Stephen Muggleton
Ramon Otero
Alireza Tamaddoni-Nezhad (Eds.)

LNAI 4455

Inductive Logic Programming

16th International Conference, ILP 2006
Santiago de Compostela, Spain, August 2006
Revised Selected Papers

Springer

Stephen Muggleton Ramon Otero
Alireza Tamaddoni-Nezhad (Eds.)

Inductive Logic Programming

16th International Conference, ILP 2006
Santiago de Compostela, Spain, August 24-27, 2006
Revised Selected Papers

Springer

Series Editors

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

Volume Editors

Stephen Muggleton
Alireza Tamaddoni-Nezhad
Imperial College London
Department of Computing
180 Queen's Gate, London SW7 2AZ, UK
E-mail: {shm,atn}@doc.ic.ac.uk

Ramon Otero
University of Corunna
Department of Computer Science
15071 Coruna, Spain
E-mail: otero@udc.es

Library of Congress Control Number: 2007931449

CR Subject Classification (1998): I.2.3, I.2.6, I.2, D.1.6, F.4.1

LNCS Sublibrary: SL 7 – Artificial Intelligence

ISSN 0302-9743
ISBN-10 3-540-73846-0 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-73846-6 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media

springer.com

Printed in Germany

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India
Printed on acid-free paper SPIN: 12097890 06/3180 5 4 3 2 1 0

Lecture Notes in Artificial Intelligence 4455
Edited by J. G. Carbonell and J. Siekmann

Subseries of Lecture Notes in Computer Science

Lecture Notes in Artificial Intelligence (LNAI)

Vol. 4660: S. Džeroski, J. Todoroski (Eds.), Computational Discovery of Scientific Knowledge. X, 327 pages. 2007.

Vol. 4617: V. Torra, Y. Narukawa, Y. Yoshida (Eds.), Modeling Decisions for Artificial Intelligence. XII, 502 pages. 2007.

Vol. 4612: I. Miguel, W. Ruml (Eds.), Abstraction, Reformulation, and Approximation. XI, 418 pages. 2007.

Vol. 4604: U. Priss, S. Polovina, R. Hill (Eds.), Conceptual Structures: Knowledge Architectures for Smart Applications. XII, 514 pages. 2007.

Vol. 4603: F. Pfenning (Ed.), Automated Deduction – CADE-21. XII, 522 pages. 2007.

Vol. 4597: P. Perner (Ed.), Advances in Data Mining. XI, 353 pages. 2007.

Vol. 4594: R. Bellazzi, A. Abu-Hanna, J. Hunter (Eds.), Artificial Intelligence in Medicine. XVI, 509 pages. 2007.

Vol. 4585: M. Kryszkiewicz, J.F. Peters, H. Rybinski, A. Skowron (Eds.), Rough Sets and Intelligent Systems Paradigms. XIX, 836 pages. 2007.

Vol. 4578: F. Masulli, S. Mitra, G. Pasi (Eds.), Applications of Fuzzy Sets Theory. XVIII, 693 pages. 2007.

Vol. 4573: M. Kauers, M. Kerber, R. Miner, W. Windsteiger (Eds.), Towards Mechanized Mathematical Assistants. XIII, 407 pages. 2007.

Vol. 4571: P. Perner (Ed.), Machine Learning and Data Mining in Pattern Recognition. XIV, 913 pages. 2007.

Vol. 4570: H.G. Okuno, M. Ali (Eds.), New Trends in Applied Artificial Intelligence. XXI, 1194 pages. 2007.

Vol. 4565: D.D. Schmorrow, L.M. Reeves (Eds.), Foundations of Augmented Cognition. XIX, 450 pages. 2007.

Vol. 4562: D. Harris (Ed.), Engineering Psychology and Cognitive Ergonomics. XXIII, 879 pages. 2007.

Vol. 4548: N. Olivetti (Ed.), Automated Reasoning with Analytic Tableaux and Related Methods. X, 245 pages. 2007.

