Klaus P. Jantke
Ayumi Shinohara (Eds.)

LNAI 2226

Discovery Science

**4th International Conference, DS 2001
Washington, DC, USA, November 2001
Proceedings**

Springer

Klaus P. Jantke Ayumi Shinohara (Eds.)

Discovery Science

4th International Conference, DS 2001
Washington, DC, USA, November 25-28, 2001
Proceedings

Springer

Series Editors

Jaime G. Carbonell, Carnegie Mellon University, Pittsburgh, PA, USA
Jörg Siekmann, University of Saarland, Saarbrücken, Germany

Volume Editors

Klaus P. Jantke
DFKI GmbH Saarbrücken
66123 Saarbrücken, Germany
E-mail: jantke@dfki.de

Ayumi Shinohara
Kyushu University, Department of Informatics
6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-8581, Japan
E-mail: ayumi@i.kyushu-u.ac.jp

Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek - CIP-Einheitsaufnahme

Discovery science : 4th international conference ; proceedings / DS 2001, Washington, DC, USA, November 25 - 28, 2001. Klaus P. Jantke ; Ayumi Shinohara (ed.). - Berlin ; Heidelberg ; New York ; Barcelona ; Hong Kong ; London ; Milan ; Paris ; Tokyo : Springer, 2001
(Lecture notes in computer science ; Vol. 2226 : Lecture notes in artificial intelligence)
ISBN 3-540-42956-5

CR Subject Classification (1998): I.2, H.2.8, H.3, J.1, J.2

ISBN 3-540-42956-5 Springer-Verlag Berlin Heidelberg New York

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

Springer-Verlag Berlin Heidelberg New York
a member of BertelsmannSpringer Science+Business Media GmbH

http://www.springer.de

© Springer-Verlag Berlin Heidelberg 2001
Printed in Germany

Typesetting: Camera-ready by author
Printed on acid-free paper SPIN: 10840973 06/3142 5 4 3 2 1 0

Preface

These are the conference proceedings of the 4th International Conference on Discovery Science (DS 2001). Although discovery is naturally ubiquitous in science, and scientific discovery itself has been subject to scientific investigation for centuries, the term Discovery Science is comparably new. It came up in connection with the Japanese Discovery Science project (cf. Arikawa's invited lecture on *The Discovery Science Project in Japan* in the present volume) some time during the last few years.

Setsuo Arikawa is the father in spirit of the Discovery Science conference series. He led the above mentioned project, and he is currently serving as the chairman of the international steering committee for the Discovery Science conference series. The other members of this board are currently (in alphabetical order) Klaus P. Jantke, Masahiko Sato, Ayumi Shinohara, Carl H. Smith, and Thomas Zeugmann.

Colleagues and friends from all over the world took the opportunity of meeting for this conference to celebrate Arikawa's 60th birthday and to pay tribute to his manifold contributions to science, in general, and to Learning Theory and Discovery Science, in particular.

Algorithmic Learning Theory (ALT, for short) is another conference series initiated by Setsuo Arikawa in Japan in 1990. In 1994, it amalgamated with the conference series on Analogical and Inductive Inference (AII), when ALT was held outside of Japan for the first time.

This year, ALT 2001 and DS 2001 were co-located in Washington D.C., held in parallel and sharing five invited talks and all social events. The proceedings of ALT 2001 are published as a twin volume of the present one as LNAI 2225.

The present volume is organized in three parts. The first part contains the five invited lectures of ALT 2001 and DS 2001 exactly in the order in which they appeared in the conferences' common advance program. The invited speakers are Setsuo Arikawa, Lindley Darden, Dana Angluin, Ben Shneiderman, and Paul R. Cohen. Because their talks were invited to both conferences, there had to be found a modus vivendi for publication. This volume contains the full versions of Lindley Darden's and Ben Shneiderman's paper as well as abstracts of the others.

The second part contains the accepted 30 regular papers of the DS 2001 conference. Last but not least, there is a third part with written versions of the posters accepted for presentation during the conference. In a sense, DS 2001 posters are posters of ALT 2001 as well, because both events shared a conference venue including the exhibition area for the posters.

The combination of ALT 2001 and DS 2001 allowed for an especially comprehensive treatment of the issues ranging from rather theoretical investigations to applications and to both psychological and sociological topics. The organizers consider this an attractive approach to both communities.

