion was confirmed by latter experiments (N.Bakalar, Study of Vision Tackles a Philosophy Riddle, *The New York Times*, April 25, 2011): The mind cannot immediately make sense of what the eyes see, and the recovered blind people cannot distinguish the two objects.

SLN and cyber-physical society can help study this philosophical issue. The following are two causes:

The first cause is that the channel of sight and the channel of touch are separated. This can be confirmed by the following test:

The child who never sees the real elephant cannot tell the feeling of touch by just watching a picture or video of an elephant.

The second cause is that the two things are isolated. If the blind person knows one relation (e.g., *aboveOf*, *leftOf*, *rightOf*) between the two objects as shown in Fig. 2.21.1, he/she should appropriately link the feeling of touch to the corresponding object although he may not recognize the two objects only by sight immediately. Afterwards, the person should be able to link the sight features to the touch features quickly. This indicates that humans can establish new concepts based on existing concepts.

People have built rich relations between semantic images of various granularities in mind through life-long learning and experience. These relations help people to emerge new semantic images. This is why people can understand each other in conversation even though some words are missed.

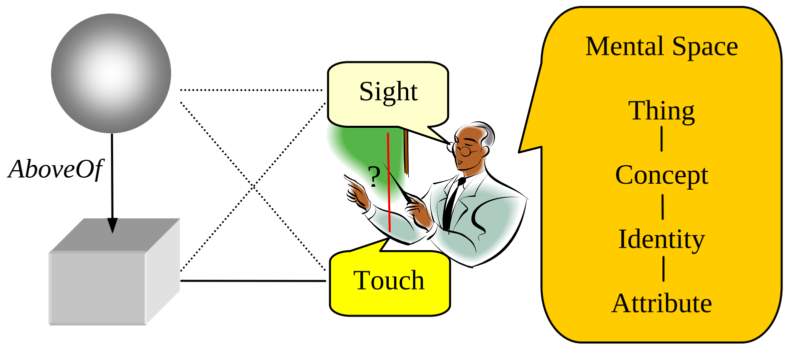


Fig. 2.21.1. Establishing new concepts with the help of relations.

The following mental process involves in establishing the initial concepts of the two physical objects:

(1) Link the touch features to the appropriate object in the physical space according to the relation between the object and the process of touching it (the motion and position of hand in the physical space).
(2) Link the touch features to the sight features through the relation between the objects.
(3) Attaching the sight features and the touch features to the objects as attributes.

The following is an important issue: *What are the most basic concepts or behaviors that enable humans to grow a network of versatile concepts*?

The following are *basic networking behaviors* in the mental space:

(1) Sense the objects in the physical space.
(2) Establish link between self and the objects, link one object to the other through self, and then establish the concept of space.
(3) Sense the individuals in the social space, link self to the individuals, and then link one individual to the other. Family relations will be established firstly.
(4) Link the features of sense to the features of emotions.

We should not forget an important fact that the blind person lives in a social space. People can obtain indirect experience through the social links. A blind person has a mental space and has established semantic images on many objects. The difference is that the concepts in their semantic images have no sight attributes. So, the recognition of objects needs the reflections from multiple channels and multiple spaces.

Natural language is a semantic indicating system created by humans for interaction. Humans acquire the ability of natural language through social interaction. Studies show that languages are processed in certain areas in human brain. Languages are generated and developed through interaction and evolution in a process of Darwin's natural selection.

With language, humans often interact with each other through the symbol space without participation of the physical space. This brings efficiency of interaction and expansion of human imagination, but leads to misunderstandings. People without experience in the physical space on the interaction topic rely on the symbol space to build semantic images. But, the symbol space is limited in ability to help humans to build the semantic images of versatile objects if people do not have relevant experience (direct mapping from the physical space). As shown in Fig.2.21.2, it is hard for individuals A and B to build appropriate semantic images according to the indicator "tiger" if they do not have the experience about tiger.

SLN can be regarded as a general language and method for exploring the semantics through spaces. The semantic space of the SLN can be regarded as the grammar of semantic indicators.

Human minds establish semantic images when sensing external stimulation, which evolves through the microscopic closed loop and macroscopic closed loop of super links through multiple spaces.

There are two types of semantic images: semantic images of physical objects and semantic images of indicators. Different people can generate the same semantic image of one physical object. This can be tested by behaviors and Q/A about the object. But, different people may generate different semantic images of one indicator. This is because different people may involve in different close-loops of hearing, speaking, reading and writing the indicator.

A well-designed semantic indicator system like maps helps humans to build abstract semantic images although people have no direct experience. Some semantic indicator systems are created by team work, where collaborators may have different types of semantic images. As shown in Fig.2.21.2, participation of individual C in the interaction will help A and B to build the appropriate semantic image on "tiger". The possible barrier refers to the distances in the physical space and the social space, as well as to the change of the characteristics of individuals (e.g., the characteristics of human-raised animals are different from those of the wild animals).