Vol. 4539: N.H. Bshouty, C. Gentile (Eds.), Learning Theory. XII, 634 pages. 2007.

Vol. 4529: P. Melin, O. Castillo, L.T. Aguilar, J. Kacprzyk, W. Pedrycz (Eds.), Foundations of Fuzzy Logic and Soft Computing. XIX, 830 pages. 2007.

Vol. 4511: C. Conati, K. McCoy, G. Paliouras (Eds.), User Modeling 2007. XVI, 487 pages. 2007.

Vol. 4509: Z. Kobti, D. Wu (Eds.), Advances in Artificial Intelligence. XII, 552 pages. 2007.

Vol. 4496: N.T. Nguyen, A. Grzech, R.J. Howlett, L.C. Jain (Eds.), Agent and Multi-Agent Systems: Technologies and Applications. XXI, 1046 pages. 2007.

Vol. 4483: C. Baral, G. Brewka, J. Schlipf (Eds.), Logic Programming and Nonmonotonic Reasoning. IX, 327 pages. 2007.

Vol. 4482: A. An, J. Stefanowski, S. Ramanna, C.J. Butz, W. Pedrycz, G. Wang (Eds.), Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. XIV, 585 pages. 2007.

Vol. 4481: J. Yao, P. Lingras, W.-Z. Wu, M. Szczuka, N.J. Cercone, D. Ślęzak (Eds.), Rough Sets and Knowledge Technology. XIV, 576 pages. 2007.

Vol. 4476: V. Gorodetsky, C. Zhang, V.A. Skormin, L. Cao (Eds.), Autonomous Intelligent Systems: Multi-Agents and Data Mining. XIII, 323 pages. 2007.

Vol. 4455: S. Muggleton, R. Otero, A. Tamaddoni-Nezhad (Eds.), Inductive Logic Programming. XII, 456 pages. 2007.

Vol. 4452: M. Fasli, O. Shehory (Eds.), Agent-Mediated Electronic Commerce. VIII, 249 pages. 2007.

Vol. 4451: T.S. Huang, A. Nijholt, M. Pantic, A. Pentland (Eds.), Artifical Intelligence for Human Computing. XVI, 359 pages. 2007.

Vol. 4438: L. Maicher, A. Sigel, L.M. Garshol (Eds.), Leveraging the Semantics of Topic Maps. X, 257 pages. 2007.

Vol. 4429: R. Lu, J.H. Siekmann, C. Ullrich (Eds.), Cognitive Systems. X, 161 pages. 2007.

Vol. 4426: Z.-H. Zhou, H. Li, Q. Yang (Eds.), Advances in Knowledge Discovery and Data Mining. XXV, 1161 pages. 2007.

Vol. 4411: R.H. Bordini, M. Dastani, J. Dix, A.E.F. Seghrouchni (Eds.), Programming Multi-Agent Systems. XIV, 249 pages. 2007.

Vol. 4410: A. Branco (Ed.), Anaphora: Analysis, Algorithms and Applications. X, 191 pages. 2007.

Vol. 4399: T. Kovacs, X. Llorà, K. Takadama, P.L. Lanzi, W. Stolzmann, S.W. Wilson (Eds.), Learning Classifier Systems. XII, 345 pages. 2007.

Vol. 4390: S.O. Kuznetsov, S. Schmidt (Eds.), Formal Concept Analysis. X, 329 pages. 2007.

Vol. 4389: D. Weyns, H.V.D. Parunak, F. Michel (Eds.), Environments for Multi-Agent Systems III. X, 273 pages. 2007.

Vol. 4384: T. Washio, K. Satoh, H. Takeda, A. Inokuchi (Eds.), New Frontiers in Artificial Intelligence. IX, 401 pages. 2007.

Vol. 4371: K. Inoue, K. Satoh, F. Toni (Eds.), Computational Logic in Multi-Agent Systems. X, 315 pages. 2007.