Over the past dozen or so years, many enterprises have begun to routinely capture paramount volumes of data describing their operations, products, services, and customers. Simultaneously, scientists and engineers have been record-

ing experimental data of a continuously growing size covering experience in many fields. The finer the measurement granularity of the engineers' equipment and the more computer power available to support scientific experiments, the larger the amounts of data captured. These huge collections of bits and bytes constitute a new challenge to those who try to separate the wheat from the chaff.

Potentially, there is more fruitful knowledge hiding in huge amounts of data, but combinatorially, there is even more rubbish. It requires a new dimension of technological investment to extract useful information, and humans must attack these problems differently. Discovery Science deals with all aspects of promoting scientific discovery, and it is changing its character within a changing world.

New questions are being asked and leading to innovative concepts. Conceptualizations are setting the stage for asking new questions. Under these circumstances, new ways of looking at the problems might arise. More traditional disciplines are invoked and innovative ideas are made precise to get computers involved in knowledge discovery. Autonomously working machinery is necessary to deal with the flood of data, thus learning becomes a core technology of discovery science. When all said and done, humans and machines must learn together and support each other.

The field of Discovery Science is evolving and frequently changing its appearance. The Discovery Science conference series aims at reflecting this development, summarizing the state of affair and helping humans to navigate in such an exciting environment.

September 2001

Klaus P. Jantke
Ayumi Shinohara

试读结束：需要全本请在线购买： www.ertongbook.com

Organization

Conference Chair

Masahiko Sato	Kyoto Univ., Japan

Program Committee

Klaus P. Jantke (Co-chair)	DFKI GmbH, Germany
Ayumi Shinohara (Co-chair)	Kyushu Univ. Japan
Diane J. Cook	Univ. of Texas at Arlington, USA
Andreas Dengel	DFKI GmbH, Germany
Michael E. Gorman	Univ. of Virginia, USA
Achim Hoffmann	UNSW, Australia
John R. Josephson	Ohio State Univ., USA
Pat Langley	ISLE, USA
Hiroshi Motoda	Osaka Univ., Japan
Ryohei Nakano	Nagoya Inst. Techn., Japan
Yukio Ohsawa	Tsukuba Univ., Japan
Jorge C.G. Ramirez	Intelligent Tech. Corp., USA
Kazumi Saito	NTT, Japan
Einoshin Suzuki	Yokohama National Univ., Japan
Stefan Wrobel	Techn. Univ. Magdeburg, Germany
Thomas Zeugmann	Med. Univ. Lübeck, Germany

Local Arrangements

Carl Smith	Univ. of Maryland, USA

Subreferees

Hiroki Arimura
Markus Bläser
Eisaku Maeda
Tsutomu Hirao
Armin Hust
Naresh Iyer
Andreas Jakoby
Markus Junker
Stefan Klink
Maciej Liskiewicz
Shigeo Matsubara
Bodo Siebert
Marin Simina
Hirotoshi Taira
Paul Thagard
Naonori Ueda
Shravan Vasishth

Contents

Invited Papers

The Discovery Science Project in Japan 1
Setsuo Arikawa

Discovering Mechanisms: A Computational Philosophy of Science Perspective 3
Lindley Darden

Queries Revisited 16
Dana Angluin

Inventing Discovery Tools: Combining Information Visualization with Data Mining 17
Ben Shneiderman

Robot Baby 2001 29
Paul R. Cohen, Tim Oates, Niall Adams and Carole R. Beal

Regular Papers

VML: A View Modeling Language for Computational Knowledge Discovery 30
Hideo Bannai, Yoshinori Tamada, Osamu Maruyama and Satoru Miyano

Computational Discovery of Communicable Knowledge: Symposium Report 45
Sašo Džeroski and Pat Langley

Bounding Negative Information in Frequent Sets Algorithms 50
I. Fortes, J.L. Balcázar and R. Morales

Functional Trees 59
João Gama

Spherical Horses and Shared Toothbrushes: Lessons Learned from a Workshop on Scientific and Technological Thinking 74
Michael E. Gorman, Alexandra Kincannon and Matthew M. Mehalik