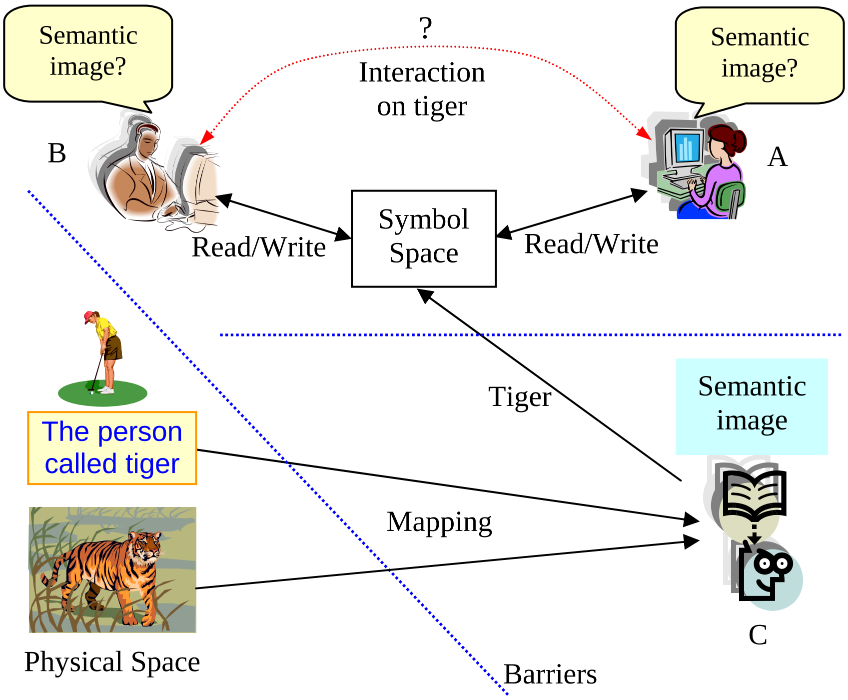


Fig. 2.21.2 The semantic images of physical objects and indicators. The dotted blue lines indicate the possible barriers.

The cyber-physical society will extend human ability by enabling experiencing the cyber space, the physical space and the social space, and reflecting themselves from these spaces while interacting with each other and with individuals in these spaces.

The SLN in the Cyber-Physical Society passes through the cyber space, physical space, socio space, and mental space for extending human ability. The objective existence of the network is linked to the subjective existence. So, research on SLN in the Cyber-Physical Society upgrades previous research in the cyber space. Only in the Cyber-Physical Society, some fundamental effects in the cyber space, social space, and mental space can be revealed.

When I was checking this chapter, Steve Jobs passed away on October 5, 2011. I watched the videos of his past activities several times. The most impressive one is his talk about *connecting the dots* in life toward success. His idea is in line with the following philosophical ideal of the Semantic Link Network:

Appropriately linking appropriate nodes through multiple spaces to wave the meaning of life.

When I was about to finish this chapter, John McCarthy passed away on October 23, 2011. He coined the artificial intelligence and invented the Lisp language. He foresaw service computing, grid computing, and cloud computing in 1961. In 1969, he studied philosophical problems of artificial intelligence. During 1963-1986, he studied commonsense, context, and causal laws, which are the basis of semantics and intelligent systems. John received the Turing Award in 1971. After 2001, he explored the emotions, Internet culture and social networking issues.

I tried to rebuild John's semantic image in my mind by reading his papers when I wrote the "Interactive Semantics" (*Artificial Intelligence*, 174(2010)190-204) and the "Semantic linking through spaces for cyber-physical-socio intelligence: A methodology" (*Artificial Intelligence*, 175(2011)988-1019). His insight on "Human-level AI" (*Artificial Intelligence*, 171(18)(2007)1174-1182) and "Well-designed child" (*Artificial Intelligence*, 172(18)(2008)2003-2014) enlightens an important way of AI development.

It is an interesting issue to build his semantic image on AI by extracting semantic link networks from his papers so that AI researchers can interact with his semantic images.

John left us, but his important semantic images will be rebuilt and propagated through various explicit and implicit links in the mental space, cyber space and social space.

Humans have been relying on languages to indicate semantic images in minds. Only the important semantic images in minds are widely propagated and inherited through generations. In the cyber-physical society, individual mental spaces will be indicated by behaviors and languages. More individual mental spaces will be preserved, and retrieved more easily.

As a general method, SLN can be applied to many special networks such as biological network, brain network, physiological network, economic network, ecological network, and social networks. For example, by creating the gene networks of cancer features, the SLN method can help find the implicit relations, know the evolution rules of the network and the stage of the network in the evolution process, and indicate the way to control the evolution of the network.

Classification plays the key role in specifying the semantic space of the semantic link network. Multi-dimensional classifications can be formalized into a resource organization model.

The following chapter will introduce a multi-dimensional classification space: the Resource Space Model (RSM), in which coordinates of dimensions (axes) play roles like those of latitudes and longitudes in geographical maps as outlined at the beginning of this chapter.

Bibliography

Aberer, K., Cudre-Mauroux, P., Hauswirth, M. and Pelt. T.V. (2004). GridVine: Building Internet-Scale Semantic Overlay Networks, *The 3rd International Semantic Web Conference*, LNCS 3298, pp.107-121.

Adamic, L. A. and Huberman, B. A. (2000). Power-law distribution of the World Wide Web, *Science*, 287, No.24, pp.2115.

Aiello, W., Chung, F. and Lu, L. (2000). A random graph model for massive graphs, *Proc. 32nd ACM Symposium on the Theory of Computing*, New York, pp.171-180.

Albert, R., Jeong, H. and Barabási, A. L. (1999). Diameter of the World-Wide Web, *Nature*, 401, pp.130-131.

Albert, R. and Barabási, A. L.(2000). Topology of evolving networks: local events and universality, *Phys. Rev. Lett.*, 85, pp.5234-5237.

Albert, R. and Barabási, A. L. (2002). Statistical mechanics of complex networks, *Rev. Mod. Phys.*, 74, pp.48-94.

Allan, J. (1997). Building hypertext using information retrieval, *Information Processing and Management*, 33, No.2, pp.145-159.

Amaral, L.A.N. et al. (2000). Classes of small-world networks, *Proc. Nat. Acad. Sci.*, 97, No.21, pp.11149-11152.

Androutsellis-Theotokis, S. and Spinellis, D. (2004). A Survey of Peer-to-Peer Content Distribution Technologies, *ACM Computing Surveys*, 36, no. 4, pp. 335-371.

Anthonisse, J. (1971). The Rush in a Graph, *Amsterdam: University of Amsterdam Mathematical Centre*.

Bachman, C. (1972). The Evolution of Storage Structures, Communications of the ACM, 15, No.7, pp.628-634.

Bachman, C. (1973) The Programmer as Navigator. *Communications of the ACM*, 16(11) 653-658.

Bailey. N. T. J. (1975) The Mathematics Theory of Infectious Diseases and Its Applications, second ed. Hafner Press.

Balakrishnan, H. et al. (2003). Looking Up Data in P2P Systems, *Communications of the ACM*, 46, no. 2, pp. 43-48.