Vol. 4369: M. Umeda, A. Wolf, O. Bartenstein, U. Geske, D. Seipel, O. Takata (Eds.), Declarative Programming for Knowledge Management. X, 229 pages. 2006.

Vol. 4342: H. de Swart, E. Orłowska, G. Schmidt, M. Roubens (Eds.), Theory and Applications of Relational Structures as Knowledge Instruments II. X, 373 pages. 2006.

Vol. 4335: S.A. Brueckner, S. Hassas, M. Jelasity, D. Yamins (Eds.), Engineering Self-Organising Systems. XII, 212 pages. 2007.

Vol. 4334: B. Beckert, R. Hähnle, P.H. Schmitt (Eds.), Verification of Object-Oriented Software. XXIX, 658 pages. 2007.

Vol. 4333: U. Reimer, D. Karagiannis (Eds.), Practical Aspects of Knowledge Management. XII, 338 pages. 2006.

Vol. 4327: M. Baldoni, U. Endriss (Eds.), Declarative Agent Languages and Technologies IV. VIII, 257 pages. 2006.

Vol. 4314: C. Freksa, M. Kohlhase, K. Schill (Eds.), KI 2006: Advances in Artificial Intelligence. XII, 458 pages. 2007.

Vol. 4304: A. Sattar, B.-h. Kang (Eds.), AI 2006: Advances in Artificial Intelligence. XXVII, 1303 pages. 2006.

Vol. 4303: A. Hoffmann, B.-h. Kang, D. Richards, S. Tsumoto (Eds.), Advances in Knowledge Acquisition and Management. XI, 259 pages. 2006.

Vol. 4293: A. Gelbukh, C.A. Reyes-Garcia (Eds.), MICAI 2006: Advances in Artificial Intelligence. XXVIII, 1232 pages. 2006.

Vol. 4289: M. Ackermann, B. Berendt, M. Grobelnik, A. Hotho, D. Mladenič, G. Semeraro, M. Spiliopoulou, G. Stumme, V. Svátek, M. van Someren (Eds.), Semantics, Web and Mining. X, 197 pages. 2006.

Vol. 4285: Y. Matsumoto, R.W. Sproat, K.-F. Wong, M. Zhang (Eds.), Computer Processing of Oriental Languages. XVII, 544 pages. 2006.

Vol. 4274: Q. Huo, B. Ma, E.-S. Chng, H. Li (Eds.), Chinese Spoken Language Processing. XXIV, 805 pages. 2006.

Vol. 4265: L. Todorovski, N. Lavrač, K.P. Jantke (Eds.), Discovery Science. XIV, 384 pages. 2006.

Vol. 4264: J.L. Balcázar, P.M. Long, F. Stephan (Eds.), Algorithmic Learning Theory. XIII, 393 pages. 2006.

Vol. 4259: S. Greco, Y. Hata, S. Hirano, M. Inuiguchi, S. Miyamoto, H.S. Nguyen, R. Słowiński (Eds.), Rough Sets and Current Trends in Computing. XXII, 951 pages. 2006.

Vol. 4253: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part III. XXXII, 1301 pages. 2006.

Vol. 4252: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part II. XXXIII, 1335 pages. 2006.

Vol. 4251: B. Gabrys, R.J. Howlett, L.C. Jain (Eds.), Knowledge-Based Intelligent Information and Engineering Systems, Part I. LXVI, 1297 pages. 2006.

Vol. 4248: S. Staab, V. Svátek (Eds.), Managing Knowledge in a World of Networks. XIV, 400 pages. 2006.

Vol. 4246: M. Hermann, A. Voronkov (Eds.), Logic for Programming, Artificial Intelligence, and Reasoning. XIII, 588 pages. 2006.

Vol. 4223: L. Wang, L. Jiao, G. Shi, X. Li, J. Liu (Eds.), Fuzzy Systems and Knowledge Discovery. XXVIII, 1335 pages. 2006.

Vol. 4213: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), Knowledge Discovery in Databases: PKDD 2006. XXII, 660 pages. 2006.