Clipping and Analyzing News Using Machine Learning Techniques 87
Hans Gründel, Tino Naphtali, Christian Wiech, Jan-Marian Gluba, Maiken Rohdenburg and Tobias Scheffer

Towards Discovery of Deep and Wide First-Order Structures: A Case Study in the Domain of Mutagenicity 100
Tamás Horváth and Stefan Wrobel

Eliminating Useless Parts in Semi-structured Documents Using Alternation Counts 113
Daisuke Ikeda, Yasuhiro Yamada and Sachio Hirokawa

Multicriterially Best Explanations 128
Naresh S. Iyer and John R. Josephson

Constructing Approximate Informative Basis of Association Rules 141
Kouta Kanda, Makoto Haraguchi and Yoshiaki Okubo

Passage-Based Document Retrieval as a Tool for Text Mining with User's Information Needs 155
Koichi Kise, Markus Junker, Andreas Dengel and Keinosuke Matsumoto

Automated Formulation of Reactions and Pathways in Nuclear Astrophysics: New Results 170
Sakir Kocabas

An Integrated Framework for Extended Discovery in Particle Physics 182
Sakir Kocabas and Pat Langley

Stimulating Discovery 196
Ronald N. Kostoff

Assisting Model-Discovery in Neuroendocrinology 214
Ashesh Mahidadia and Paul Compton

A General Theory of Deduction, Induction, and Learning 228
Eric Martin, Arun Sharma and Frank Stephan

Learning Conformation Rules 243
Osamu Maruyama, Takayoshi Shoudai, Emiko Furuichi, Satoru Kuhara and Satoru Miyano

Knowledge Navigation on Visualizing Complementary Documents 258
Naohiro Matsumura, Yukio Ohsawa and Mitsuru Ishizuka

KeyWorld: Extracting Keywords from a Document as a Small World 271
Yutaka Matsuo, Yukio Ohsawa and Mitsuru Ishizuka

A Method for Discovering Purified Web Communities 282
Tsuyoshi Murata

Divide and Conquer Machine Learning for a Genomics Analogy Problem .. 290
Ming Ouyang, John Case and Joan Burnside

Towards a Method of Searching a Diverse Theory Space for Scientific Discovery 304
Joseph Phillips

Efficient Local Search in Conceptual Clustering 323
Céline Robardet and Fabien Feschet

Computational Revision of Quantitative Scientific Models 336
Kazumi Saito, Pat Langley, Trond Grenager, Christopher Potter, Alicia Torregrosa and Steven A. Klooster

An Efficient Derivation for Elementary Formal Systems Based on Partial Unification 350
Noriko Sugimoto, Hiroki Ishizaka and Takeshi Shinohara

Worst-Case Analysis of Rule Discovery 365
Einoshin Suzuki

Mining Semi-structured Data by Path Expressions 378
Katsuaki Taniguchi, Hiroshi Sakamoto, Hiroki Arimura, Shinichi Shimozono and Setsuo Arikawa

Theory Revision in Equation Discovery 389
Ljupčo Todorovski and Sašo Džeroski

Simplified Training Algorithms for Hierarchical Hidden Markov Models ... 401
Nobuhisa Ueda and Taisuke Sato

Discovering Repetitive Expressions and Affinities from Anthologies of Classical Japanese Poems 416
Koichiro Yamamoto, Masayuki Takeda, Ayumi Shinohara, Tomoko Fukuda and Ichirō Nanri

Poster Papers

Web Site Rating and Improvement Based on Hyperlink Structure 429
Hironori Hiraishi, Hisayoshi Kato, Naonori Ohtsuka and Fumio Mizoguchi

A Practical Algorithm to Find the Best Episode Patterns 435
Masahiro Hirao, Shunsuke Inenaga, Ayumi Shinohara, Masayuki Takeda and Setsuo Arikawa

Interactive Exploration of Time Series Data 441
Harry Hochheiser and Ben Shneiderman

Clustering Rules Using Empirical Similarity of Support Sets 447
Shreevardhan Lele, Bruce Golden, Kimberly Ozga and Edward Wasil

Computational Lessons from a Cognitive Study of Invention 452
Marin Simina, Michael E. Gorman and Janet L. Kolodner