Barabási, A. L. and Albert, R. (1999a). Emergence of scaling in random networks, *Science*, 286, pp. 509-512.

Barabási, A. L., Albert, R. and Jeong, H. (1999b). Mean-field theory for scale-free random networks, *Physica A*, 272, pp.173-187.

Barabási, A. L., Jeong, H., Neda, Z., Ravasz, E., Schubert, A. and Vicsek, T. (2002). Evolution of the social network of scientific collaborations, *Physica A*, 311, pp.590-614.

Barabási, A. L. and Bonabeau, E. (2003). Scale-free networks, *Scientific American*, May, pp.50-59.

Barrat, A. and Weigt, M. (2000). On the properties of small-world networks, *Eur. Phys. J. B*, 13, pp.547-560.

Barabasi, A., Jeong, H., Neda, Z., Ravasz, E.. Schubert A., and Vicsek T. (2002). Evolution of the social network of scientific collaborations, *Physica A: Statistical Mechanics and its Applications*, 311, no. 3-4, pp. 590–614.

Batagelj, V. and Mrvar, A. (2000). Some analyses of the erdös collaboration graph, *Social Networks*, 22, pp.173-186.

Batty, C. D. (1966) *An introduction to the Dewey decimal classification. Melbourne*, Cheshire.

Beckmann, N., Kriegel, H. P., Schneider, R. and Seeger, B. (1990) The R* tree: an efficient and robust index method for points and rectangles. In *SIGMOD*, pp322-331.

Beckwith, R., Fellbaum, C., Gross, C. and Miller, G. (1991). WordNet: a lexical database organized on psycholinguistic principles, lexical acquisition: exploiting on-line resources to build a lexicon, U.Zernik, ed., *Lawrence Erlbaum*, pp.211-231.

Berners-Lee, T., Hendler, J. and Lassila, O. (2001). Semantic Web, *Scientific American*, 284, No.5, pp.34-43.

T. Berners-Lee, et al. (2006). A Framework for Web Science, Now Publishers Inc.

Berry, M.W. et al. (1999). Matrices, Vector Spaces, and Information Retrieval, *Society for Industrial and Applied Mathematics Review*, 41, pp.335-362.

Bianconi, G. and Barabási, A. L. (2001a). Bose-Einstein condensation in complex networks, *Phys. Rev. Lett.*, 86, pp.5632-5635.

Bianconi, G. and Barabási, A. L. (2001b). Competition and multiscaling in evolving networks, *Europhys. Lett.*, 54, pp.436-442.

Bocy, R. et al. (1975). Specifying queries as relational expressions, *Communications of the ACM*, 18, No.11, pp.621-628.

Booch, G., Rumbaugh, J. and Jacobson, I. (1999). The unified modeling language: user guide. Reading, Mass. *Addison-Wesley*.

Bornholdt, S. and Ebel, H. (2001). World Wide Web scaling exponent from Simon's 1955 model, *Phys. Rev. E*, 64, 035104(R).

Bonacich, P. (1972). Factoring and weighting approaches to status scores and clique identification, *Journal of Mathematical Sociology*, 2, no. 1, pp. 113–120.

Bradshaw, J. M., Greaves, M. and Holmback, H. (1999). Agents for the masses? *IEEE Intelligent Systems*, 14, No.2, Mar./Apr., pp.53-63.

Broder, A. et al. (2000). Graph structure of the Web, *Proc. of the 9th WWW Conf.*, Amsterdam, May, pp.309-320.

Broekstra, J. et al, (2001). Enabling knowledge representation on the Web by extending RDF schema, *Proc. 10th International WWW Conf.*, Hong Kong, pp.467-478, available at http://www.cs.vu.nl/~frankh/abstracts/www01.html.

Bush, V. (1945). As we may think. *The Atlantic Monthly*, 176, No.1, pp.101-108.

Callaway, D. S. et al. (2001). Are randomly grown graphs really random? *Phys. Rev. E.*, 64, cond-mat/0104546.

Cascia, M. L., Sethi, S. and Sclaroff, S. (1998). Combining textual and visual cues for content-based image retrieval on the worldwide web. *Proc. IEEE Workshop on Content-based Access of Image and Video Libraries*, pp.24-28.

Cass, S. (2004). A fountain of knowledge. *IEEE Spectrum*, 41, No.1, pp.60-67.

Chakrabarti, S., Dom, B., Gibson, D., Kleinberg, J., Raghavan, P. and Rajagopalan, S. (1998). Automatic resource list compilation by analyzing hyperlink structure and associated text, *Proc. 7th International WWW Conf.*, pp.65-74.

Chen, P. (1976) The Entity-Relationship Model: Toward a Unified View of Data, *ACM Transactions on Database Systems*, 1, No.1, pp.9-36.

Chen, Q. et al. (2002). The origin of power laws in Internet topologies revisited, *Proc. of the IEEE INFOCOM Conf.*, pp.608-617.

Clauset, A., Newman, M., and Moore, C. (2004). Finding community structure in very large networks, *Physical Review E*, 70, no. 6, p. 66111.

Codd, E. F. (1970). A relational model of data for large shared data banks, *Communications of the ACM*, 13, No.6, pp.377-387.

Cohen, E., Fiat, A. and Kaplan, H. (2007). Associative search in peer to peer networks: Harnessing latent semantics, *Computer Networks*, 51, no. 8, pp. 1861-1881.

Colizza, V., Flammini, A., Serrano, M., and Vespignani, A. (2006). Detecting rich-club ordering in complex networks, *Arxiv preprint physics/0602134*.

Cortes, C. and Vapnik, V. (1995), Support-Vector Networks. *Machine Learning*, 20(3)(1995) 273-297.

Crespo, A. and Garcia-Molina, H. (2002). Routing Indices for Peer-to-peer Systems, *Proceedings of the 22nd International Conference on Distributed Computing Systems*, pp. 23-32.

Dachselt, R. Frisch, M. Weiland, and M. (2007) FacetZoom: a continuous multi-scale widget for navigating hierarchical metadata, CHI08, 2008, pp.1353-1356.