Vol. 4212: J. Fürnkranz, T. Scheffer, M. Spiliopoulou (Eds.), Machine Learning: ECML 2006. XXIII, 851 pages. 2006.

Vol. 4211: P. Vogt, Y. Sugita, E. Tuci, C.L. Nehaniv (Eds.), Symbol Grounding and Beyond. VIII, 237 pages. 2006.

Vol. 4203: F. Esposito, Z.W. Raś, D. Malerba, G. Semeraro (Eds.), Foundations of Intelligent Systems. XVIII, 767 pages. 2006.

Vol. 4201: Y. Sakakibara, S. Kobayashi, K. Sato, T. Nishino, E. Tomita (Eds.), Grammatical Inference: Algorithms and Applications. XII, 359 pages. 2006.

Vol. 4200: I.F.C. Smith (Ed.), Intelligent Computing in Engineering and Architecture. XIII, 692 pages. 2006.

Vol. 4198: O. Nasraoui, O. Zaïane, M. Spiliopoulou, B. Mobasher, B. Masand, P.S. Yu (Eds.), Advances in Web Mining and Web Usage Analysis. IX, 177 pages. 2006.

Vol. 4196: K. Fischer, I.J. Timm, E. André, N. Zhong (Eds.), Multiagent System Technologies. X, 185 pages. 2006.

Vol. 4188: P. Sojka, I. Kopeček, K. Pala (Eds.), Text, Speech and Dialogue. XV, 721 pages. 2006.

Vol. 4183: J. Euzenat, J. Domingue (Eds.), Artificial Intelligence: Methodology, Systems, and Applications. XIII, 291 pages. 2006.

Vol. 4180: M. Kohlhase, OMDoc – An Open Markup Format for Mathematical Documents [version 1.2]. XIX, 428 pages. 2006.

Vol. 4177: R. Marín, E. Onaindía, A. Bugarín, J. Santos (Eds.), Current Topics in Artificial Intelligence. XV, 482 pages. 2006.

Vol. 4160: M. Fisher, W. van der Hoek, B. Konev, A. Lisitsa (Eds.), Logics in Artificial Intelligence. XII, 516 pages. 2006.

Vol. 4155: O. Stock, M. Schaerf (Eds.), Reasoning, Action and Interaction in AI Theories and Systems. XVIII, 343 pages. 2006.

Vol. 4149: M. Klusch, M. Rovatsos, T.R. Payne (Eds.), Cooperative Information Agents X. XII, 477 pages. 2006.

Vol. 4140: J.S. Sichman, H. Coelho, S.O. Rezende (Eds.), Advances in Artificial Intelligence - IBERAMIA-SBIA 2006. XXIII, 635 pages. 2006.

Preface

The inherent dangers of change are often summed up in the misquoted Chinese curse "May you live in interesting times." The submission procedure for the 16th International Conference of Inductive Logic Programming (ILP 2006) was a radical (hopefully interesting but not cursed) departure from previous years. Submissions were requested in two phases. The first phase involved submission of short papers (three pages) which were then presented at the conference and included in a short papers proceedings. In the second phase, reviewers selected papers for long paper submission (15 pages maximum). These were then assessed by the same reviewers, who then decided which papers to include in the journal special issue and proceedings. In the first phase there were a record 77 papers, compared to the usual 20 or so long papers of previous years. Each paper was reviewed by three reviewers. Out of these, 71 contributors were invited to submit long papers. Out of the long paper submissions, 7 were selected for the *Machine Learning Journal* special issue and 27 were accepted for the proceedings. In addition, two papers were nominated by Program Committee referees for the applications prize and two for the theory prize. The papers represent the diversity and vitality in present ILP research including ILP theory, implementations, search and phase transition, distributed and large-scale learning, probabilistic ILP, biological applications, natural language learning and planning and action learning.