Component-Based Framework for Virtual Information Materialization..... 458
Yuzuru Tanaka and Tsuyoshi Sugibuchi

Dynamic Aggregation to Support Pattern Discovery: A Case Study with Web Logs 464
Lida Tang and Ben Shneiderman

Separation of Photoelectrons via Multivariate Maxwellian Mixture Model . 470
Genta Ueno, Nagatomo Nakamura and Tomoyuki Higuchi

Logic of Drug Discovery: A Descriptive Model of a Practice in Neuropharmacology 476
Alexander P.M. van den Bosch

SCOOP: A Record Extractor without Knowledge on Input..... 482
Yasuhiro Yamada, Daisuke Ikeda and Sachio Hirokawa

Meta-analysis of Mutagenes Discovery 488
Premysl Zak, Pavel Spacil and Jaroslava Halova

Author Index 493

The Discovery Science Project in Japan

Setsuo Arikawa

Department of Informatics, Kyushu University
Fukuoka 812-8581, Japan
arikawa@i.kyushu-u.ac.jp

Abstract. The Discovery Science project in Japan in which more than sixty scientists participated was a three-year project sponsored by Grant-in-Aid for Scientific Research on Priority Area from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. This project mainly aimed to (1) develop new methods for knowledge discovery, (2) install network environments for knowledge discovery, and (3) establish Discovery Science as a new area of Computer Science / Artificial Intelligence Study.

In order to attain these aims we set up five groups for studying the following research areas:

(A) Logic for/of Knowledge Discovery
(B) Knowledge Discovery by Inference/Reasoning
(C) Knowledge Discovery Based on Computational Learning Theory
(D) Knowledge Discovery in Huge Database and Data Mining
(E) Knowledge Discovery in Network Environments

These research areas and related topics can be regarded as a preliminary definition of Discovery Science by enumeration. Thus Discovery Science ranges over philosophy, logic, reasoning, computational learning and system developments.

In addition to these five research groups we organized a steering group for planning, adjustment and evaluation of the project. The steering group, chaired by the principal investigator of the project, consists of leaders of the five research groups and their subgroups as well as advisors from the outside of the project. We invited three scientists to consider the Discovery Science overlooking the above five research areas from viewpoints of knowledge science, natural language processing, and image processing, respectively.

The group A studied discovery from a very broad perspective, taking into account of historical and social aspects of discovery, and computational and logical aspects of discovery. The group B focused on the role of inference/reasoning in knowledge discovery, and obtained many results on both theory and practice on statistical abduction, inductive logic programming and inductive inference. The group C aimed to propose and develop computational models and methodologies for knowledge discovery mainly based on computational learning theory. This group obtained some deep theoretical results on boosting of learning algorithms and the minimax strategy for Gaussian density estimation, and also methodologies specialized to concrete problems such as algorithm for finding best subsequence patterns, biological sequence compression algorithm, text categorization, and MDL-based compression. The group D aimed to create computational strategy for speeding up the discovery process in total. For this purpose,

试读结束：需要全本请在线购买： www.ertongbook.com

the group D was organized with researchers working in scientific domains and researchers from computer science so that real issues in the discovery process can be exposed out and practical computational techniques can be devised and tested for solving these real issues. This group handled many kinds of data: data from national projects such as genomic data and satellite observations, data generated from laboratory experiments, data collected from personal interests such as literature and medical records, data collected in business and marketing areas, and data for proving the efficiency of algorithms such as UCI repository. So many theoretical and practical results were obtained on such a variety of data. The group E aimed to develop a unified media system for knowledge discovery and network agents for knowledge discovery. This group obtained practical results on a new virtual materialization of DB records and scientific computations that help scientists to make a scientific discovery, a convenient visualization interface that treats web data, and an efficient algorithm that extracts important information from semi-structured data in the web space.

This lecture describes an outline of our project and the main results as well as how the project was prepared. We have published and are publishing special issues on our project from several journals [5],[6],[7],[8],[9],[10]. As an activity of the project we organized and sponsored Discovery Science Conference for three years where many papers were presented by our members [2],[3],[4]. We also published annual progress reports [1], which were distributed at the DS conferences. We are publishing the final technical report as an LNAI[11].