Dakka, W., Ipeirotis, P.G., and Wood, K.R. (2007) Faceted Browsing over Large Databases of Text-Annotated Objects, ICDE 2007, pp.1489 – 1490.

Davenport, T. H., Jarvenpaa, S. L. and Beer, M. C. (1996). Improving knowledge work process, *Sloan Management Review*, 34, No.4, pp. 53-65.

Davis, G. F. and Greve, H. R. (1997). Corporate élite networks and governance changes in the 1980s, *Am. J. Sociol.*, 103, pp.1-37.

Dean, J. and Henzinger, M. R. (1999). Finding related pages in the World Wide Web, *Proc. 8th International WWW Conf.*, pp.1467-1476.

Decker, S., et al. (2000). The Semantic Web: the roles of XML and RDF, *IEEE Internet Computing*, 4, No. 5, pp.63-74.

Desouza, K. C. (2003). Facilitating tacit knowledge exchange, *Communications of the ACM*, 46, No.6, pp.85-88.

Dieng, R. (2000). Knowledge management and the Internet, *IEEE Intelligent Systems*, 15, No.3, May/June, pp.14-17.

Dorogovtsev, S. N. and Mendes, J. F. F. (2000a). Scaling behavior of developing and decaying networks, *Europhys. Lett.*, 52, pp.33-39.

Dorogovtsev, S. N. and Mendes, J. F. F. (2000b). Evolution of reference networks with aging, *Phys. Rev. E.*, 62, pp.1842-1845.

Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. (2000c). Structure of growing networks with preferential linking, *Phys. Rev. Lett.*, 85, pp.4633-4636.

Dorogovtsev, S. N., Mendes, J. F. F. and Samukhin, A. N. (2000d). WWW and Internet models from 1955 till our days and the "popularity is attractive" principle, *Condensed Matter Archive*, cond-mat/0009090.

Dorogovtsev, S. N. and Mendes, J. F. F. (2001). Language as an evolving word Web, *Proc. Royal Soc. London B*, 268, No.2603, cond-mat/0105093.

Drucker, P. F. (1998). Harvard business review on knowledge management, Boston, *Harvard Business School Press*.

Faloutsos, M., Faloutsos, P. and Faloutsos, C. (1999). On power law relationships of the Internet topology, *Comput. Commun. Rev.*, 29, pp.251-262.

Fensel, D., et al. (2001). OIL: an ontology infrastructure for the Semantic Web, *IEEE Intelligent Systems*, 16, No.2, Mar./Apr., pp.38-45.

Ferrer, R., Janssen, C. and Sole, R. V. (2001b). The topology of technology graphs: small-world patterns in electronic circuits, *Phys. Rev.E.*, 64, No.32767.

Ferrer, R. and Sole, R. V. (2001a). The small-world of human language, *Working Papers of Santa Fe Institute*, http://www.santafe.edu/sfi/publications/Abstracts/01-03-004abs.htm.

Franz, T. et al. (2009) TripleRank: Ranking Semantic Web Data by Tensor Decomposition. In *Proceedings of the International Semantic Web Conference*, 2009, pp.213-228.

Freeman, L. (1979), Centrality in social networks: Conceptual clarification, *Social Networks*, 1, no. 3, pp. 215–239.

Freeman, L. (1977). A Set of Measures of Centrality Based on Betweenness, *Sociometry*, 40, no. 1, pp. 35–41.

Fikes, R. and Farquhar, A. (1999). Distributed repositories of highly expressive reusable ontologies, *IEEE Intelligent Systems*, 14, No. 2, Mar/April, pp.73-79.

Fortunato, S., Latora, V., and Marchiori, M. (2004). Method to find community structures based on information centrality, *Physica Review E*, 70, no. 5, p. 56104.

Foster, I. (2000). Internet computing and the emerging Grid, 408, No.6815, *Nature*, http://www.nature.com/nature/webmatters/grid/grid.html.

Foster, I., Kesselman, C., Nick, J. M. and Tuecke, S. (2002). Grid services for distributed system integration, *Computer*, 35, pp.37-46.

Ganesan, P. Bawa, M. and Garcia-Molina, H. (2004) Online Balancing of Range-Partitioned Data with Applications to Peer-to-Peer Systems. *VLDB '04*. pp.444-455.

Gevers, T. and Smeulders, A. W. M. (1999). The PicToSeek WWW image search system, *Proc. IEEE International Conf. on Multimedia Computing and Systems*, pp.264-269.

Girvan, M. and Newman, M. (2002). Community structure in social and biological networks, *Proceedings of the National Academy of Sciences*, 99, no. 12, p. 7821.

Goldberg, A. (1984). SMALLTALK-80: the interactive programming environment, Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.

Gong, L. (2001). JXTA: a network programming environment, *IEEE Internet Computing*, 5, no.3, pp.88-95.

Gorlitz, O., Sizov, S. and Staab, S. (2008). PINTS: Peer-to-Peer Infrastructure for Tagging Systems, *The 7th International Workshop on Peer-to-Peer Systems*.

Govindan, R. and Tangmunarunkit, H. (2000). Heuristics for Internet map discovery, *Proc. of the IEEE INFOCOM Conf*., Tel Aviv, Israel, March, pp.1371-1380.

Gray, J. (2003). What next? A dozen information-technology research goals, *Journal of the ACM*, 50, No.1, pp.41-57.

Green, S. J. (1999). Building hypertext links by computing semantic similarity, *IEEE Trans. On Knowledge and Data Engineering*, 11, No.5, pp.713-730.

Grosz, B. J. and Kraus, S. (1996). Collaborative plans for complex group action, *Artificial Intelligence*, 86, pp.269-357.

Gupta, A. and Jain, R. (1997). Visual information retrieval, *Communications of the ACM*, 40, No.5, pp.71-79.

Gyssens, M .et al., (1994). A graph-oriented object database model, *IEEE Transactions on Knowledge and Data Engineering*, 6, No.4, pp.572-586.

Han, J. and Kambr, M. (2000). Data mining: concepts and techniques, *Morgan Kaufmann Publishers*.