ILP 2006 was held in Santiago de Compostela under the auspices of the University of Corunna and the University of Santiago de Compostela. The annual meeting of ILP researchers acts as the premier forum for presenting the latest work in the field. In addition to the many technical paper presentations, the invited talks this year were given by some of the most distinguished names in artificial intelligence research, namely, Vladimir Lifschitz, John McCarthy, Stuart Russell, Bart Selman and Ehud Shapiro.

We gratefully acknowledge support of the PASCAL network of excellence, the Spanish National Commissions of Science and Technology, the Galicia-Spain Secretary of R&D, the University of Corunna, Imperial College London, the University of Santiago de Compostela, the Spanish Association of AI and the *Machine Learning Journal*. Finally we would like to thank the many individuals involved in the preparation of the conference. These include the Journal Special Issue organizer (Simon Colton), the Local Chair (David Losada), the Local Organizers (Jorge Gonzalez and Miguel Varela) as well as Bridget Gundry, who organized and distributed the conference poster.

March 2007

Stephen Muggleton
Ramon Otero
Alireza Tamaddoni-Nezhad

Organization

Organizing Committee

Program Chair:	Stephen Muggleton (Imperial College, UK)
Program Chair:	Ramon Otero (University of Corunna, Spain)
Special Issue Organizer:	Simon Colton (Imperial College, UK)
Proceedings Organizer:	Alireza Tamaddoni-Nezhad (Imperial College, UK)
Local Organization Chair:	David Losada (University of Santiago, Spain)
Local Organizer:	Jorge Gonzalez (University of Corunna, Spain)
Local Organizer:	Miguel Varela (University of Corunna, Spain)

Program Committee

Hendrik Blockeel, Belgium
Rui Camacho, Portugal
James Cussens, UK
Luc Dehaspe, Belgium
Luc De Raedt, Germany
Saso Dzeroski, Slovenia
Floriana Esposito, Italy
Peter Flach, UK
Tamas Horvath, Germany
Katsumi Inoue, Japan
Andreas Karwath, Germany
Roni Khardon, USA
Joerg-Uwe Kietz, Switzerland
Ross King, UK
Stefan Kramer, Germany
Nada Lavrac, Slovenia
Francesca Lisi, Italy
John Lloyd, Australia
Donato Malerba, Italy
Stephen Muggleton, UK
Ramon Otero, Spain
David Page, USA
Bernhard Pfahringer, New Zealand
Jan Ramon, Belgium
Celine Rouveirol, France
Michele Sebag, France
Jude Shavlik, USA
Takayoshi Shoudai, Japan
Arno Siebes, Netherlands
Ashwin Srinivasan, India
Tomoyuki Uchida, Japan
Lyle Ungar, USA
Christel Vrain, France
Stefan Wrobel, Germany
Akihiro Yamamoto, Japan
Mohammed Zaki, USA
Gerson Zaverucha, Brazil
Filip Zelezny, Czech Republic
Jean-Daniel Zucker, France

Invited Speakers

Vladimir Lifschitz	University of Texas at Austin, USA
John McCarthy	Stanford University, USA
Stuart Russell	University of California at Berkeley, USA
Bart Selman	Cornell University, USA
Ehud Shapiro	Weizmann Institute, Israel

Sponsoring Institutions

PASCAL European Network of Excellence
Consellería de Innovación, Industria e Comercio, Xunta de Galicia
Consellería de Educación e Ordenación Universitaria, Xunta de Galicia
Universidade da Coruña
Imperial College London
Universidade de Santiago de Compostela
Asociación Española de Inteligencia Artificial (AEPIA)
Machine Learning Journal

Additional Referees

Annalisa Appice
Gilles Bisson
Janez Brank
Damjan Demsar
Aloisio De Pina
Nicola Di Mauro
Anton Dries
Stefano Ferilli
Kristian Kersting
Jiri Klema
Yoshiaki Okubo
Aline Paes
Andrea Passerini
Joel Plisson
Chiaki Sakama
Vitor Santos Costa
Leander Schietgat
Ivica Slavkov
Jan Struyf
Anneleen Van Assche
Takeaki Uno