References

1. S. Arikawa, M. Sato, T. Sato, A. Maruoka, S. Miyano, and Y. Kanada. Discovery Science Progress Report No.1 (1998), No.2 (1999), No.3 (2000). *Department of Informatics, Kyushu University.*
2. S. Arikawa and H. Motoda. Discovery Science. *LNAI, Springer* 1532, 1998.
3. S. Arikawa and K. Furukawa. Discovery Science. *LNAI, Springer* 1721, 1999.
4. S. Arikawa and S. Morishita. Discovery Science. *LNAI, Springer* 1967, 2000.
5. H. Motoda and S. Arikawa (Eds.) Special Feature on Discovery Science. *New Generation Computing,* 18(1): 13–86, 2000.
6. S. Miyano (Ed.) Special Issue on Surveys on Discovery Science. *IEICE Transactions on Information and Systems,* E83-D(1): 1–70, 2000.
7. H. Motoda (Ed.) Special Issue on Discovery Science. *Journal of Japanese Society for Artificial Intelligence,* 15(4):592–702, 2000.
8. S. Morishita and S. Miyano(Eds.) Discovery Science and Data Mining (in Japanese). *bit special volume , Kyoritsu Shuppan,* 2000.
9. S. Arikawa, M. Sato, T. Sato, A. Maruoka, S. Miyano, and Y. Kanada. The Discovery Science Project. *Journal of Japanese Society for Artificial Intelligence,* 15(4) 595–607, 2000.
10. S. Arikawa, H. Motoda, K. Furukawa, and S. Morishita (Eds.) Theoretical Aspects of Discovery Science. *Theoretical Computer Science* (to appear)
11. S. Arikawa and A. Shinohara (Eds.) Progresses in Discovery Science. *LNAI, Springer* (2001, to appear)

Discovering Mechanisms: A Computational Philosophy of Science Perspective

Lindley Darden

Department of Philosophy
University of Maryland
College Park, MD 20742
darden@carnap.umd.edu
http://www.inform.umd.edu/PHIL/faculty/LDarden/

Abstract. A task in the philosophy of discovery is to find reasoning strategies for discovery, which fall into three categories: strategies for generation, evaluation and revision. Because mechanisms are often what is discovered in biology, a new characterization of mechanism aids in their discovery. A computational system for discovering mechanisms is sketched, consisting of a simulator, a library of mechanism schemas and components, and a discoverer for generating, evaluating and revising proposed mechanism schemas. Revisions go through stages from how possibly to how plausibly to how actually.

1 Introduction

Philosophers of discovery look for reasoning strategies that can guide discovery. This work is in the framework of Herbert Simon's (1997) view of discovery as problem solving. Given a problem to be solved, such as explaining a phenomenon, one goal is to find a mechanism that produces that phenomenon. For example, given the phenomenon of the production of a protein, the goal is to find the mechanism of protein synthesis. The task of the philosopher of discovery is to find reasoning strategies to guide such discoveries. Strategies are heuristics for problem solving; that is, they provide guidance but do not guarantee success.

Discovery is not viewed as something that occurs in a single a-ha moment of insight. Instead, discovery is construed as a process that occurs over an extended period of time, going through cycles of generation, evaluation, and revision (Darden 1991).

The history of science is a source of "compiled hindsight" (Darden 1987) about reasoning strategies for discovering mechanisms. This paper will use examples from the history of biology to illustrate general reasoning strategies for discovering mechanisms. Section 2 puts this work into the broader context of a matrix of biological knowledge. Section 3 discusses a new characterization of mechanism, based on an ontology of entities, properties, and activities. Section 4 outlines components of a mechanism discovery system, including a simulator, a library of mechanism designs and components, and a discoverer.