Haridasan, M. and Renesse, R. (2008). Gossip-based Distribution Estimation in Peer-to-Peer Networks, *the 7th International Workshop on Peer-to-Peer Systems*.

Harmandas, V., Sanderson, M. and Dunlop, M. D. (1997). Image retrieval by hypertext links. *Proc. 20th Annual International ACM SIGIR Conf. on Research and Development in Information Retrieval*, pp.296-303.

Hammer, M. and McLeod, D. (1978) The Semantic Data Model: A Modeling Mechanism for Database Applications, *SIGMOD*78, pp.26-36.

Havil, J. and Dyson, F. (2003). The Harmonic Series, Chapter.2 in Gamma: Exploring Euler's Constant. *Princeton, NJ: Princeton University Press.*

Heflin, J. and Hendler, J. (2001). A portrait of the Semantic Web in action, *IEEE Intelligent Systems*, 16, No.2, Mar./Apr., pp.54-59.

Hendler, J and McGuinness, D. (2000). The DARPA agent markup language, *IEEE Intelligent Systems*, 15, No.6, Nov./Dec., pp.72-73.

Hendler, J. (2001). Agents and the Semantic Web, *IEEE Intelligent Systems*, 16, No.2, Mar./Apr., pp.30-37.

Henzinger, M. R. (2001). Hyperlink analysis for the Web, *IEEE Internet Computing*, 5, No.1, Jan./Feb., pp.45-50.

Heylighen, F. (1992). A Cognitive-systemic reconstruction of maslow's theory of self-actualization, *Behavioral Science*, 37, pp.39-58.

Huberman, B. A. and Adamic, L. A. (1999). Growth dynamics of the World-Wide Web, *Nature*, 401, 131p.

Huynh D.F. and Karger D.R. (2009). Parallax and companion: Set-based browsing for the data web, WWW09, pp.145-156.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabási, A. L. (2000). The large-scale organization of metabolic networks, *Nature*, 407, pp.651-654.

Jeong, H. Neda, Z. and Barabási, A. L. (2003). Measuring preferential attachment for evolving networks, *Europhysics letters*, 61, pp.567-572.

Kanth, K., Agrawal, D. and Singh, A. (1998). Dimensionality reduction for similarity searching in dynamic databases, *Proc. ACM SIGMOD International Conf. on Management of Data*, 75, pp.166-176.

Kaashoek, M. F. and Karger and Koorde, D. R. (2003). A Simple Degree-Optimal Distributed Hash Table. *Proceeding of International Workshop on Peer-to-Peer Systems (IPTPS)*. (2003)98-107.

Kee, K.-P., Swearingen, K., Li, K. and Hearst, M. (2003). Faceted metadata for image search and browsing, CHI03' *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2003, April 5-10, 2003, Florida, USA.

Kermarrec, A.M., Massoulié, L. and Ganesh, A.J. (2003). Probabilistic Reliable Dissemination in Large-scale Systems, *IEEE Transactions on Parallel and Distributed Systems*, 14, no. 3, pp. 248-258.

Klein, M. (2001). XML, RDF, and relatives, *IEEE Internet Computing*, 16, No.2, Mar./Apr., pp.26-28.

Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM*, 46, No.5, pp.604-632.

Kleinberg, J and Lawrence, S. (2001). The structure of the Web, *Science*, 294, No.5548, pp.1849-1850.

Kleinberg, J. (2001) Small-World Phenomena and the Dynamics of Information, *Proceeding of Advances in Neural Information Processing Systems* (NIPS) 14, pp.431-415.

Krapivsky, P. L. Redner, S. and Leyvraz, F. (2000). Connectivity of growing random networks, *Phys. Rev. Lett.*, l.85, pp.4629-4632.

Krapivsky, P. L. and Redner, S. (2001a). Organization of growing random networks, *Phys. Rev. E.*, 63, cond-mat/0011094.

Krapivsky, P. L., Rodgers, G. J. and Redner, S. (2001b). Degree distributions of growing networks, *Phys. Rev. Lett.*, 86, pp.5401-5404.

Krapivsky, P. L. and Redner, S. (2002). A statistical physics perspective on Web growth, *Computer Networks*, 39, pp.261-276.

Kumar, R. et al. (1999). Extracting large-scale knowledge bases from the Web, *Proc. 25th VLDB Conf.*, Edinburgh, Scotland, September, pp.639-650.

Larry, P., Sergey, B., Motwani, R. *et al.* (1998). The PageRank citation ranking: Bringing order to the web, *Online: http://citeseer. nj.Nec com/page98pagerank. html* [*04.06. 2003*].

Latora, V. and Marchiori, M. (2004). A measure of centrality based on the network efficiency, *Arxiv preprint cond-mat/0402050*.

Lawrence, S. and Giles, C. L. (1998). Searching the World Wide Web, *Science*, 280, No.5360, pp.98-100.

Lawrence, S. and Giles, C. L. (1999). Accessibility of information on the Web, *Nature*, 400, No.8, pp.107-109.

Lazer, D. et al. (2009), Computational social science, *Science*, 323, no.5915, pp.721-723.

Lee, S., Yook, S., and Kim, Y. (2009). Centrality measure of complex networks using biased random walks, *The European Physical Journal B*, 68, no. 2, pp. 277–281.

Leland, W. E. et al. (1994). On the self-similar nature of Ethernet traffic, *IEEE/ACM Transactions on Networking*, 2, no.1, pp.1-15.

Lempel, R. and Moran, S. (2000). The stochastic approach for link-structure analysis (SALSA) and the TKC effect. *Proc. 9th International WWW Conf.*, pp.387-401.

Lempel, R. and Soffer, A. (2001). PicASHOW: pictorial authority search by hyperlinks on the Web, *Proc. 10th International WWW Conf.*, pp.438-448.

Levene, M., Fenner, T., Loizou, G. and Wheeldon, R. (2002). A stochastic model for the evolution of the Web, *Computer Networks*, 39, pp.277-287.