Table of Contents

Invited Papers

Special Issue Extended Abstracts

Research Papers

Actions, Causation and Logic Programming

Vladimir Lifschitz

Department of Computer Sciences
University of Texas at Austin, USA

Reasoning about changes caused by the execution of actions has long been at the center of attention of researchers in the area of logic-based AI. Logical properties of causal dependencies turned out to be similar to properties of rules in logic programs. This fact allows us to apply methods of logic programming to computational problems related to action and change. Ideas of answer set programming, based on the concept of a stable model, turned out to be particularly useful. In the past they have been applied primarily to the problem of plan generation. There is now increasing interest also in using logic programming for learning action descriptions.

S. Muggleton, R. Otero, and A. Tamaddoni-Nezhad (Eds.): ILP 2006, LNAI 4455, p. 1, 2007.

Challenges to Machine Learning: Relations Between Reality and Appearance

John McCarthy

Stanford University, USA

Abstract. Machine learning research, e.g. as described in [4], has as its goal the discovery of relations among observations, i.e. appearances. This is inadequate for science, because there is a reality behind appearance, e.g. material objects are built up from atoms. Atoms are just as real as dogs, only harder to observe, and the atomic theory arose long before there was any idea of how big atoms were. This article discusses how atoms were discovered, as an example of discovering the reality behind appearance. We also present an example of the three-dimensional reality behind a two-dimensional appearance, and how that reality is inferred by people and might be inferred by computer programs. Unfortunately, it is necessary to discuss the philosophy of appearance and reality, because the mistaken philosophy of taking the world (or particular phenomena) as a structure of sense data has been harmful in artificial intelligence and machine learning research, just as behaviorism and logical positivism harmed psychology.

1 Introduction

Apology: My knowledge of of machine learning research is no more recent than Tom Mitchell's book [4]. Its chapters describe, except for inductive logic programming, programs solely aimed at classifying appearances.

We live in a complicated world that existed for billions of years before there were humans, and our sense organs give us limited opportunities to observe it directly. Four centuries of science tell us that we and the objects we perceive are built in a complicated way from atoms and, below atoms, quarks. Maybe there is something below quarks.

Science, since 1700, is far better established than any kind of philosophy. Bad philosophy, proposing to base research entirely on appearances, has stunted AI, just as behaviorism stunted psychology for many decades.

Here's the philosophy in a nutshell. As emphasized by Descartes, all a human's information comes through the senses. Therefore, it is tempting to try to base science on relations among sense data and relations between actions that may be performed and subsequent sense data. [6] is an important source for this approach. Unfortunately for this approach, humans and our environment are complicated structures built of vastly smaller objects that our senses do not directly observe. Science had to discover atoms.

Besides the fundamental realities behind appearance studied by science, there are hidden every day realities—the three dimensional reality behind two

S. Muggleton, R. Otero, and A. Tamaddoni-Nezhad (Eds.): ILP 2006, LNAI 4455, pp. 2–9, 2007.

dimensional images, hidden surfaces, objects in boxes, people's names, what people really think of us.

Human common sense also reasons in terms of the realities that give rise to the appearances our senses provide us. Thus young babies have some initial knowledge of the permanence of physical objects. This initial knowledge seems not to be expressed in terms of particular senses. Blind babies have it too, and so do babies whose sense of touch is compromised by lack of arms. See [7] for experiments related to initial knowledge.

Perhaps if your philosophy rejects the notion of reality as a fundamental concept, you'll accept a notion of *relative reality* appropriate for the design and debugging of robots. Thus the robot needs to be designed to determine this relative reality from the appearance given by its inputs.

We'll discuss:

Dalton's atomic theory as a discovery of the reality behind appearance.

A simple problem involving changeable two dimensional appearances and a three dimensional reality.

Some formulas relating appearance and reality in particular cases.

What can one know about a three dimensional object and how to represent this knowledge.