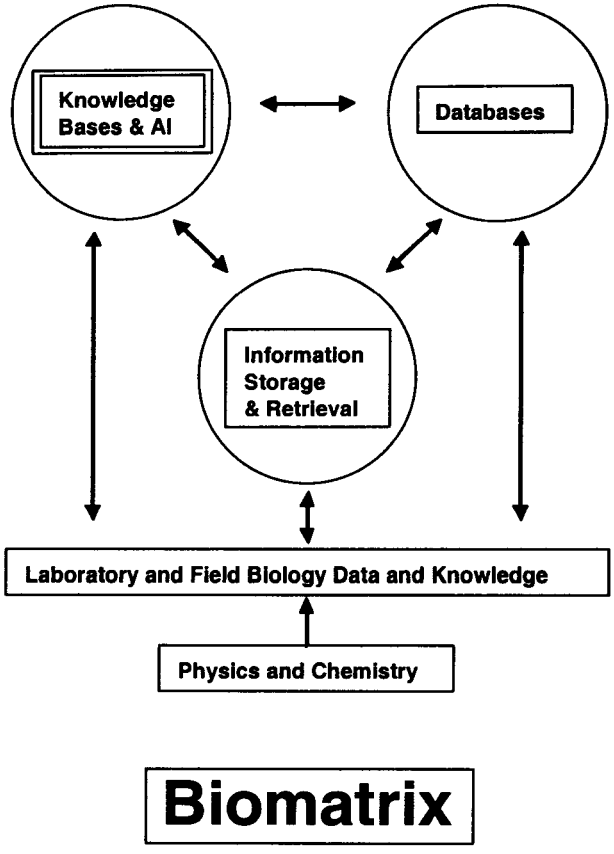

Fig. 1. Matrix of Biological Knowledge

2 Biomatrix

This work is situated in a larger framework. In the 1980s, Harold Morowitz (1985) chaired a National Academy of Sciences workshop on models in biology. As a result of that workshop, a society was formed with the name, "Biomatrix: A Society for Biological Computation and Informatics" (Morowitz and Smith 1987). This society was ahead of its time; it has splintered into different groups and its grand vision has yet to be realized. Nonetheless, its vision is worth revisiting in order to put the work to be discussed in this paper into a broader context. As Figure 1 shows, the biomatrix vision included relations among three areas: first, databases; second, information storage and retrieval by literature cataloging (e.g., Medline); and, third, artificial intelligence and knowledge bases. Discovery science has worked in all three areas since the 1980s. Knowledge discovery in databases is a booming area (e.g., Piatetsky-Shapiro and Frawley, eds., 1991).

Discovery using abstracts available from literature catalogues has been developed by Don Swanson (1990) and others. The area of discovery using knowledge based systems is an active area, especially in computational biology. The meetings on Intelligent Systems in Molecular Biology and the International Society for Computational Biology arose from that part of the biomatrix. It is in the knowledge based systems box that my work today will fall. Relations to databases and information retrieval as related to mechanism discovery will perhaps occur to the reader.

3 Mechanisms, Schemas, and Sketches

Often in biology, what is to be discovered is a mechanism. Physicists often aim to discover general laws, such as Newton's laws of motion. However, few biological phenomena are best characterized by universal, mathematical laws (Beatty 1995). The field of molecular biology, for example, studies mechanisms, such as the mechanisms of DNA replication, protein synthesis, and gene regulation. The lively area of functional genomics is now attempting to discover the mechanisms in which the gene sequences act. Such mechanisms include gene expression, during both embryological development and normal gene activities in the adult. The field of biochemistry also studies mechanisms when it find the activities that transform one stage in a pathway to next, such as the enzymes, reactants and products in the Krebs cycle that produces the energy molecule ATP. An important current scientific task is to connect genetic mechanisms studied by molecular biology with metabolic mechanisms studied by biochemistry. As that task is accomplished, science will have a unified picture of the mechanisms that carry out the two essential features of life according to Aristotle: reproduction and nutrition.

Given this importance of mechanisms in biology, a correspondingly important task for discovery science is to find methods for discovering mechanisms. If the goal is to discover a mechanism, then the nature of that product shapes the process of discovery. A new characterization of mechanism aids the search for reasoning strategies to discover mechanisms.

A mechanism is sought to explain how a *phenomenon* is produced (Machamer, Darden, Craver 2000) or how some *task* is carried out (Bechtel and Richardson 1993) or how the mechanism as a whole *behaves* (Glennan 1996). Mechanisms may be characterized in the following way:

> Mechanisms are entities and activities organized such that they are productive of regular changes from start or set-up to finish or termination conditions. (Machamer, Darden, Craver 2000, p. 3).