Leymann, F. and Roller, D. (1997). Workflow-based applications, *IBM Systems Journa*l, 36, No.1, pp.102-122.

Licklider, J.C.R. (1960). Man-Computer Symbiosis, Transactions on Human Factors in Electronics, vol. HFE-1, pp.4–11, March.

Lieberman, H. (2008). A Creative Programming Environment, in HCI Remixed, T.Erickson and D.W.McDonald (Ed).

B. Libet, Mind Time, Harvard University Press, 2005.

Liu, Z. Q. and Satur, R. (1999). Contextual fuzzy cognitive maps for decision support in geographic information systems, *IEEE Transactions on Fuzzy Systems*, 7, No.10, pp.495-502.

Mack, R., Ravin, Y. and Byrd, R. J. (2001). Knowledge portals and the emerging knowledge workplace, *IBM Systems Journal*, 40, No. 4, pp. 925-955.

Maedche, A. and Staab, S. (2001). Ontology learning for the Semantic Web, *IEEE Intelligent Systems*, 16, No.2, Mar./Apr., pp.72-79.

Malkhi, D., Naor, M. and Ratajczak, D. (2002). Viceroy: A Scalable and Dynamic Emulation of the Butterfly. *Proceeding of the 21st ACM SIGACT-SIGOPS Symposium on Principles of Distributed Computing* (PODC '02), pp.183-192.

Manku, G. Bawa, M. and Raghavan, P. Symphony, (2003), Distributed Hashing in a Small World. *Proceeding of the 4th USENIX Symposium on Internet Technologies and Systems*, pp.10-22.

Marchionini, G., Dwiggins, S., Katz, A. and Lin, X. (1990). Information roles of domain and search expertise, *Library and Information Science Research*, 15, no.1, pp.391-407.

Martin, P. and Eklund, P. W. (2000). Knowledge retrieval and the World Wide Web, *IEEE Intelligent Systems*, 15, no.3, May/June, pp. 18-25.

Martinez, N. D. (1991). Artifacts or attributes? Effects of resolution on the Little Rock Lake food web, *Ecological Monographs*, 61, pp.367-392.

Maybury, M. (2001). Collaborative virtual environments for analysis and decision support, *Communication of the ACM*, 14, no.12, pp. 51-54.

McCreight, E. M. (1976) A Space-Economical Suffix Tree Construction Algorithm. *Journal of ACM*, 23, no.2, pp. 262-272.

McHraith, S. A., Son, T. C. and Zeng, H. (2001). Semantic Web services, *IEEE Intelligent Systems*, 16, No.2, Mar./Apr., pp. 46-53.

Miao, Y., et al. (2001). Dynamic cognitive network, *IEEE Transactions on Fuzzy System*, 9, no.5, pp. 760-770.

Milgram, S. (1967) The Small World Problem. *Psychology Today*, 1, no.61.

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network Motifs: Simple Building Blocks of Complex Networks, pp. 824–827.

Mitri, M. (2003). Applying tacit knowledge management techniques for performance assessment, *Computers & Education*, 4, pp.173-189.

Mok, W.Y. (2002). A comparative study of various nested normal forms, *IEEE Trans. on Knowledge and Data Engineering*, 14, no.2, pp. 369-385.

Nasukawa, T. and Nagano, T. (2001). Text analysis and knowledge mining system, *IBM Systems Journa*l, 40, no. 4, pp. 967-984.

Nejdl, W. et al. (2002). Edutella: a P2P Networking Infrastructure Based on RDF, *Proceedings of the 11th International Conference on World Wide Web*, pp. 604-615.

Newman, M. E. J. et al.(2001a). Random graphs with arbitrary degree distribution and their applications", *Phys. Rev. E*, 64, no.026118, cond-mat/0007235.

Newman, M. E. J. (2001b). The structure of scientific collaboration networks, *Proc. Nat. Acad. Sci. U.S.A.*, 98, pp.404-409.

Newman, M. E. J. (2001c). Scientific collaboration networks: I. Network construction and fundamental results, *Phys. Rev. E*, 64, No.016131.

Newman, M. E. J. (2001d). Scientific collaboration networks: II. Shortest paths, weighted networks, and centrality, *Phys. Rev. E*, 64, No.016132.

Newman, M. E. J. (2001e). Who is the best connected scientist? A study of scientific coauthorship networks, *Phys. Rev. E,* 64, No.016131.

Newman, M.(2004). Analysis of weighted networks, *Physical Review E*, 70, no. 5.

Ng, W.S. et al. (2003). PeerDB: A P2P-Based system for distributed data sharing, *Proc. of International Conf. on Data Engineering*, ICDE2003, Bangalore, India.

Nieminen, J. (1974). On the centrality in a graph, *Scandinavian Journal of Psychology*, 15, no. 1, pp. 332–336, 1974.

Nissen, M. E. (2002). An extended model for knowledge-flow dynamics, *Communications of the Associations for Information Systems*, 8, pp. 251-266.

Nonaka, I. (1994). A dynamic theory of organizational knowledge creation, *Organization Science*, 5, No.1, pp. 14-37.

Nowak, M. A. Komarova, N. L., and Niyogi, P., Computational and Evolutionary Aspects of Language, *Nature*, 417(2002)611-617.

Oida, K. (2002). The birth and death process of hypercycle spirals, in: R.K. Standish, M.A. Bedau, H.A. Abbass, edd., Artificial Life VIII, *MIT Press*.

Oren, E., Delbru R. and Decker, S. (2006). Extending Faceted Navigation for RDF Data, ISWC2006, pp.559-572.

O'Leary, D. E. (1998). Enterprise knowledge management, *Computer*, 31, no.3, pp.54-61.

O'Neil, P. O'Neil, E. Pal, S. Cseri, E. and Schaller, G. (2004), ORDPATHs: Insert-Friendly XML Node Labels, *SIGMOD*, pp. 903-908.

Ozsoyoglu, Z. M. and Yuan, L. Y. (1987). A new normal form for nested relations, *ACM Trans. Database Systems*, 12, no.1, pp.111-136.