The use of touch in finding the shape of an object.

How scientific study and the use of instruments extends what can be learned from the senses. Thus a doctor's training involving dissection of cadavers enables him to determine something about the liver by palpation.

2 Elements, Atoms, and Molecules

Some scientific discoveries like Galileo's $s = \frac{1}{2}gt2$ involve discovering the relations between known entities. Patrick Langley's Bacon program [1] did that.

John Dalton's postulation of atoms and molecules made up of fixed numbers of atoms of two or more kinds was much more creative and will be harder to make computers do. That's the reason for this section of the paper.

The ancient ideas of Democritus and Lucretius that matter was made up from atoms had no important or even testable consequences. Dalton's did.

Giving each kind of atom its own atomic mass explained the complicated ratios of masses in a compound as representing small numbers of atoms in a molecule. Thus a sodium chloride (NaCl) molecule would have one atom of each of its elements. Water came out as H_2O.

The simplest forms of the atomic theory were inaccurate. [Thus early 19th century chemists didn't soon realize that the hydrogen and oxygen molecules are H_2 and O_2 and not just H and O.] Computers also need to be able to propose theories adventurously and fix their inaccuracies later later.

Only the relative masses of atoms could be proposed in Dalton's time. The first actual way of estimating these masses was made by Maxwell and Boltzmann about 60 years after Dalton's proposal. They realized that the coefficients of viscosity, heat conductivity, and diffusion of gases as explained by the kinetic theory of gases depended on the actual sizes of molecules.

The last important scientific holdout against the reality of atoms, the chemist Wilhelm Ostwald, was convinced by Einstein's 1905 quantitative explanation of Brownian motion as caused by liquid molecules striking a suspended object. The philosopher Ernst Mach was unconvinced.

Long after the reality of atoms was accepted in science, it was still believed that individual atoms could not be observed. The first actual pictures of atoms in the 1990s were a big surprise. Now quarks are accepted as real although an actual picture of a proton showing the quarks would be even more surprising and seems quite unlikely, because the quarks move too fast.

Philosophical point: Atoms cannot be regarded as just an explanation of the observations that led Dalton to propose them. Maxwell and Boltzmann used the notion to explain entirely different observations, and modern explanations of atoms are not at all based on the law of combining proportions. In short, atoms were discovered, not invented.

Reality is usually more stable than appearance, i.e. changes more slowly. Formulas giving the effects of events (including actions) are almost always written in terms of reality. Getting reality from appearance is an *inverse problem.* Geologists, oil companies, and astronomers are faced with inverse problems. Their solution is intellectually difficult and computationally intensive. Human-level AI systems will also have to be able to infer reality from appearances related to them in complex ways.

3 Elements, Atoms, Molecules - Formulas

Most likely, it is still too hard to make programs that will discover elements, atoms, and molecules. Let's therefore try to write logical sentences that will introduce these concepts to a knowledge base that has no ideas of them.

We assume that the notions of a body being composed of parts and of mass have already been formalized, but the idea of atom has not. The ideas of bodies being disjoint is also assumed to be formalized.

The following formulas approximate a fragment of high school chemistry and should be somewhat *elaboration tolerant* [2], e.g. should admit additional information about the structure of molecules. The situation argument s is included only to point out that material bodies change in chemical reactions.

$$\begin{aligned}
&Body(b,s) \rightarrow (\exists u \subset Molecules(b,s))(\forall y \in u)(Molecule(y) \wedge Part(y,b)),\\
&y1 \in Molecules(b,s) \wedge y2 \in Molecules(b,s) \wedge y1 \neq y2 \rightarrow Disjoint(y1,y2),\\
&Part(x,b,s) \rightarrow (\exists y \in Molecules(b,s))\neg Disjoint(y,x),\\
&Body(b,s) \rightarrow Mass(b,s) = \textstyle\sum_{x \in Molecules(b,s)} Mass(x,s).
\end{aligned} \tag{1}$$