Mechanisms are regular in that they usually work in the same way under the same conditions. The regularity is exhibited in the typical way that the mechanism runs from beginning to end; what makes it regular is the *productive continuity* between stages. Mechanisms exhibit productive continuity without gaps

from the set up to the termination conditions; that is, each stage gives rise to, allows, drives, or makes the next.

The ontology proposed here consists of entities, properties, and activities. Mechanisms are composed of both *entities* (with their properties) and *activities*. Activities are the producers of change. Entities are the things that engage in activities. Activities require that entities have specific types of properties. For example, two entities, a DNA base and its complement, engage in the activity of hydrogen bonding because of their properties of geometric shape and weak polar charges.

For a given scientific field, there are typically entities and activities that are accepted as relatively fundamental or taken to be unproblematic for the purposes of a given scientist, research group, or field. That is, descriptions of mechanisms in that field typically bottom out somewhere. Bottoming out is relative: different types of entities and activities are where a given field stops when constructing its descriptions of mechanisms. In molecular biology, mechanisms typically bottom out in descriptions of the activities of cell organelles, such as the ribosome, and molecules, including macromolecules, smaller molecules, and ions. The most important kinds of activities in molecular biology are geometrico-mechanical and electro-chemical activities. An example of a geometrico-mechanical activity is the lock and key docking of an enzyme and its substrate. Electro-chemical activities include strong covalent bonding and weak hydrogen bonding.

Entities and activities are interdependent (Machamer, Darden, Craver 2000, p. 6). For example, appropriate chemical valences are necessary for covalent bonding. Polar charges are necessary for hydrogen bonding. Appropriate shapes are necessary for lock and key docking. This interdependence of entities and activities allows one to reason about one, based on what is known or conjectured about the other, in each stage of the mechanism (Darden and Craver, in press).

A *mechanism schema* is a truncated abstract description of a mechanism that can be filled with more specific descriptions of component entities and activities. An example is the following:

DNA → RNA → protein.

This is a diagram of the central dogma of molecular biology. It is a very abstract, schematic representation of the mechanism of protein synthesis.

A schema may be even more abstract if it merely indicates functional roles played in the mechanism by fillers of a place in the schema (Craver 2001). Consider the schema

DNA → template → protein.

The schema term "template" indicates the functional role played by the intermediate between DNA and protein. Hypotheses about role-fillers changed during the incremental discovery of the mechanism of protein synthesis in the 1950s and 1960s. Thus, mechanism schemes are particularly good ways of representing functional roles. (For discussion of "local" and "integrated" functions and a less schematic way of representing them in a computational system, see Karp 2000.)

Table 1. Constraints on the Organization of Mechanisms

Character of phenomenon

Componency Constraints
- Entities and activities
- Modules

Spatial Constraints
- Compartmentalization
- Localization
- Connectivity
- Structural
- Orientation

Temporal Constraints
- Order
- Rate
- Duration
- Frequency

Hierarchical Constraints
- Integration of levels

(from Craver and Darden 2001)

Mechanism *sketches* are incomplete schemas. They contain black boxes, which cannot yet be filled with known components. Attempts to instantiate a sketch would leave a gap in the productive continuity; that is, knowledge of the needed particular entities and activities is missing. Thus, sketches indicate what needs to be discovered in order to find a mechanism schema.

Once a schema is found and instantiated, a detailed description of a mechanism results. For example, a more detailed description of the protein synthesis mechanism (often depicted in diagrams) satisfies the constraints that any adequate description of a mechanism must satisfy. It shows how the phenomenon, the synthesis of a protein, is carried out by the operation of the mechanism. It depicts the entities–DNA, RNA, and amino acids–as well as implicitly, the activities. Hydrogen bonding is the activity operating when messenger RNA is copied from DNA. There is a geometrico-mechanical docking of the messenger RNA and the ribosome, a particle in the cytoplasm. Hydrogen bonding again occurs as the codons on messenger RNA bond to the anticodons on transfer RNAs carrying amino acids. Finally, covalent bonding is the activity that links the amino acids together in the protein. Good mechanism descriptions show the spatial relations of the components and the temporal order of the stages.

A detailed description of a mechanism satisfies several general constraints. (They are listed in Table 1 and indicated here by italics.) There is a *phenomenon* that the mechanism, when working, produces, for example, the synthesis of a pro-

试读结束：需要全本请在线购买： www.ertongbook.com