Pansiot, J. J. and Grad, D. (1998). On routes and multicast trees in the Internet, *Computer Communications Review*, 28, pp.41-50.

Parise, S. and Henderson, J. C. (2001). Knowledge resource exchange in strategic alliances, *IBM Systems Journal*, 40, no. 4, pp.908-924.

Peltz, C. (2003). Web Services Orchestration and Choreography, *Computer*, 36, no.10, pp. 46-52

Perra, N. and Fortunato, S. (2008). Spectral centrality measures in complex networks, *Physical Review E*, vol. 78, no. 3, pp. 036107.

Pittel, B. On Spreading a Rumor, (1987). *SIAM Journal of Applied Mathematics*, 47, no. 1, pp. 213-223.

Raman, V. and Hellerstein (2001), J. M. Potter's wheel: An interactive data cleaning system, *VLDB2001*, Roma, Italy.

Ratnasamy, S. et al. (2001). A Scalable Content-addressable Network, *ACM SIGCOMM Symposium on Communication, Architecture, and Protocol*, pp.161-172.

Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution, *European Physics Journal B*, 4, pp.131-134.

Renesse, R. V., Minsky, Y. and Hayden, M. (1998). A Gossip-style Failure Detection Service, *Middleware98: IFIP International Conference, Distributed Systems and Platforms and Open Distributed Processing*, pp. 55-70.

Ritter, F. E., Baxter, G. E., Jones, G and Young, R.M. (2000). Supporting cognitive models as users, *ACM Trans. on Computer-Human Interaction*, 7, no.2, pp.141-173.

Robertson, S., Walker, S., and Beaulieu, M. (1998). Okapi at TREC-7: Automatic ad hoc, filtering, VLC and interactive. *Proceedings of the* 7^{th} *text retrieval conference (TREC-7)*, NIST special publication, pages 500–242.

Rowstron, A. and Druschel, P. (2001). Pastry: Scalable, Distributed Object Location and Routing for Large-scale Peer-to-Peer Systems, *International Conference on Distributed Systems Platforms*, pp. 329-350.

Sabidussi, G.(1966). The centrality index of a graph, *Psychometrika*, 31, no. 4, pp. 581–603, 1966.

Simon, H. A. (1955). On a class of skew distribution functions, *Biometrika*, 42, pp.425-440.

Simon, H. A. and Van Wormer, T. (1963). Some Monte Carlo estimates of the Yule distribution, *Behavioral Sci.*, 8, pp.203-210.

Srihari, R. K., Zhang, Z. and Rao, A. (2000). Intelligent indexing and semantic retrieval of multimodel documents, *Information Retrieval*, 2, no.2, pp. 245-275.

Stoica, I. et al., (2003). Chord: A scalable Peer-to-Peer lookup protocol for Internet applications, *IEEE/ACM Transactions on Networking*, 11, pp. 17–32.

Sun, X. (2010), OSLN: An Object-Oriented Semantic Link Network language for complex object description and operation. *Future Generation Computer Systems*, 26, no.3, pp.389-399

Tari, Z., Stokes, J. and Spaccapietra, S. (1997). Object normal forms and dependency constraints for object-oriented schemata, *ACM Trans. Database Systems*, 22, no.4, pp.513-569.

Tatarinov, I. et al. (2003). The Piazza Peer Data Management Project, *SIGMOD Record*, 32, no. 3, pp. 47-52.

Tempich, C., Staab, S. and Wranik. A. (2004). REMINDIN': Semantic Query Routing in Peer-to-Peer Networks Based on Social Metaphors, *Proceedings of the 13th international conference on World Wide Web*, pp. 640-649.

Thistlewaite, P. (1997). Automatic construction and management of large open Webs, *Information Processing and Management*, 33, no.2, pp. 145-159.

Thomas, J. C., Kellogg, W. A. and Erickson, T. (2001). The knowledge management puzzle: Human and social factors in knowledge management, *IBM Systems Journal*, 40, no. 4, pp.863-884.

Tudhope, D. and Taylor, C. (1997). Navigation via similarity: automatic linking based on semantic closeness, *Information Processing and Management*, 33, no.2, pp.233-242.

Turing, A.M. (1950). Computing machinery and intelligence, *Mind*, vol.59, no.236, pp.433-460.

Ullman, J. D. (1988). Principles of database and knowledge-base systems, *Computer Science Press, Inc.*, Valverde, S., Cancho, R. F. and Sole, R. V. (2002). Scale-free networks from optimal design, *Europhysics Letters*, 60, pp.512-517.

W. Vogels, R. V. Renesse, and K. Birman. (2003) The Power of Epidemics: Robust Communication for Large-Scale Distributed Systems. *ACM SIGCOMM Computer Communication Review*. 33, no.1, pp.131-135.

Warshall, S.(1962). A theorem on boolean matrices, *J. ACM*, 9, no. 1, pp. 11–12.

Wasserman, S. and Faust, K. (1994). *Social Network Analysis: Methods and Applications*. Cambridge University Press.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of “Small-world” networks, *Nature*, 393, pp.440-442.

Watts, D. J. (1999). Networks, dynamics, and the small-world phenomenon, *Am. J. Sociol.*, 105, pp.493-592.

Wilks, Y. Is There Progress on Talking Sensibly to Machines? *Science*, 5852(318)927-928.

Yang, B. and Garcia-Molina, H. (2002). Efficient Search in Peer-to-peer Networks, *International Conference on Distributed Computing Systems*, pp.5-14.

Yook, S. H. et al. (2002). Modeling the Internet’s large-scale topology, *Proc. the National Academy of Sciences*, pp. 13382-13386, cond-mat/0107417.

Zeng Q., Jiang X. and Zhuge, H. (2012). Adding logical operators to tree pattern queries on graph-structured data, *VLDB* 2012.

Zhao, B. Y., Kubiatowicz, J.D. and Joseph, A.D. (2001). Tapestry: An Infrastructure for Fault-resilient Wide-area Location and Routing, Technical Report UCB/CSD-01-1141, U.C.Berkeley.

Zhu, Y. W., Wang, H. and Hu, Y. M. (2003). Integrating Semantic-Based Access Mechanisms with Peer-to-Peer File System, *Proceedings of the 3rd International Conference on Peer-to-Peer Computing*, pp. 118-125.

Zhuge, H. (1998). Inheritance rules for flexible model retrieval, *Decision Support Systems*, 22, no.4, pp. 383-394.

Zhuge, H. (2000a). Conflict decision training through multi-space co-operation, *Decision Support Systems*, 29, no. 5, pp.111-123.

Zhuge, H., Cheung, T. Y. and Pung, H. K. (2000b). A timed workflow process model, *Journal of Systems and Software*, 55, no.3, pp.231-243.

Zhuge, H. (2002a). A Knowledge Grid model and platform for global knowledge sharing, *Expert Systems with Applications*, 22, no.4, pp.313-320.

Zhuge, H. (2002b). A knowledge flow model for peer-to-peer team knowledge sharing and management, *Expert Systems with Applications*, 23, no.1, pp.23-30.

Zhuge, H. (2002c). Clustering soft-device in the Semantic Grid, *IEEE Computing in Science and Engineering*, 4, no.6, Nov./Dec., pp.60-62.

Zhuge, H. (2003b). China e-science Knowledge Grid Environment, *IEEE Intelligent Systems*, 19, No.1, Sep./Oct., pp. 13-17.

Zhuge, H. (2003c). Workflow-based cognitive flow management for distributed team cooperation, *Information and Management*, 40, no.5, pp.419-429.

Zhuge, H. (2003d). Active e-Document Framework ADF: model and tool, *Information and Management*, 41, no.1, pp. 87-97.

Zhuge, H. and Shi, X. (2003e). Fighting epistemology in knowledge and information age, *Computer*, 36, no.10, pp.114-116.

Zhuge, H. (2004a). Resource Space Model, its design method and applications, *Journal of Systems and Software*, 72, no.1, pp. 71-81.

Zhuge, H. (2004b). Semantics, resources and Grid, *Future Generation Computer Systems*, Special issue on *Semantic Grid and Knowledge Grid: the Next-generation Web*, editorial, 20, no.1, pp.1-5.

Zhuge, H. (2004c). Fuzzy resource space model and platform, *Journal of Systems and Software*, 73, no.3, pp.389-396.

Zhuge, H. (2004d). Resource Space Grid: model, method and platform, *Concurrency and Computation: Practice and Experience*, 16, no.13.

Zhuge, H. (2004e). Retrieve images by understanding semantic links and clustering image fragments, *Journal of Systems and Software*, 73, no.3, pp.455-466.

Zhuge, H. (2004f). Toward the eco-grid: a harmoniously evolved interconnection environment, *Communications of the ACM*, 47, no.9, pp.78-83.

Zhuge, H. (2004g). Semantics, Resource and Grid, *Future Generation Computer Systems*, 20, no.1, pp.1-5.

Zhuge, H. (2005a). The Future Interconnection Environment, *IEEE Computer*, 38, no.4, pp.27-33.

Zhuge, H. (2005b). Exploring an Epidemic in an E-Science Environment, *Communications of the ACM*, 48, no.9, pp.109-114.

Zhuge, H. et al. (2005c), Query Routing in a Peer-to-Peer Semantic Link Network *Computational Intelligence*, 21, no. 2, pp. 197–216.

Zhuge, H. (2005d). Semantic Grid: Scientific Issues, Infrastructure, and Methodology. *Communications of the ACM*, 48, no.4, pp.117-119.

Zhuge, H., Sun, X., Liu, J., Yao, E. and Chen, X., (2005). A Scalable P2P Platform for the Knowledge Grid, *IEEE Transactions on Knowledge and Data Engineering*, 17, no.12, pp.1721-1736.

Zhuge, H. (2006). Discovery of Knowledge Flow in Science, *Communications of the ACM*, 49, no. 5, pp. 101-107.

Zhuge, H. (2007). Autonomous Semantic Link Networking Model for the Knowledge Grid, *Concurrency and Computation: Practice and Experience*, no.19, pp.1065–1085.

Zhuge H. and Li, X. (2007), Peer-to-Peer in Metric Space and Semantic Space, *IEEE Transactions on Knowledge and Data Engineering*, 19, no. 6, pp. 759–771.

Zhuge H. and Feng, L. (2008), Distributed Suffix Tree Overlay for Peer-to-Peer Search, *IEEE Transactions on Knowledge and Data Engineering*, 20, no.2, pp.276–285.

Zhuge, H. (2008). The Knowledge Grid Environment, *IEEE Intelligent Systems*, 23, no. 6, pp.63-71.

Zhuge, H. and Sun, X. (2008). A Virtual Ring Method for Building Small-World Structured P2P Overlays, *IEEE Transactions on Knowledge and Data Engineering*, 20, no.12, pp.1712-1725.

Zhuge, H. Xing, Y.and Shi, P. (2008). Resource Space Model, OWL and Database: Mapping and Integration, *ACM Transactions on Internet Technology*, 8, no.4.

Zhuge, H., Chen, X., Sun, X. and Yao, E. (2008), HRing: A Structured P2P Overlay Based on Harmonic Series, *IEEE Transactions on Parallel and Distributed Systems*, 19, no.2, pp.145-158.

Zhuge, H. and Feng, L. (2008). Distributed Suffix Tree Overlay for Peer-to-Peer Search, *IEEE Transactions on Knowledge and Data Engineering*, 20, no.2, pp.276-285.

Zhuge, H. (2009), Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning, *IEEE Transactions on Knowledge and Data Engineering*, 21, no.6, pp.785-799.

Zhuge, H. and Sun, Y. (2010) The schema theory for semantic link network. *Future Generation Computer Systems*, 26, no.3, pp. 408-420.

Zhuge, H. (2010) Interactive Semantics, *Artificial Intelligence*, 174, no.2, pp.190-204.

Zhuge, H. (2011) Semantic linking through spaces for cyber-physical-socio intelligence: A methodology, *Artificial Intelligence*, 175, pp.988-1019.